Some Asymptotic Results in Discounted Repeated Games of One-Sided Incomplete Information*

Martin W. Cripps
University of Warwick,
Coventry CV4 7AL, UK.

Jonathan P. Thomas
University of St Andrews,
St Andrews KY16 9AL, UK.

This version: May 2000

ABSTRACT: The paper analyzes the Nash equilibria of two-person discounted repeated games with one-sided incomplete information and known own payoffs. If the informed player is arbitrarily patient relative to the uninformed player, then the characterization for the informed player's payoffs 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, and under an assumption that strictly individually rational payoffs exist, 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.

* We are very grateful to two anonymous referees for numerous helpful comments. We also thank seminar participants at Erasmus University, Carlos III Madrid, Cambridge, Keele, London School of Economics, Notttingham, Warwick, Essex, University College London, Nuffield College Oxford, Southampton, York, and also to Mike Peters, Carlo Perroni and Marco Celentani for their comments and suggestions. The current version is based on a paper entitled "The Folk Theorem in Repeated Games of Incomplete Information" presented at the 1996 summer workshop at MEDS Northwestern University, and the 1995 Fields Institute Conference Toronto.

1 Introduction

In this paper we consider discounted non-zero-sum repeated games between two players with one-sided incomplete information and known own payoffs. We shall investigate equilibrium payoffs as the players become patient. We consider two cases concerning relative discount factors. Our first main result, in Section 3, states that for arbitrary given initial beliefs, for a fixed value of the uninformed player's (player 2) discount factor, and if the informed player's (player 1) discount factor is sufficiently close to one, the equilibrium payoffs to player 1 (for each of a finite number of types) must approximately satisfy the conditions of an equilibrium in which the informed player acts to reveal her information at the start of the game. This implies a continuity result[1] 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 Section 4, the symmetric discounting case is analysed. Under an assumption on the existence of strictly individually rational payoffs, we establish a continuity result with complete information games as the probability of one of the types goes to one: for any degree of approximation, provided the players are sufficiently patient and *provided initial beliefs put sufficiently high probability on this type,* then given any feasible strictly individually rational payoff vector in the game between this type and player 2, there is a Nash equilibrium of the incomplete information game with approximately these payoffs (to this type of player 1 and to player 2). 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, and notably it is *not* the limit of the discounted case if both players' discount factors are equal.

The situation where one or more players' preferences may be unknown to the opponent(s) has received relatively little attention in the non-zero-sum discounted repeated games literature, despite considerable work on 'reputation' models where perturbations of preferences are in terms of irrational or commitment types. *Undiscounted* repeated games

[1]This continuity property is not uniform with respect to initial beliefs.

of incomplete information have, however, been studied in some depth (see Section 3). Some recent results exist for the discounted case, however. Kalai and Lehrer (1993) and Jordan (1995) have established that play must converge to Nash play of the true game. Jordan (1995) has also proved the existence of an equilibrium for this class of games. Perfect Bayesian equilibria of such games must have a Markov property (Bergin (1989)). 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 necessary to know something about how rapidly convergence takes place relative to the rate of discounting of payoffs and also, possibly, what happens in the shorter run. By exploiting a result due to Fudenberg and Levine (1992) on the speed of learning (see also Sorin (1999) for a synthesis of a number of the results in this literature) 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. A more detailed consideration of the shorter run is needed for the symmetric discounting case as the speed of learning is crucial.[2] Finally, in a recent paper, equilibrium payoffs in discounted repeated *zero-sum* games with incomplete information have been studied by Lehrer and Yariv (1999), who show that as both players become infinitely and equally patient the equilibrium payoffs converge to those with no discounting, whereas if the informed player is infinitely more patient than the uninformed an example is given to show that this is not true.

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 $\Re^K$), and informed to

[2]This contrast is why the characterization for the case of a relatively patient informed player holds for all priors which assign positive probability to all types: equilibria are shown to be approximately equivalent in terms of player 1's payoffs to an equilibrium where information is revealed at the start of play; prior beliefs are unimportant for such equilibria. In the symmetric discounting case, where the speed of learning matters, priors play an important role and they determine the characterization of equilibrium payoffs. In this case we only provide a characterization for priors putting almost all weight on a particular type.

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, while player 2 simultaneously chooses an action j^t from the finite set J, where I and J have at least two elements. 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 $\delta_i \in (0, 1)$, with the payoff to type k of player 1 being $\tilde{a}_k = (1 - \delta_1) \sum_{t=0}^{\infty} \delta_1^t A_k(i^t, j^t)$, and that to player 2 being $\tilde{b} = (1 - \delta_2) \sum_{t=0}^{\infty} \delta_2^t B(i^t, j^t)$. 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 (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 strategy for player 2 is a sequence of maps $\tau = (\tau^0, \tau^1, \cdots)$, $\tau^t : H^{t-1} \to \Delta^J$. 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 maximize expected payoffs, and a Nash equilibrium is defined as a pair of strategies (σ, τ) such that, for each k, $E_{\mathbf{p},\sigma,\tau}[\tilde{a}_k \mid k] \geq E_{\mathbf{p},\sigma',\tau}[\tilde{a}_k \mid k]$ for all σ', and $E_{\mathbf{p},\sigma,\tau}[\tilde{b}] \geq E_{\mathbf{p},\sigma,\tau'}[\tilde{b}]$ for all τ'. Finally we shall need the following. Let $\hat{a}_k := \min_{g \in \Delta^J} \max_{f \in \Delta^I} A_k(f, g)$ 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 player 2's minmax payoff is given by $\hat{b} := \min_{f \in \Delta^I} \max_{g \in \Delta^J} B(f, g)$.

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 (asymptotically) satisfied in the discounted case as $\delta_1 \to 1$. These necessary conditions can be interpreted as requiring payoff equivalence to some fully revealing equilibrium.

Hart (1985) gave a complete characterization for the general class of undiscounted games (payoffs evaluated according to a Banach limit) with one-sided incomplete information, which includes the possibility that the uninformed player is unaware of his own payoff function. For the case we are interested in, namely "known own payoffs" but where one of the players does not know the payoffs of the other player, a simpler characterization has been provided by Shalev (1994) (see also Koren (1988), and Forges (1992) for a survey of the literature.) Denote this game by $\Gamma(\mathbf{p}, 1, 1)$. 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. Punishment strategies for player 2 are more complex than in the complete information setting, because all possible types of player 1 must *simultaneously* be punished. Let $\mathbf{x} := (x_k)_{k\in K} \in \Re^K$ be a vector of payoffs for the types of player 1. For $\mathbf{q} \in \Delta^K$, let $a(\mathbf{q})$ be player 1's minmax payoff in the one-shot game with payoffs given by $\sum_{k\in K} q_k A_k(i,j)$.The set of payoffs $\{\mathbf{y} \in \Re^K | \mathbf{y} \leq \mathbf{x}\}$ is said to be *approachable* by player 2 if and only if

$$(1) \qquad \mathbf{q}\cdot\mathbf{x} \geq a(\mathbf{q}) \quad \text{for all } \mathbf{q} \in \Delta^K.$$

Blackwell's approachability result (Blackwell (1956)) then implies that player 2 has a strategy, τ, that guarantees type k gets average (i.e., undiscounted) payoffs of no more than x_k whatever strategy, σ, player 1 uses. Thus if the set $\{\mathbf{y}|\mathbf{y} \leq \mathbf{x}\}$ is approachable then $\mathbf{x}$ is a vector of feasible punishment payoffs for player 2 to impose on the types of player 1. We will say that 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. 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

$$(2) \qquad y \geq \hat{b}.$$

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 $A_k(\pi) = \sum_{i\in I, j\in J} \pi^{ij} A_k(i,j)$ and $B(\pi) = \sum_{i\in I, j\in J} \pi^{ij} B(i,j)$ respectively. 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 $k, k' \in K$.*

Shalev (1994) showed that 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 first main result needed is Lemma 2, which states that equilibrium play between type k and player 2, as summarised in the average (using player 1's discount factor in the weighted average) frequencies over action profiles, must approximately satisfy the individual rationality condition of Definition 1 *for player 2*. Its proof depends on two main ideas. First (Lemma 1), if player 2's equilibrium strategy gives him less than $\hat{b}$ when he plays against k, then he must anticipate that the probability distribution over outcomes if he is facing type k's strategy differs from the one generated by the "expected" equilibrium strategy of player 1 (averaging 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 idea (Result 1) states 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 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 is playing type k's strategy, and he will choose a response which gives him at least his minmax payoff. For a fixed value of δ_2, Result 1 implies an upper bound on how long this learning takes. Consequently if a sufficiently high discount factor (i.e. δ_1 as opposed to δ_2) is used to evaluate player 2's payoffs, this learning phase will be insignificant and player 2 must get approximately his minmax payoff against type k.

For a fixed equilibrium, we define the average frequencies over action profiles conditional on type k when the discount factor is δ as: $\pi_k^{ij}(\delta) = (1-\delta) E\left[\sum_{t=0}^{\infty} \delta^t \mathbf{1}\{i,j,t\} \middle| k\right]$, for each i and j, where $\mathbf{1}\{i,j,t\}$ is the indicator function for the action profile (i,j) occurring at date t. It is easy to check that the equilibrium payoffs are $E\left[\tilde{a}_k \mid k\right] = A_k(\pi_k(\delta_1))$ for each k and $E\left[\tilde{b}\right] = \sum_{k \in K} p_k B(\pi_k(\delta_2))$. Let $b_{min} = \min_{i \in I} \min_{j \in J} B(i,j)$ be the worst payoff player 2 can get in the stage game. 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 $y^N = \left((i^{t+1}, j^{t+1}), \ldots, (i^{t+N}, j^{t+N})\right)$. For given equilibrium strategies (σ, τ) we let $\mathbf{q}^N(\cdot \mid h^t)$ be the distribution over these outcomes (i.e., $\mathbf{q}^N(y^N \mid h^t)$ =prob$[h^{t+N} = (h^t, y^N) \mid h^t]$, using obvious notation) and likewise $\mathbf{q}^N(\cdot \mid h^t, k)$ 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 $\mathbf{q}^N$ and $\hat{\mathbf{q}}^N$, $\parallel \mathbf{q}^N - \hat{\mathbf{q}}^N \parallel \; := \; \max_{y^N} \left|\mathbf{q}^N(y^N) - \hat{\mathbf{q}}^N(y^N)\right|$. Finally, define the continuation payoff for player 1 type k, discounted to period $t+1$, as: $\tilde{a}_k^{t+1} := (1-\delta_1) \sum_{r=t+1}^{\infty} \delta_1^{r-t-1} A_k(i^r, j^r)$, and that for player 2 as $\tilde{b}^{t+1} := (1-\delta_2) \sum_{r=t+1}^{\infty} \delta_2^{r-t-1} B(i^r, j^r)$.

Lemma 1 *Let $\delta_2 \in (0,1)$ 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*

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

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

$$\parallel \mathbf{q}^N(\cdot \mid h^t) - \mathbf{q}^N(\cdot \mid h^t, k) \parallel \; > \; \eta \; . \tag{4}$$

<u>Proof:</u> Straightforward.

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 $T(n, N) = \{n, n+N, n+2N, \ldots\}$. The result is a straightforward adaptation of the main theorem of Fudenberg and Levine (1992, Theorem 4.1) which is stated for the case $N = 1$.

Result 1 (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

$$\| \mathbf{q}^N(\cdot \mid h^t) - \mathbf{q}^N(\cdot \mid h^t, k) \| > \psi \tag{5}$$

is less than ξ.

Lemma 2 states that equilibrium play between type k and player 2, as summarised in the average (using player 1's discount factor in the weighted average) frequencies over action profiles, must approximately satisfy the individual rationality condition of Definition 1 *for player 2* (see Cripps *et al.* (1996) for a related argument in the 'reputation' context).

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*

$$B(\pi_k(\delta_1)) \geq \hat{b} - \phi \ . \tag{6}$$

<u>PROOF:</u> 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 (4) holds whenever (3) holds. Set $\psi = \eta$ in Result 1, 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 1 there is an m (finite) such that the probability that inequality (4) holds more than m times in $T(n, N)$ is less than ξ, so the probability that inequality (3) 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

$$E\left[\tilde{b}^{t+1} \mid h^t, k\right] \leq \hat{b} - \phi/3 \tag{7}$$

holds more than Nm times is smaller than $N\xi = \frac{\phi}{3(\hat{b} - b_{min})}$. Next, $E\left[\tilde{b}^{t+1} \mid k\right] = E\left[(1-\delta_2)B(i^{t+1}, j^{t+1}) + \delta_2 \tilde{b}^{t+2} \mid k\right]$, so $(1-\delta_2)E\left[B(i^{t+1}, j^{t+1}) \mid k\right] = E\left[\tilde{b}^{t+1} - \delta_2 \tilde{b}^{t+2} \mid k\right]$. Hence, player 2's payoff against type k in the equilibrium, calculated using player 1's discount factor, is

$$B(\pi_k(\delta_1)) \ = \ (1-\delta_1) \sum_{t=0}^{\infty} \delta_1^t E\left[B(i^t, j^t) \mid k\right]$$

$$= \frac{1-\delta_1}{1-\delta_2}\sum_{t=0}^{\infty}\delta_1^t E\left[\tilde{b}^t - \delta_2\tilde{b}^{t+1} \mid k\right]$$

(8) $$= \frac{1-\delta_1}{1-\delta_2}\left\{E\left[\tilde{b}^0 \mid k\right] + E\left[\sum_{t=0}^{\infty} E\left[\delta_1^t(\delta_1-\delta_2)\tilde{b}^{t+1}\Big| h^t, k\right]\Bigg| k\right]\right\} .$$

Using the result on the number of times (7) holds, for $\delta_1 > \delta_2$ the random variable $\sum_{t=0}^{\infty} E\left[\delta_1^t(\delta_1-\delta_2)\tilde{b}^{t+1}\Big| h^t, k\right] \geq \left\{\frac{\delta_1-\delta_2}{1-\delta_1}(\hat{b}-\frac{\phi}{3}) - (\delta_1-\delta_2)(\hat{b}-b_{min})Nm\right\}$ with probability at least $(1-N\xi)$ conditional on k, where we are using the fact that in the event that (7) 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 (8) 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 $\delta_1 \to 1$ yields $\lim_{\delta_1\to 1}\Phi(\delta_1,\delta_2) = (1-N\xi)\left(\hat{b}-\frac{\phi}{3}\right) + N\xi b_{min}$; hence, since $N\xi = \frac{\phi}{3(\hat{b}-b_{min})}$, we get

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

(9) $$= \hat{b}-\frac{2\phi}{3}+\frac{\phi^2}{9(\hat{b}-b_{min})} > \hat{b}-\frac{2\phi}{3} .$$

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 $\delta_1 \geq \underline{\delta}_1^{(k)}$, $B(\pi_k(\delta_1)) \geq \hat{b}-\phi$. Set $\underline{\delta}_1 = \max_{k\in K}\{\underline{\delta}_1^{(k)}\}$ and the result follows. Q.E.D.

We are now in a position to establish that Shalev's equilibrium characterization holds approximately as a necessary condition provided that player 1 is sufficiently patient relative to player 2. 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. 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

(10) $$\mathbf{A}^* = \{(A_1(\pi_1), A_2(\pi_2), \ldots, A_K(\pi_K)) : (\pi_k)_{k\in K} \in \Pi_0\} .$$

We can state

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 player 1 has equilibrium payoffs $\mathbf{a}$ in $\Gamma(\mathbf{p}, \delta_1, \delta_2)$, then*

$$min_{\mathbf{x} \in \mathbf{A}^*} \parallel \mathbf{a} - \mathbf{x} \parallel < \epsilon \ . \tag{11}$$

<u>PROOF:</u> 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 (1), there exists $\mathbf{q}^* \in \Delta^K$ such that $\mathbf{q}^* \cdot \mathbf{a} < a(\mathbf{q}^*)$. By the minimax theorem,

$$\mathbf{q}^* \cdot \mathbf{a} < \max_{f \in \Delta^I} \min_{g \in \Delta^J} \sum_k q_k^* A_k(f, g) \ , \tag{12}$$

so that if player 1 plays a mixed action f^* which attains the maximum in (12), $\mathbf{q}^* \cdot \mathbf{a} < \sum_k q_k^* A_k(f^*, g)$ 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 $E_{\mathbf{p},\sigma^*,\tau}\left[(1-\delta_1)\sum_{t=0}^{\infty}\delta_1^t \sum_k q_k^* A_k(i^t, j^t)\right] > \mathbf{q}^* \cdot \mathbf{a}$ (NB. k is not a random variable), so that

$$\sum_k q_k^* E_{\mathbf{p},\sigma^*,\tau}\left[\tilde{a}_k \mid k\right] = \sum_k q_k^* E_{\mathbf{p},\sigma^*,\tau}\left[(1-\delta_1)\sum_{t=0}^{\infty}\delta_1^t A_k(i^t, j^t) \mid k\right] > \mathbf{q}^* \cdot \mathbf{a} \ , \tag{13}$$

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 $E_{\mathbf{p},\sigma^*,\tau}\left[\tilde{a}_k \mid k\right] > a_k$ for at least one k, contradicting the definition of equilibrium. Hence individual rationality must be satisfied for player 1 for any value of δ_1; that is, $\mathbf{a}$ satisfies (1). 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

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

and define the compact valued correspondence

$$\Psi : [0, \infty) \to\to \Pi by \Psi(\phi) = \left\{(\pi_k)_{k \in K} \mid B_k(\pi_k) \geq \hat{b} - \phi \text{ for all } k \in K\right\} .$$

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 $\mathbf{A}((\pi_k)_{k \in K}) := (A_1(\pi_1), A_2(\pi_2), \ldots, A_K(\pi_K))$ 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 1 to be $\bar{\phi}$; the corresponding $\underline{\delta}_1$ is therefore as required for (11) to hold. *Q.E.D.*

Theorem 1 developed necessary conditions which equilibrium payoffs must satisfy asymptotically. In the undiscounted model, the necessary conditions were also *sufficient*. A similar result can be established with discounting provided the inequalities in the conditions of Definition 1 are assumed to hold strictly (as they will at any interior point of Shalev's set). 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.

Theorem 2 *Suppose that* $(\pi_k)_{k \in K} \in \Pi_0$ *satisfies* (i) *:* $(A_k(\pi_k))_{k \in K}$ *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) *:* $A_k(\pi_k) > A_k(\pi_{k'})$ *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* $|\sum_{k \in K} p_k B(\pi_k) - b| < \epsilon$*.*

The proof is straightforward and is omitted; it follows closely the argument for the undiscounted case given in Koren (1988) which constructs 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 Section 4 for a discussion of Blackwell punishment strategies with discounting.

4 Symmetric Discounting

In this section we consider games where the two players are equally patient. We denote this class of games by $\Gamma(\mathbf{p}, \delta)$, so $\Gamma(\mathbf{p}, \delta) := \Gamma(\mathbf{p}, \delta, \delta)$. We show, in a sense to be made more precise, that the (Nash) Folk Theorem for complete information games is robust to small perturbations in the information structure; specifically it can be extended to the repeated games $\Gamma(\mathbf{p}, \delta)$ when p_1 is close to one. In the previous section, by contrast, the characterization was valid for all values of $\mathbf{p}$. (For symmetric discounting, it is easy to construct examples in which the Folk Theorem characterization fails when p_1 is not close to one.) In the repeated game of complete information played between, say, type 1 of player 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, given an assumption on the existence of strictly individually rational payoffs, that there exists $\delta_\nu, p_1^\nu < 1$ such that the pair (a_1, b) can be approximately sustained as equilibrium payoffs in $\Gamma(\mathbf{p}, \delta)$ if $\delta > \delta_\nu$ and $p_1 > p_1^\nu$. 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, and equally, patient.

4.0.1 Example

To illustrate what is to come, we consider an example, where $2 > c \geq 1$ (which satisfies (A.1) below provided $c > 1$). In this example, Shalev's (1994) results (discussed in

	L	R
T	3, 1	0, 0
B	0, 0	1, 3

(A_1, B)

	L	R
T	c, 1	1, 0
B	0, 0	0, 3

(A_2, B)

Section 3) imply that for $c < 2$, there is a lower bound on type 1's equilibrium payoff in the *undiscounted* case strictly above her minmax payoff of 3/4; individual rationality for type 2 and for player 2 ($A_2(\pi_2) \geq 1, B(\pi_2) \geq 3/4$), together with incentive compatibility, implies $A_1(\pi_1) \geq A_1(\pi_2) \geq 3(c+2)/4(3c-2)$. (This is clearest for the case where $c=1$, since $A_2(\pi_2) \geq 1$ then implies $\pi_2(T,L) + \pi_2(T,R) = 1$ and $\pi_2(T,L) \geq 3/4$, so that $A_1(\pi_1) \geq 9/4$). Here we show in the symmetric discounting case that as $\delta \to 1$, a_1 can be driven down to 3/4.

Let $\epsilon > 0$ be given. Consider first the following (pooling) equilibrium of $\Gamma(\mathbf{p}, \delta)$: both types of player 1 play U and player 2 plays L in every period, irrespective of past history. Player 1 gets $(3, c)$ and player 2 gets a payoff of 1 (this plays the role of the equilibrium of Lemma 5). This will be our "terminal equilibrium". Next, precede this equilibrium by the repeated play of (T, R) by both types and by player 2 ((T,R) plays the role of $\underline{\pi}_2$ in Lemma 6, and is played to reduce type 1's payoff; note that $\mathbf{z}$ as defined there, using this $\underline{\pi}_2$, *is* individually rational: player 2 need only play a punishment strategy which minmaxes type 1; the "finite sequence" of Lemma 7 is just a single play of (T,R)). Punishments in all earlier periods involve player 2 being minmaxed thereafter for observable deviations, and type 1 being minmaxed for observable deviations by player 1 (so type 2 gets $(3+c)/4$ after any observable deviation); in the general proof we shall need to vary the punishment with type 1's payoff. The constraint that limits the length of the phase where (T,R) is played in such a pooling equilibrium concerns player 2's individual rationality. Thus (T,R) is played out N times before the above terminal equilibrium is played, where N is the *largest* integer satisfying $(1-\delta^N)0 + \delta^N 1 \geq (1-\delta)3 + \delta 3/4$ (the LHS is player 2's payoff from the strategy specified, and he can get at most 1 in the period of deviation and is minmaxed thereafter). When δ is close to 1, δ^N is close to 3/4, so player 2's payoff is also close to 3/4 : there exists $\delta^*(\epsilon) < 1$ such that for $\delta > \delta^*(\epsilon)$, player 2's payoff δ^N is within $\epsilon/3$ of 3/4,[3] and thus type 1's payoff $\delta^N 3$ is no more than ϵ above 9/4. Payoffs to type 1 and player 2 at this (pooling) equilibrium are shown by point C in Figure 1.

To reduce type 1's payoff further, we introduce a randomization by type 1 in the first period of this equilibrium: suppose that type 1 plays B with probability q such that $p_1 q = 0.5$, where p_1 is player 2's prior at the start of the period (so that from player 2's point of view B is played with probability 0.5). If B is played, so that player 1 signals she

[3] The continuation payoff received by player 2 at any date can change between consecutive dates by at most $2M(1-\delta) < \epsilon/6$ for $\delta > 1 - \epsilon/12M = 1 - \epsilon/36$; likewise the RHS of the inequality defining N given above is within $\epsilon/6$ of 3/4 if $\delta > 1 - \frac{9}{24}\epsilon$; on the other hand δ^N cannot be below 3/4 or else 2's constraint would be violated. Consequently for $\delta > \delta^*(\epsilon) := 1 - \epsilon/36$, $\delta^N \in [3/4, 3/4 + \epsilon/3]$.

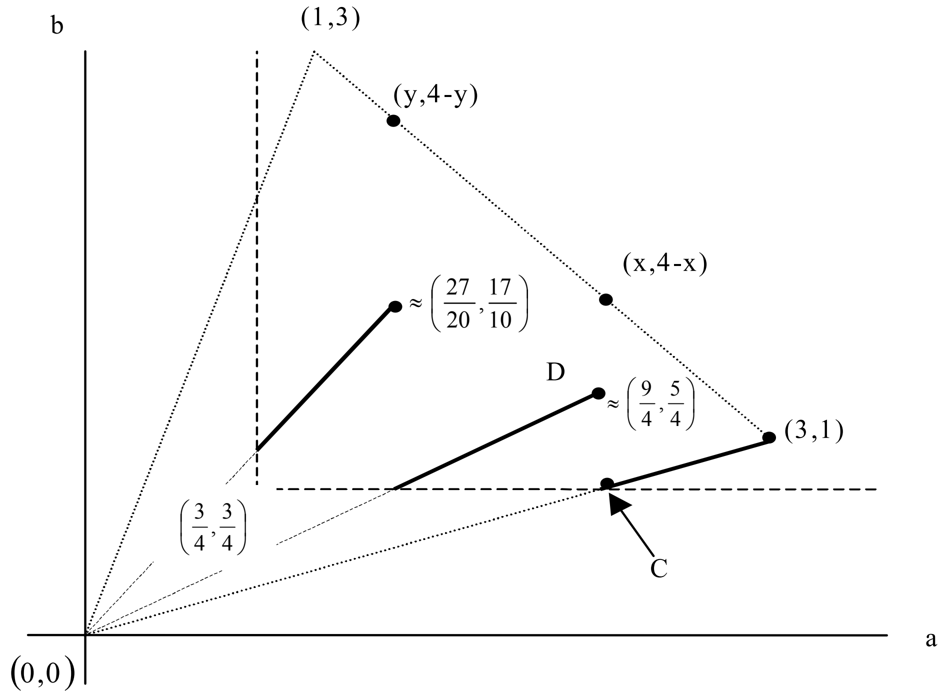

Figure 1: Payoffs to type 1 and player 2

is type 1, then from the start of the following period an equilibrium of $\Phi_1(\delta)$ is played in which, to ensure type 1's indifference, the payoff to type 1, say x, satisfies $(1-\delta)1+\delta x = \delta^N 3$, and player 2 gets $4-x$ (on the frontier of feasible set). Consequently payoffs at this equilibrium to type 1 and player 2 are $(3\delta^N, (\delta^N+(1-\delta)3+\delta(4-x))/2) = (3\delta^N, 2-\delta^N)$, after substitution for x. The purpose of the randomization is to increase the payoff that player 2 receives so as to relax his incentive compatibility constraint, thus allowing further plays of (T,R). The equilibrium just described (see point D in the figure) now replaces the initially described pooling equilibrium in a repetition of the argument. N' rounds of (T,R) are added at the start until again player 2's individual rationality constraint binds: $(1-\delta^{N'})0+\delta^{N'}(2-\delta^N) \geq (1-\delta)3+\delta 3/4$. Repeating the argument given earlier, for $\delta > \delta^*(\epsilon)$, $\delta^{N'}(2-\delta^N)$ is within $\epsilon/3$ of $3/4$. Again add an initial randomization of say q' of playing B by type 1 so that $p'_1 q' = 0.5$, where p'_1 is player 2's prior at the start of the period, and an equilibrium of the complete information game played by type 1 and player 2, which we denote $\Phi_1(\delta)$, with payoffs $(y, 4-y)$ to follow. Payoffs are then $(3\delta^{N+N'}, (\delta^{N'}(2-\delta^N)+(1-\delta)3+\delta(4-y))/2) = (3\delta^{N+N'}, 2-2\delta^{N+N'}+\delta^{N'})$, which for

$\epsilon < \widetilde{\epsilon}$ for some $\widetilde{\epsilon} > 0$ and $\delta > \delta^*(\epsilon)$, lie above the 45^o line being sufficiently close to $(27/20, 17/10)$.[4] A further repetition of the argument, so that more plays of (T, R) are appended at the beginning, then implies that the payoff of type 1 will reach 3/4 before that of player 2 does, so that the latter constraint no longer prevents type 1 receiving a low payoff. By choosing $\epsilon < \widetilde{\epsilon}$ small enough, type 1 can be held as close to 3/4 as desired provided $\delta > \delta^*(\epsilon)$. Observable deviations cannot be optimal as all continuation payoffs are above punishment levels: this is clear for type 1 and for player 2; type 2 gets a continuation payoff of $(1 - \delta^n)1 + \delta^n c$ where there are n periods to go before the final pooling equilibrium, and $\delta^n \geq 1/4$ by type 1's individual rationality, whereas deviation yields at most $\delta(3+c)/4$. We also need to check that type 2 cannot benefit from mimicking type 1 revealing her type; since the equilibria with payoffs $(x, 4-x)$ and $(y, 4-y)$ involve play fluctuating between (T, L) and (B, R) with less weight on the former than on the equilibrium path, mimicking cannot be profitable. As there were two randomizations (at each of which the total probability of player 1 revealing herself to be type 1 is 1/2), the strategies above are an equilibrium of $\Gamma(\mathbf{p}, \delta)$ provided $p_1 \geq 3/4$. To obtain higher payoffs to type 1, it is only necessary to stop the above process earlier; to obtain arbitrary payoffs to player 2, we append an initial randomization by type 1, as described earlier, but in which the equilibrium of $\Phi_1(\delta)$ gives player 2 close to the desired payoffs. Provided type 1's probability is sufficiently close to 1, this will provide any desired degree of approximation.

In what follows, we shall split the above constructions into two steps, first ignoring type $k = 2$ and constructing the equilibrium as an equilibrium of a complete information game, before introducing the possibility of a second type. Finally we deal with more than two types.

4.1 An Equilibrium of the Complete Information Game

The first step in our argument is the construction of an equilibrium of the complete information game played by type 1 and player 2, $\Phi_1(\delta)$. In Lemma 4 we construct a particular type of equilibrium where any feasible and strictly individually rational payoff

[4] Specifically, given that $\delta^N \in [3/4, 3/4 + \epsilon/3]$, and $\delta^{N'} \cdot (2 - \delta^N) \leq 3/4 + \epsilon/3$, it follows that $\delta^{N'} \leq 3/5 + 32\epsilon/5(15 - 4\epsilon) \equiv 3/5 + \Delta$. Thus type 1's payoff $\delta^{N+N'} \cdot 3 \leq 3(3/4 + \epsilon/3)(3/5 + \Delta)$, while player 2's payoff $2 - 2\delta^{N+N'} + \delta^{N'} \geq 17/10 - \Delta$, and thus there exists $\widetilde{\epsilon} > 0$ such that for $\epsilon < \widetilde{\epsilon}$ payoffs lie above the 45^o line.

to type 1 can be obtained as an equilibrium payoff. In Section 4.2 we shall use these equilibrium strategies to construct an equilibrium of a two-type incomplete information game.

Some additional notation is now necessary. 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: $G_k(\epsilon) := \{(A_k(\pi), B(\pi)) | A_k(\pi) \geq \hat{a}_k + \epsilon, B(\pi) \geq \hat{b} + \epsilon, \pi \in \Delta^{IJ}\}$, $k \in K$. Next define $\bar{a}_k(\epsilon)$ to be the largest payoff to type k in $G_k(\epsilon)$ and $\underline{a}_k(\epsilon)$ to be the smallest such payoff, that is $\bar{a}_k(\epsilon) := \max\{ a_k \mid (a_k, b) \in G_k(\epsilon) \}$, and $\underline{a}_k(\epsilon) := \min\{ a_k \mid (a_k, b) \in G_k(\epsilon) \}$. Also define $\bar{\mathbf{a}} := (\bar{a}_1, ..., \bar{a}_K) \in \Re^K$, where $\bar{a}_k = \bar{a}_k(0)$. We also use M to denote an upper bound on the absolute magnitude of the players' payoffs, so that $M \geq |A_k(i,j)|$, $|B(i,j)|$, for all (i,j), k. Define the function f, where $f : [\underline{a}_1(0), \bar{a}_1(0)] \rightarrow \Re$, to be the maximum feasible payoff to player 2 given that type 1 receives the payoff a_1, that is, $f(a_1) := \max\{ b \mid (a_1, b) \in G_1(0) \}$. The function $f(.)$ is made up of a finite number of linear segments. Define S to be the maximum absolute value of the slopes of these segments (this is finite).

We start with two preliminary results. The first is an approximation result which allows correlated strategies to be approximated by average behaviour along deterministic sequences of action profiles.

Result 2 *Let $\epsilon > 0$ be given. There is a $\hat{\delta}(\epsilon) < 1$ such that if $\delta > \hat{\delta}(\epsilon)$ and given any correlated strategy $\pi \in \Delta^{IJ}$, then there exists a sequence of actions $\{(i^t, j^t)\}_{t=0}^{\infty}$ such that: $A_k(\pi) = (1-\delta)\sum_{t=0}^{\infty} \delta^t A_k(i^t, j^t)$, for all $k \in K$, and $B(\pi) = (1-\delta)\sum_{t=0}^{\infty} \delta^t B(i^t, j^t)$; moreover*

$$\left| (1-\delta)\sum_{t=s}^{\infty} \delta^{t-s} A_k(i^t, j^t) - A_k(\pi) \right| \leq \epsilon/2 \qquad s = 0, 1, 2, ..., \forall k \in K,$$

$$\left| (1-\delta)\sum_{t=s}^{\infty} \delta^{t-s} B(i^t, j^t) - B(\pi) \right| \leq \epsilon/2 \qquad s = 0, 1, 2, ... \,.$$

The proof of Result 2 can be adapted from the proof of Lemma 2 in Fudenberg and Maskin (1991). It follows immediately that given $\epsilon > 0$, there is a $\underline{\delta}(\epsilon) \geq \hat{\delta}(\epsilon)$ such that $(a_k, b) \in G_k(\epsilon)$ are equilibrium payoffs for any $\delta > \underline{\delta}(\epsilon)$.

Next, consider the following strategies, which will be used to construct an equilibrium in which a single randomization occurs. The proof of Lemma 4 will require the iteration of this construction. Take $\epsilon > 0$ to be given and also a sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ and an arbitrary $(a_1^*, b^*) \in G_1(3\epsilon)$ and $(x, f(x)) \in G_1(2\epsilon)$. Assume that $\delta > \underline{\delta}(\epsilon)$.

Type 1 : In period 0 play $\hat{\imath}^0$ with probability 1/2 and $\tilde{\imath} \neq \hat{\imath}^0$ with probability 1/2. If $(\hat{\imath}^0, \hat{\jmath}^0)$ is played in period zero, continue to play the sequence $\{\hat{\imath}^t\}_{t=0}^{T-1}$ n times and then in period nT begin playing the equilibrium strategy to get the payoffs $(a_1^*, b^*) \in G_1(3\epsilon)$. If $(\tilde{\imath}, \hat{\jmath}^0)$ is played in period zero, play the infinite sequence of stage-game actions, determined by Result 2, to get the payoffs $(x, f(x)) \in G_1(2\epsilon)$. (Both payoffs are equilibrium payoffs by the assumption that $\delta > \underline{\delta}(\epsilon)$.) Minmax all deviations by player 2.

Player 2 : In period 0 play $\hat{\jmath}^0$. If $(\hat{\imath}^0, \hat{\jmath}^0)$ is played in period zero continue to play the sequence $\{\hat{\jmath}^t\}_{t=0}^{T-1}$ n times and then in period nT begin playing the equilibrium strategy to get the payoffs $(a_1^*, b^*) \in G_1(3\epsilon)$. If $(\tilde{\imath}, \hat{\jmath}^0)$ is played in period zero play the infinite sequence of stage-game actions, determined by Result 2, to get the payoffs $(x, f(x)) \in G_1(2\epsilon)$. Minmax all deviations by player 1.

Call the strategies defined above $\hat{\sigma}(n; a_1^*, b^*, x)$ for type 1 and $\hat{\tau}(n; a_1^*, b^*, x)$ for player 2 (we suppress the implicit dependence of the continuation equilibria on δ). Also define the strategies $\hat{\hat{\sigma}}(n; a_1^*, b^*)$ for type 1 and $\hat{\hat{\tau}}(n; a_1^*, b^*)$ for player 2, which are the same as $\hat{\sigma}(n; a_1^*, b^*, x)$ and $\hat{\tau}(n; a_1^*, b^*, x)$ except that they do not involve a randomization in period 0, that is, type 1 always plays $\hat{\imath}^0$ in period zero. Define payoffs when there are n complete rounds of the sequence to be played as $a_1(n) := (1 - \delta^{nT})\hat{A}_1 + \delta^{nT} a_1^*$ and $b(n) := (1 - \delta^{nT})\hat{B} + \delta^{nT} b^*$. We will now establish the following result.

Lemma 3 *Let $\epsilon > 0$ be given; also let $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ and $\delta^*(\epsilon) < 1$ be so that $\hat{A}_1 := ((1-\delta)/(1-\delta^T)) \sum_{s=0}^{T-1} \delta^s A_1(\hat{\imath}^s, \hat{\jmath}^s) < \hat{a}_1 + \epsilon$ for $1 > \delta > \delta^*(\epsilon)$, and let $(a_1^*, b^*) \in G_1(3\epsilon)$ with $\underline{a}_1(2\epsilon) + \epsilon < a_1^* < \bar{a}_1(2\epsilon) - \epsilon/2$, also be given. If $\delta > \max\{\underline{\delta}(\epsilon), \delta^*(\epsilon), [4M/(\epsilon + 4M)]^{1/T}\}$ then (i) there exists $(x, f(x)) \in G_1(2\epsilon)$ so that $(\hat{\sigma}(n; a_1^*, b^*, x), \hat{\tau}(n; a_1^*, b^*, x))$ is an equilibrium of $\Phi_1(\delta)$, where $n \geq 1$ is the largest integer satisfying*

$$b(n) > \hat{b} + 2\epsilon, \tag{14}$$

(15) $$a_1(n) > \underline{a}_1(2\epsilon) + \epsilon/2;$$

(ii) $(\hat{\hat{\sigma}}(n; a_1^*, b), \hat{\hat{\tau}}(n; a_1^*, b))$ *is an equilibrium of* $\Phi_1(\delta)$.

Proof: We will first show that $n \geq 1$. We have

$$\begin{aligned} a_1(1) - \underline{a}_1(2\epsilon) - \epsilon/2 &= a_1^* - \underline{a}_1(2\epsilon) - \epsilon/2 + (1 - \delta^T)(\hat{A}_1 - a_1^*) \\ &> a_1^* - \underline{a}_1(2\epsilon) - \epsilon/2 - (1 - \delta^T)2M. \end{aligned}$$

By our assumption on a_1^* and $(1 - \delta^T)2M < \epsilon/2$ by the assumption on δ, the bottom line is positive. A similar argument shows $b(1) > \hat{b} + 2\epsilon$.

To prove (i), the strategies are an equilibrium of $\Phi_1(\delta)$ provided: (a) type 1 is indifferent when she randomizes in period zero, and (b) no player prefers to deviate when playing out the sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ n times. Part (ii) follows if (b) holds. Type 1 is indifferent in period zero if we can find an equilibrium with the payoffs $(x, f(x)) \in G_1(2\epsilon)$ where the payoff x satisfies

(16) $$x = \frac{a_1(n)}{\delta} - \frac{(1 - \delta)}{\delta} A_1(\tilde{\imath}, \hat{\jmath}^0).$$

But (16) implies that $|a_1(n) - x| < 2M(1 - \delta)/\delta < \epsilon/2$, where the last inequality follows from our assumptions on δ. This implies $\underline{a}_1(2\epsilon) < x < \bar{a}_1(2\epsilon)$; the lower bound follows as $a_1(n)$ satisfies $\underline{a}_1(2\epsilon)+\epsilon/2 < a_1(n)$, and the upper bound is true since $x \leq a_1^*+\epsilon/2 < \bar{a}_1(2\epsilon)$. So there exists a pair $(x, f(x)) \in G_1(2\epsilon)$ where x satisfies (16).

Type 1's expected payoff from continuing to play the sequence when there are t periods of the current sequence and $n' \leq n$ repetitions of the sequence left to play satisfies

$$\begin{aligned} (1 - \delta) \sum_{s=0}^{t-1} \delta^s A_1(\hat{\imath}^{T-t+s}, \hat{\jmath}^{T-t+s}) + \delta^t a_1(n') &\geq -M(1 - \delta^T) + \delta^T a_1(n) \\ &\geq -M(1 - \delta^T) + \delta^T(\hat{a}_1 + 2\epsilon) \ . \end{aligned}$$

This follows as $a_1(n') \geq a_1(n)$. Type 1's payoff from deviation is bounded above by $(1 - \delta^T)M + \delta^T \hat{a}_1$, so a sufficient condition for deviation not to be profitable, $\delta^T(\epsilon + M) \geq M$, is given in the proposition. An identical argument using the fact that $b(n') \geq \min\{b(n), b^*\}$ shows that player 2 also does not benefit from deviating when they are playing out the sequence n times. *Q.E.D.*

In the next lemma, we start with an equilibrium of $\Phi_1(\delta)$ with payoffs (a_1^*, b^*) close to the maximum feasible and individually rational payoff to type 1 in $G_1(3\epsilon)$. Using this equilibrium we find a new equilibrium with the payoffs $(a_1(n), (b(n) + (1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x))/2)$, where, by construction, $a_1(n) < a_1^*$. If the payoffs at this new equilibrium satisfy $(a_1(n), (b(n) + (1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x))/2) \in G_1(3\epsilon)$ and the condition $\underline{a}_1(2\epsilon) + \epsilon < a_1(n)$ then it is possible to apply the lemma a second time to find a further equilibrium of $\Phi_1(\delta)$ where type 1 receives the payoff $a_1(n+n') < a_1(n) < a_1^*$. Again if this new equilibrium gives payoffs in $G_1(3\epsilon)$ and satisfying the same condition, it will be possible to iterate the lemma a third time, to find further equilibria of $\Phi_1(\delta)$ where type 1 receives even lower payoffs, and so on.

We define $(\hat{\sigma}(N), \hat{\tau}(N))$ to be the strategies that iteratively apply Lemma 3 to the equilibrium with payoffs (a_1^*, b^*) where the sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ is played out in total N times; each iteration uses the strategies $(\hat{\sigma}, \hat{\tau})$ defined above Lemma 3except for the last which uses $(\hat{\hat{\sigma}}, \hat{\hat{\tau}})$ (so there is no initial randomization). (The dependence on δ, ϵ and (a_1^*, b^*) is suppressed.) There is an upper bound on the number of times Lemma 3 can be applied, and hence on N; let N_{max} be this upper bound on N. (We show that the strategies $(\hat{\sigma}(N_{max}), \hat{\tau}(N_{max}))$ will imply that a_1 is close to $\underline{a}_1(4\frac{1}{16}\epsilon)$.) Randomizations by player 1 occur at each new application of Lemma 3.

Lemma 4 *Let $\epsilon > 0$ and $C > 0$ be given and let $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ and $\delta^*(\epsilon) < 1$ be so that $\hat{A}_1 := ((1-\delta)/(1-\delta^T)) \sum_{s=0}^{T-1} \delta^s A_1(\hat{\imath}^s, \hat{\jmath}^s) < \hat{a}_1 + \epsilon$ for $1 > \delta > \delta^*(\epsilon)$. There exists $\underline{r} > 0$ and $\tilde{\delta}(\epsilon) \geq \delta^*(\epsilon)$ such that: given $(a_1^*, b^*) \in G_1(3\epsilon)$ which satisfies $\bar{a}_1(2\epsilon) - \epsilon/2 > a_1^* > \bar{a}_1(3\epsilon) - C\epsilon$, $a_1 \in [\underline{a}_1(4\frac{1}{16}\epsilon) + \epsilon, \bar{a}_1(3\epsilon) - C\epsilon]$, and $\delta > \tilde{\delta}(\epsilon)$, then there exists an N such that $(\hat{\sigma}(N), \hat{\tau}(N))$ is an equilibrium of $\Phi_1(\delta)$ with a payoff to type 1 of $a_1(N)$ within $\epsilon/32$ of a_1, and at this equilibrium type 1 departs from repeated play of the sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ (by playing $\tilde{\imath}$ instead of $\hat{\imath}^0$ at the points of randomisation) with a total probability of at most $1 - \underline{r}$.*

<u>Proof:</u> Let $\tilde{\delta}(\epsilon) := \max\{\underline{\delta}(\epsilon), \delta^*(\epsilon), [32M/(\epsilon + 32M)]^{1/T}, 1 - \epsilon/[32M(S+1)]\}$. This lower bound on δ implies that if x and y are any two feasible payoffs for player i, then

$$(17) \qquad |x - [(1-\delta^T)y + \delta^T x]| = (1-\delta^T)|x-y| < (1-\delta^T)2M < \epsilon/16.$$

We will first show that the payoff to type 1 at the equilibrium $(\hat{\sigma}(N_{max}), \hat{\tau}(N_{max}))$ is no greater than $\underline{a}_1(4\frac{1}{16}\epsilon) + \epsilon$. It is impossible to apply Lemma 3 another time if $a_1(N_{max}) \leq \underline{a}_1(2\epsilon) + \epsilon$, but in this case the result is proved. We will now suppose that $a_1(N_{max}) > \underline{a}_1(2\epsilon) + \epsilon$, which implies that in the last feasible iteration of Lemma 3 the constraint $a_1(n) > \underline{a}_1(2\epsilon) + \epsilon/2$ does not bind (*cf.* the argument in the first paragraph of the proof of Lemma 3). Thus, instead, in the last feasible iteration of Lemma 3 the constraint $b(n) > \hat{b} + 2\epsilon$ binds and Lemma 3 cannot be reapplied because $(a_1(n), (b(n) + (1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x))/2) \notin G_1(3\epsilon)$. There are now two separate cases to consider: (1) If $[b(n) + (1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x)]/2 > \hat{b} + 3\epsilon$, but $(a_1(n), [b(n) + (1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x)]/2) \notin G_1(3\epsilon)$, then it must be that $a_1(n) < \underline{a}_1(3\epsilon)$. (2) If $[b(n)+(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f(x)]/2 < \hat{b}+3\epsilon$, then $b(n) > \hat{b} + 2\epsilon$ implies

$$(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x) < \hat{b} + 4\epsilon. \tag{18}$$

Player 1's equilibrium payoff is $a_1(n) = (1-\delta)A(\tilde{\imath}, \hat{\jmath}^0) + \delta x$, by indifference. The point $(a_1(n), (1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f(x))$ is in the feasible set and is within $\epsilon/16$ of the point $(x, f(x))$, by (17). We know that $f(x) < \hat{b} + 4\frac{1}{16}\epsilon$, from (17) and (18). It therefore follows that $x < \underline{a}_1(4\frac{1}{16}\epsilon)$. This and (17) applied again implies $a_1(n) < \underline{a}_1(4\frac{1}{16}\epsilon) + \frac{1}{16}\epsilon$.

The payoff to type 1 at the equilibrium $(\hat{\sigma}(N_{max}), \hat{\tau}(N_{max}))$ is thus no greater than $\underline{a}_1(4\frac{1}{16}\epsilon) + \epsilon$. Therefore, type 1's payoff at the equilibrium $(\hat{\sigma}(N), \hat{\tau}(N))$ ranges from less than $\underline{a}_1(4\frac{1}{16}\epsilon) + \epsilon$ (for N large) to $a_1^* > \bar{a}_1(3\epsilon) - C\epsilon$ (for $N = 0$). By (17), type 1's payoff at the equilibrium $(\hat{\sigma}(N), \hat{\tau}(N))$ increases by at most $\epsilon/16$ as N increases in integer steps. Thus there must be a value N for which type 1's payoff is within $\epsilon/32$ of any point in $[\underline{a}_1(4\frac{1}{16}\epsilon) + \epsilon, \overline{a}_1(3\epsilon) - C\epsilon]$.

Fix a particular (a_1^*, b^*) satisfying the conditions of the lemma statement and a $\delta > \tilde{\delta}(\epsilon)$. The equilibrium $(\hat{\sigma}(N_{\max}), \hat{\tau}(N_{\max}))$ is well defined, so: there are only a finite number of periods when the sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ is played and there are only a finite number of occasions when type 1 randomizes over the actions $\hat{\imath}^0$ and $\tilde{\imath}$. Thus, there is a strictly positive probability $\underline{r}$ of always playing $\hat{\imath}^0$ and not deviating from the sequence. We now need to prove that the number of randomizations between $n = N_{\max}$ and $n = 0$ is bounded above by a number independent of δ and (a_1^*, b^*). For a given δ and (a_1^*, b^*), at the equilibrium $(\hat{\sigma}(N_{\max}), \hat{\tau}(N_{\max}))$, let $a_1(n)$ and $a_1(n+n')$ be player 1's payoff at two consecutive randomizations (assuming there are at least 2 randomizations). Recall that there is no randomization at the start of the last iteration, so $n + n' < N_{\max}$. We must have

$$b(n+n') = (1-\delta^{n'T})\hat{B} + \delta^{n'T}\frac{1}{2}[b(n) + (1-\delta)B(\tilde{\imath},\tilde{\jmath}^0) + \delta f(x)] > \hat{b} + 2\epsilon, \tag{19}$$

where x is chosen as in (16). (Note: if there are *any* randomizations, then $\hat{B} < \hat{b} + 2\epsilon$.) By definition of there being a randomization at $n + n'$ the inequality in (19) must be violated for $n + n' + 1$ (since the constraint $a_1(n + n') > \underline{a}_1(2\epsilon) + \frac{1}{2}\epsilon$ can only bind — in the sense that additional play of the sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ would lead to its violation — at $n + n' = N_{\max}$), and since δ^T is bounded below by the assumption $\delta > \tilde{\delta}$ there is an upper bound on $\delta^{Tn'}$:

$$\frac{(1+\epsilon/32M)(\hat{b} + 2\epsilon - \hat{B})}{\frac{1}{2}[b(n) + (1-\delta)B(\tilde{\imath},\tilde{\jmath}^0) + \delta f(x)] - \hat{B}} > \delta^{n'T}.$$

But $\hat{A}_1 < \hat{a}_1 + \epsilon$, so the upper bound on $\delta^{n'T}$ gives an upper bound on $a_1(n + n')$:

$$a_1(n+n') - \hat{A}_1 < (a_1(n) - \hat{A}_1)\left\{\frac{(1+\epsilon/32M)(\hat{b} + 2\epsilon - \hat{B})}{\frac{1}{2}[b(n) + (1-\delta)B(\tilde{\imath},\tilde{\jmath}^0) + \delta f(x)] - \hat{B}}\right\}.$$

A sufficient condition for the term in braces to be strictly bounded below unity for all $\delta > \tilde{\delta}(\epsilon)$ is that there exists an $\eta > 0$ such that

$$1 + \frac{\epsilon}{32M} + \eta < \frac{\frac{1}{2}[b(n) + (1-\delta)B(\tilde{\imath},\tilde{\jmath}^0) + \delta f(x)] - \hat{B}}{\hat{b} + 2\epsilon - \hat{B}}, \qquad \forall\, 1 > \delta > \tilde{\delta}(\epsilon). \tag{20}$$

Subtracting unity from each side and then noticing that the denominator of the right is strictly less than $2M$ gives the following sufficient condition

$$\frac{\epsilon}{16} < \frac{1}{2}[b(n) + (1-\delta)B(\tilde{\imath},\tilde{\jmath}^0) + \delta f(x)] - \hat{b} - 2\epsilon, \qquad \forall\, 1 > \delta > \tilde{\delta}(\epsilon).$$

Let $f(a_1) = G + Ha_1$ be the equation of the tangent to the frontier of the feasible set at a_1. Substituting this in for $f(.)$ and then for x from (16) and multiplying by 2:

$$\frac{\epsilon}{8} < G + Ha_1(n) - b(n) - (1-\delta)(G + HA_1(\tilde{\imath},\tilde{\jmath}^0) - B(\tilde{\imath},\tilde{\jmath}^0)) + 2(b(n) - \hat{b} - 2\epsilon). \tag{21}$$

The last term on the right is positive by (14). The second last term equals $(1-\delta)(G + Hx - B(\tilde{\imath},\tilde{\jmath}^0) + H(A_1(\tilde{\imath},\tilde{\jmath}^0) - x)))$, which is bounded above by $(1-\delta)(S+1)2M$, as $|H| \le S$. By the assumption that $\delta > \tilde{\delta}(\epsilon)$, a sufficient condition for (21) and hence for (20) is

$$\frac{\epsilon}{8} < G + Ha_1(n) - b(n) - \frac{\epsilon}{16}.$$

The construction of the strategies ensures that $a_1(n) \leq a_1^* < \bar{a}_1(2\epsilon) - \epsilon/2$, so at $a_1(n)$ the maximum feasible payoff to player 2 is at least $\hat{b} + 2\epsilon$. The line $G + Ha_1$ graphs payoffs to player 2 that are weakly greater than his maximum feasible payoff so $G + Ha_1(n) \geq \hat{b} + 2\epsilon$. (17) ensures that $b(n)$ is within $\epsilon/16$ of $\hat{b}$, so the right of the above expression is at least $2\epsilon - \epsilon/16$. We have shown that after the first randomization the value $a_1(n) - \hat{A}_1$ declines (at least) exponentially with each randomization at some constant rate, say $\psi < 1$, so $a_1(n+n') - \hat{A}_1 < \psi[a_1(n) - \hat{A}_1]$ (where n and $(n+n')$ refer to consecutive randomizations, as before). Since $\hat{A}_1 < \hat{a}_1 + \epsilon$ this implies $a_1(n+n') - (\hat{a}_1 + \epsilon) < \psi[a_1(n) - (\hat{a}_1 + \epsilon)]$. Thus even if the first iteration (i.e., up to the first randomization) had an arbitrarily small effect, and since a_1 at the first randomization is bounded above by $\bar{a}_1$, it follows that after h randomizations $a_1(n) - (\hat{a}_1 + \epsilon) < \psi^{h-1}[\bar{a}_1 - (\hat{a}_1 + \epsilon)]$. If h^* satisfies $\psi^{h^*-1} < \epsilon[\bar{a}_1 - (\hat{a}_1 + \epsilon)]^{-1}$ we can be certain that at most h^* randomizations are required before $a_1(n) \leq \underline{a}_1(2\epsilon) + \frac{1}{2}\epsilon$, and that there is a strictly positive lower bound $\underline{r} \geq 2^{-h^*}$ on the probability of sticking to repeated play of the sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$. *Q.E.D.*

The lemma asserts that the total probability with which player 1 departs from repetitions of the sequence (by playing $\tilde{\imath}$ at one of the points of randomization) is bounded below one. Lemma 4 is essential because we can adapt its construction to build an equilibrium where player 1 is one of *two* different types: type k always plays the fixed sequence of actions and type 1 plays the sequence with occasional randomizations. By requiring the probability of type k to be sufficiently small (in particular it must be less than $\underline{r}$), and by adjusting the probability that type 1 plays $\tilde{\imath}$, the actions of the two types will combine to reproduce the strategy $\hat{\sigma}(N)$ and the optimal response by player 2 thus remains $\hat{\tau}(N)$.

4.2 The Repeated Game of Incomplete Information

The definition of individual rationality given in Section 3.1 applies to player 1's undiscounted payoffs. In *discounted* games as the players become more patient, player 2 is able to approximate these punishments arbitrarily closely. First we define the notion of ϵ-IR payoffs.

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

(The notation **1** is used to denote a $K-$ dimensional vector of 1's.) There is a lower threshold on the discounting, δ_ϵ, so that if $\delta > \delta_\epsilon$ then player 2 can hold player 1 down any ϵ-IR payoff in $\Gamma(\mathbf{p}, \delta)$. 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 (1). Then 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 $(-A_k(i,j))_{k \in K}$ (e.g., Zamir (1992, p.126)). Now consider the zero-sum discounted repeated game of incomplete information with the same payoffs. The value function for this game, $v_\delta(\mathbf{p})$, exists and satisfies $0 \leq v_\delta(\mathbf{p}) - \text{Cav } a(\mathbf{p}) \leq M\sqrt{\{(K-1)(1-\delta)/(1+\delta)\}}$ (see Zamir (1992, pp.119-125)). This implies that 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 (details of this final step available on request).

The Folk Theorem for discounted repeated games of complete information, as usually stated, applies only to strictly individually rational payoffs. Likewise, we shall assume (in (A.1)) that we can find strictly (by a margin of at least $\bar{\epsilon}$) individually rational payoffs for the repeated game of incomplete information $\Gamma(\mathbf{p}, \delta)$.

(A.1) *There exists* $(\breve{\pi}_1, \breve{\pi}_2, ..., \breve{\pi}_K) \in (\Delta^{IJ})^K$ *and* $\bar{\epsilon} > 0$ *such that* $(A_k(\breve{\pi}_k))_{k \in K}$ *is* $\bar{\epsilon}$*-IR and* $B(\breve{\pi}_k) > \hat{b} + \bar{\epsilon}$ *for all* $k \in K$.

We define strict individual rationality 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 $(\breve{\pi}_k)_{k \in K} \in (\Delta^{IJ})^K$ 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 $k \in K$, but A.1 requires more. In particular, it 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. It is, nevertheless, a natural extension of the implicit restriction made in the complete information case.

Using A.1 we can now describe a particular equilibrium, which we refer to as the *terminal equilibrium*. The terminal equilibrium is revealing in the sense that there is an initial signalling phase, where each player signals her type with possible pooling, and no information is revealed thereafter. In general the incentive compatibility conditions (that each type should have no incentive to mimic another type) will bind most tightly at such an equilibrium. We therefore choose the payoffs at the equilibrium so type k receives a payoff close to $\bar{a}_k(\epsilon)$. (This was why, in Lemma 4, terminal payoffs were restricted to be high.) The terminal equilibrium will serve to describe the players' long-run behaviour in $\Gamma(\mathbf{p}, \delta)$, apart from on paths on which player 1 reveals herself to be type 1 earlier in the game.

Lemma 5 *Given A.1, there exists an* $\tilde{\epsilon} > 0$ *such that for all* $\epsilon < \tilde{\epsilon}$*: there exists a* $\bar{\delta}(\epsilon) < 1$ *such that for all* $\delta > \bar{\delta}(\epsilon)$ *and all* $\mathbf{p} \in \Delta^K$ *the game* $\Gamma(\mathbf{p}, \delta)$ *has an equilibrium with payoffs,* $((\bar{\alpha}_1, ..., \bar{\alpha}_K), \bar{\beta})$*, that satisfy:*
(a) $\bar{a}_k(3\epsilon) - \frac{1}{2}\epsilon \geq \bar{\alpha}_k > \bar{a}_k(3\epsilon) - C\epsilon$ *for some constant* C*, independent of* ϵ *and* δ*, and for* $k = 1, 2, ..., K$*;*
(b) $\bar{\beta} \geq \hat{b} + 3\epsilon$.

<u>PROOF:</u> We start by constructing correlated strategies that give the players payoffs close to their maximum feasible and individually rational payoffs. Consider the convex set

$$D_\epsilon := \bigcap_{k=1}^{K} \{ \ \pi \in \Delta^{IJ} \mid A_k(\pi) \leq \bar{a}_k(3\epsilon) - \frac{3}{4}\epsilon, \ B(\pi) \geq \hat{b} + 4\epsilon \}.$$

D_0 has a non-empty interior, by A.1. D_ϵ is defined by $K+1$ linear inequalities which are continuous in ϵ and become tighter as ϵ increases. Define $\hat{\epsilon} > 0$ to be the largest ϵ such that $D_\epsilon \neq \emptyset$ for all $\epsilon \leq \hat{\epsilon}$. For $k = 1, 2, ..., K$ and $\epsilon \leq \hat{\epsilon}$, choose $\pi_k^*(\epsilon)$ to maximize $A_k(.)$ on the constraint set D_ϵ; obviously $A_k(\pi_k^*(0)) = \bar{a}_k(0)$. We will define $\tilde{\epsilon}$ to be the largest value of $\epsilon \leq \hat{\epsilon}$ such that the vector $(A_k(\pi_k^*(\epsilon))_{k \in K})$ is 3ϵ-IR.

We will now show that there exists a constant C^o, independent of ϵ and δ, so that

$$C^o\epsilon > \bar{a}_k(3\epsilon) - A_k(\pi_k^*(\epsilon)), \qquad \text{for } \epsilon \leq \tilde{\epsilon}, \forall\ k. \tag{22}$$

Let k be given and let $\pi^o := \pi_k^*(0)$. Also, for $\lambda \in [0, 1]$ define $\pi^\lambda := \lambda\pi^\dagger + (1 - \lambda)\pi^o$, where $\pi^\dagger \in D_{\tilde{\epsilon}}$. By linearity $B(\pi^\lambda) \geq \lambda(\hat{b} + \hat{\epsilon}) + (1 - \lambda)\hat{b}$, so π^λ is a feasible solution to $\max\{ \ A_k(\pi) \mid B(\pi) \geq \hat{b} + \lambda\hat{\epsilon}\}$. Thus $\bar{a}_k(\lambda\hat{\epsilon}) \geq A_k(\pi^\lambda) = \lambda A_k(\pi^\dagger) + (1 - \lambda)\bar{a}_k(0)$. Let

$\lambda = \epsilon/\tilde{\epsilon}$ for $0 \leq \epsilon \leq \hat{\epsilon}$; then this implies

$$\bar{a}_k(\epsilon) \geq \bar{a}_k(0) - \epsilon\frac{\bar{a}_k(0) - A_k(\pi^\dagger)}{\tilde{\epsilon}}\ , \qquad \forall\ \epsilon < \tilde{\epsilon}.$$

Define C_k to be the term that multiplies ϵ; then for $\epsilon < \tilde{\epsilon}$ and $\forall\ k$,

$$\bar{a}_k(\epsilon) \geq \bar{a}_k(0) - C_k\epsilon, \tag{23}$$

and note that C_k is a constant independent of ϵ and δ. Consider again, for a fixed k, the correlated strategy π^λ. If $\lambda \geq \epsilon/\hat{\epsilon}$, then π^λ satisfies the constraint $B(\pi^\lambda) \geq \hat{b} + 4\epsilon$. If $\lambda \geq \epsilon(\frac{3}{4} + 3C_{k'})/(\bar{a}_{k'}(0) - A_{k'}(\pi^\dagger))$ for all k', then π^λ satisfies the constraint $A_{k'}(\pi^\lambda) \leq \bar{a}_{k'}(3\epsilon) - \frac{3}{4}\epsilon$ for all k'. This second condition follows from rearranging the below sufficient condition for the constraint:

$$(1-\lambda)\bar{a}_{k'}(0) + \lambda A_{k'}(\pi^\dagger) \leq \bar{a}_{k'}(0) - C_{k'}3\epsilon - \frac{3}{4}\epsilon \tag{24}$$

(it is sufficient since the LHS of (24) is an upper bound for $A_{k'}(\pi^\lambda)$, while the RHS is no greater than $\bar{a}_{k'}(3\epsilon) - \frac{3}{4}\epsilon$ by (23)). Thus $\pi^\lambda \in D_\epsilon$ if $\lambda \geq E\epsilon$, where E is a positive constant. The value $A_k(\pi^{E\epsilon})$ is, therefore, a lower bound on $A_k(\pi_k^*(\epsilon))$ for $\epsilon < 1/E$. This implies that

$$\bar{a}_k(3\epsilon) - A_k(\pi_k^*(\epsilon)) \leq \bar{a}_k(0) - A_k(\pi^{E\epsilon}) = E[\bar{a}_k(0) - A_k(\pi^\dagger)]\epsilon$$

for $\epsilon < x$, for some $x > 0$, and thus a constant C_k^o exists such that for $\epsilon < x$, $C_k^o\epsilon > \bar{a}_k(3\epsilon) - A_k(\pi_k^*(\epsilon))$. It follows that on any compact interval for which $\bar{a}_k(3\epsilon) - A_k(\pi_k^*(\epsilon))$ is defined a linear upper bound exists with finite slope, and in particular it has a linear upper bound on $[0, \tilde{\epsilon}]$, and (22) follows.

By Result 2, for any $\delta > \hat{\delta}(\epsilon)$ we can specify K sequences of action profiles $\{(i_k^t, j_k^t)\}_{t=0}^\infty$ such that

$$\begin{aligned} A_{k'}(\pi_k^*(\epsilon)) &= (1-\delta)\sum_{s=0}^{\infty}\delta^t A_{k'}(i_k^t, j_k^t), \qquad \forall k, k' \in K, \\ B(\pi_k^*(\epsilon)) &= (1-\delta)\sum_{s=0}^{\infty}\delta^t B(i_k^t, j_k^t), \qquad \forall k \in K. \end{aligned}$$

By Result 2 we can also choose these sequences so that, for all k, k', player k''s continuation payoffs, if play follows $\{i_k^t, j_k^t)\}_{t=0}^\infty$, are within $\epsilon/2$ of $A_{k'}(\pi_k^*(\epsilon))$ at all future times. These sequences will be our equilibrium path actions. As $(A_k(\pi_k^*(\epsilon))_{k\in K})$ is 3ϵ-IR there is a profile of IR payoffs $(\breve{\omega}_k)_{k\in K}$, satisfying $\breve{\omega}_k + 3\epsilon < A_k(\pi_k^*(\epsilon))$, and player 1 will be punished for an observable deviation by being held down to $\breve{\omega}_k + \epsilon$ for all k.

In this proof we will choose $\bar{\delta}(\epsilon) < 1$ so that (i) $\bar{\delta}(\epsilon) > \hat{\delta}(\epsilon)$, (ii) $\bar{\delta}(\epsilon) > \delta_\epsilon$, (iii) $\bar{\delta}(\epsilon) > [16M/(16M+\epsilon)]^{1/K}$, (iv) $\bar{\delta}(\epsilon) > [(\hat{b}+3\epsilon+M)/(\hat{b}+4\epsilon+M)]^{1/K}$ for all k. The second condition ensures that player 2 can hold the types of player 1 to within ϵ of any IR payoffs. The third ensures that the loss from signalling is at most $\epsilon/8$ and the last condition will ensure that player 2 never gets less than $\hat{b}+3\epsilon$.

We now take $\epsilon < \tilde{\epsilon}$ to be given. We now show that the following strategies are an equilibrium of $\Gamma(\mathbf{p},\delta)$: Player 2 begins by playing the fixed sequence of actions associated with type 1, $\{j_1^t\}$, and if he observes player 1 deviating from her corresponding sequence $\{i_1^t\}$ in period t, for $t = 0, 1, ..., K-2$, he interprets this move as a signal that player 1 is type $k = t+2$. When type k is signalled he then begins to play out the sequence $\{j_k^t\}_{t=0}^\infty$ from the beginning and expects player 1 to play out the corresponding sequence $\{i_k^t\}_{t=0}^\infty$. If player 1 deviates from the sequence $\{i_1^t\}$ in period $t > K-2$, or deviates from the sequence $\{i_k^t\}$ once type k has been signalled, then player 2 punishes these deviations by holding her to the payoffs $(\breve{\omega}_k)_{k\in K} + \epsilon\mathbf{1}$ (defined above). This is possible as $\delta > \delta_\epsilon$. Each of player 1's types plays a best response to this strategy of player 2 and minmaxes player 2 if he deviates from the above strategy.

If type k signals truthfully, then her expected payoff is bounded below by $\bar{a}_k(3\epsilon) - C^o\epsilon - \frac{1}{8}\epsilon$. (We have shown that $A_k(\pi_k^*(\epsilon)) > \bar{a}_k(3\epsilon) - C^o\epsilon$ and the assumption $16M(1-\delta^K) < \epsilon\delta^K$ implies that the payoffs over the first $K-1$ periods contribute at most $\epsilon/8$ to her total payoff.) Thus the optimal response of type k to 2's strategy must give her a payoff, $\bar{\alpha}_k$, satisfying $\bar{\alpha}_k > \bar{a}_k(3\epsilon) - (C^o + \frac{1}{8})\epsilon$, since she always has the option of signalling truthfully. Then once we have established equilibrium, the lower bound on equilibrium payoffs to player 1 will be as required with $C = C^o + \frac{1}{8}$. In general the optimal response for type k will be to signal some type k' (which may be k itself) and never to trigger the punishment from player 2. Suppose this is false, so that it is optimal for type k to signal type k' and to trigger the punishment after s periods of mimicking type k'. Her payoff from playing out the sequence $\{(i_{k'}^t, j_{k'}^t)\}_{t=0}^\infty$ in its entirety can be decomposed into her average payoff over the first s periods, x, and her average payoff over the remaining periods, y, that is, $A_k(\pi^*_{k'}(\epsilon)) = (1-\delta^s)x + \delta^s y$. By the construction of the sequence of actions, at any point in time the continuation payoff satisfies $y \geq A_k(\pi^*_{k'}(\epsilon)) - \epsilon/2$. These two facts imply an upper bound on x: $(1-\delta^s)x \leq (1-\delta^s)A_k(\pi^*_{k'}(\epsilon)) + \delta^s\epsilon/2$. Her payoff (discounted to the period after the signal is sent) from mimicking type k' and then deviating in period s is thus bounded above by

$$(1-\delta^s)A_k(\pi^*_{k'}(\epsilon)) + \delta^s\epsilon/2 + (1-\delta)\delta^s M + \delta^{s+1}(\breve{\omega}_k + \epsilon). \tag{25}$$

If she prefers to be punished from time s, then $A_k(\pi^*_{k'}(\epsilon)) \leq \breve{\omega}_k + 25\epsilon/16$, because her payoff from continuing to play $\{i^t_{k'}\}_{t=0}^{\infty}$ is at least $A_k(\pi^*_{k'}(\epsilon)) - \epsilon/2$ by the construction of the action sequences, and the deviation payoff is at most $(1-\delta)M + \delta(\breve{\omega}_k + \epsilon) \leq \breve{\omega}_k + \epsilon(1+1/16)$. This upper bound for $A_k(\pi^*_{k'}(\epsilon))$ and the bound on δ implies that (25) is less than $\breve{\omega}_k + 2\epsilon$. By the definition of $\tilde{\epsilon}$ the payoffs $(A_k(\pi^*_k(\epsilon))_{k\in K}$ are 3ϵ-IR, so this is strictly less than the payoff from truthful revelation, described above, which gives a contradiction. Likewise, an observable deviation during the signalling leads to a payoff of at most $\breve{\omega}_k + \epsilon + \frac{1}{8}\epsilon$, which is less than the payoff from truthful revelation. Type k's equilibrium payoffs can now be broken down into a payoff from signalling and a payoff $A_k(\pi^*_{k'}(\epsilon))$ after signalling. This is bounded above by $(1-\delta^K)M + \delta^K(\bar{a}_k(3\epsilon) - \frac{3}{4}\epsilon)$, by definition of $\pi^*_{k'}(\epsilon)$. Assumption (iii) on δ ensures that this is less than $\bar{a}_k(3\epsilon) - \frac{1}{2}\epsilon$. The upper bound on equilibrium payoffs is established.

Player 2's expected payoff is determined by playing at most $K-1$ arbitrary actions followed by one of the fixed sequences $\{(i^t_k, j^t_k)\}$. His equilibrium payoff is therefore no less than $(1-\delta^K)(-M) + \delta^K(\hat{b} + 4\epsilon)$. This lower bound is strictly greater than $\hat{b} + 3\epsilon$ (by the fourth assumption on δ). This proves part (b) of the Lemma. His payoff from a deviation is at most $(1-\delta)(M) + \delta\hat{b}$, so we have also shown that player 2 cannot profitably deviate from the strategy above. *Q.E.D.*

The next result determines $K-1$ correlated strategies $(\underline{\pi}_2, ..., \underline{\pi}_K) \in (\Delta^{IJ})^{K-1}$, and each correlated strategy $\underline{\pi}_k$ will be mimicked by the finite sequence of actions played by type k. It shows that: (a) each correlated strategy holds type 1 to her minmax level or lower; (b) normalizing for the effect on type 1's payoff, each correlated strategy satisfies an incentive compatibility condition; (c) there is an individually rational point $\mathbf{z} \in \Re^K$ where type 1 receives her minmax payoff and type $k > 1$ receives a convex combination of her payoff $\bar{a}_k$ and the payoff she gets from playing the correlated strategy, that is

$\bar{a}_k + \lambda_k(A_k(\underline{\pi}_k) - \bar{a}_k)$, where the weight λ_k is chosen to hold type 1 to her minmax level when using the same correlated strategy, $\bar{a}_1 + \lambda_k(A_1(\underline{\pi}_k) - \bar{a}_1) = \hat{a}_1$.

Lemma 6 *Assume A.1, then there exist correlated strategies* $(\underline{\pi}_2, ..., \underline{\pi}_K) \in (\Delta^{IJ})^{K-1}$ *such that:*
(a) $A_1(\underline{\pi}_k) \leq \hat{a}_1$ *for all* $k = 2, 3, ..., K$,
(b) $(A_k(\underline{\pi}_k) - \bar{a}_k)/(\bar{a}_1 - A_1(\underline{\pi}_k)) \geq (A_k(\underline{\pi}_{k'}) - \bar{a}_k)/(\bar{a}_1 - A_1(\underline{\pi}_{k'}))$
for all $k, k' = 2, 3, ..., K$,
(c) $\mathbf{z}$ *is individually rational, where*

$$\mathbf{z} := \left(\hat{a}_1, \bar{a}_2 + \frac{\bar{a}_1 - \hat{a}_1}{\bar{a}_1 - A_1(\underline{\pi}_2)}(A_2(\underline{\pi}_2) - \bar{a}_2), ..., \bar{a}_K + \frac{\bar{a}_1 - \hat{a}_1}{\bar{a}_1 - A_1(\underline{\pi}_K)}(A_K(\underline{\pi}_K) - \bar{a}_K)\right).$$

<u>PROOF:</u> Consider the constrained optimization

$$\max_{\pi \in \Delta^{IJ}} \frac{A_k(\pi) - \bar{a}_k}{\bar{a}_1 - A_1(\pi)}, \qquad \text{subject to } A_1(\pi) \leq \hat{a}_1. \tag{26}$$

As $\bar{a}_1 > \hat{a}_1$, by assumption A.1, the maximand is well defined. As the constraint set is non-empty (by the Minimax Theorem) and compact there is a solution $\underline{\pi}_{k'}$ to the optimization for all $k' > 1$.

We aim to show that the point $\mathbf{z}$, defined above, is individually rational. We must, therefore, show that the set $\{\mathbf{x}|\mathbf{x} \leq \mathbf{z}\}$ is approachable. By Zamir (1992), for example, it is sufficient to show that for any $\mathbf{q} \in \Re^K$ with $\mathbf{q} \geq 0$ there exists a mixed action, g, for player 2 such that

$$\mathbf{q}((A_1(i, g), ..., A_K(i, g)) - \mathbf{z}) \leq 0, \qquad \forall i \in I. \tag{27}$$

Let $\hat{g}$ be a mixed strategy that ensures player 2 receives his minmax level ($B(i, \hat{g}) \geq \hat{b}$ for all $i \in I$) and let $\hat{g}_1$ be a mixed strategy that minmaxes type 1 ($A_1(i, \hat{g}_1) \leq \hat{a}_1$ for all $i \in I$). We will show that for any $\mathbf{q} \geq 0$ either $g = \hat{g}$ or $g = \hat{g}_1$ will ensure (27) holds. If (27) holds for all $\mathbf{q}$ when $g = \hat{g}$ then there is nothing to prove. Suppose that for some $\mathbf{q} \geq 0$ (27) does not hold with $g = \hat{g}$; then there exists $i \in I$ such that $\mathbf{q}((A_1(i, \hat{g}), ..., A_K(i, \hat{g})) - \mathbf{z}) > 0$.

By the definition of $\bar{\mathbf{a}}$, $\bar{a}_k \geq A_k(i, \hat{g})$, and together with the fact that $\mathbf{q} \geq 0$, this implies $\mathbf{q}(\bar{\mathbf{a}} - \mathbf{z}) > 0$. A substitution from the definition of $\mathbf{z}$ shows this is equivalent to

$$(28) \qquad (\bar{a}_1 - \hat{a}_1) \left(q_1 + \sum_{k=2}^{K} q_k \frac{A_k(\underline{\pi}_k) - \bar{a}_k}{A_1(\underline{\pi}_k) - \bar{a}_1} \right) > 0.$$

We must show that if (28) holds, $\mathbf{q}((A_1(i, \hat{g}_1), ..., A_K(i, \hat{g}_1)) - \mathbf{z}) \leq 0$ for all $i \in I$. It is sufficient to show $\mathbf{q}((A_1(\pi), ..., A_K(\pi)) - \mathbf{z}) \leq 0$ for all π such that $A_1(\pi) \leq \hat{a}_1$. A substitution for $\mathbf{z}$ then gives

$$\begin{aligned}
&\mathbf{q}((A_1(\pi), ..., A_K(\pi)) - \mathbf{z}) \\
&\quad = q_1(A_1(\pi) - \hat{a}_1) + \sum_{k=2}^{K} q_k \left(A_k(\pi) - \bar{a}_k + (\bar{a}_1 - \hat{a}_1) \frac{A_k(\underline{\pi}_k) - \bar{a}_k}{A_1(\underline{\pi}_k) - \bar{a}_1} \right) \\
&\quad = (A_1(\pi) - \hat{a}_1) q_1 + (\bar{a}_1 - A_1(\pi)) \sum_{k=2}^{K} q_k \left(\frac{A_k(\pi) - \bar{a}_k}{\bar{a}_1 - A_1(\pi)} + \frac{\bar{a}_1 - \hat{a}_1}{\bar{a}_1 - A_1(\pi)} \frac{A_k(\underline{\pi}_k) - \bar{a}_k}{A_1(\underline{\pi}_k) - \bar{a}_1} \right) \\
&\quad \leq (A_1(\pi) - \hat{a}_1) \left(q_1 + \sum_{k=2}^{K} q_k \frac{A_k(\underline{\pi}_k) - \bar{a}_k}{A_1(\underline{\pi}_k) - \bar{a}_1} \right) \leq 0 \qquad \forall \pi \text{such that } A_1(\pi) \leq \hat{a}_1.
\end{aligned}$$

The first inequality arises because π is replaced by $\underline{\pi}_k$ in $(A_k(\pi) - \bar{a}_k)/(\bar{a}_1 - A_1(\pi))$ and this is therefore maximized on the set of π's with $A_1(\pi) \leq \hat{a}_1$. The final inequality then follows from (28). Thus if $\mathbf{q}((A_1(i, \hat{g}), ..., A_K(i, \hat{g})) - \mathbf{z}) > 0$ it must be true that $\mathbf{q}((A_1(i, \hat{g}_1), ..., A_K(i, \hat{g}_1)) - \mathbf{z}) \leq 0$. We can conclude that $\mathbf{z}$ is individually rational. *Q.E.D.*

In Lemma 7 we define $K-1$ finite sequences of actions that approximate the correlated strategies $(\underline{\pi}_2, ..., \underline{\pi}_K)$.

Lemma 7 *For any $\epsilon > 0$ there exists $\delta'(\epsilon) < 1$, a finite integer $T > 0$ and $K - 1$ sequences of actions $\{(\hat{\imath}^s_{k'}, \hat{\jmath}^s_{k'})\}_{s=0}^{T-1}$, for $k' = 2, 3, ..., K$, such that for all $1 > \delta > \delta'(\epsilon)$:* (a) *$|\hat{A}_{k,k'} - A_k(\underline{\pi}_{k'})| < \epsilon/2$ for $k \in K$, $k' = 2, 3, ..., K$;* (b) *$|\hat{B}_{k'} - B(\underline{\pi}_{k'})| < \epsilon/2$ for $k' = 2, 3, ..., K$; where*

$$(29) \qquad \hat{A}_{k,k'} := \frac{1-\delta}{1-\delta^T} \sum_{s=0}^{T-1} \delta^s A_k(\hat{\imath}^s_{k'}, \hat{\jmath}^s_{k'}), \qquad \hat{B}_{k'} := \frac{1-\delta}{1-\delta^T} \sum_{s=0}^{T-1} \delta^s B(\hat{\imath}^s_{k'}, \hat{\jmath}^s_{k'}).$$

<u>Proof:</u> For $k' = 2, 3, ..., K$, let $\pi(k')$ be a rational approximation to the correlated strategy $\underline{\pi}_{k'}$, such that $\|\underline{\pi}_{k'} - \pi(k')\| < \epsilon/4$ for $k' = 2, 3, ..., K$. There exists a positive integer

T such that $T\pi(k')_{ij}$ is an integer for all $k' = 2, 3, ..., K$, $i \in I$ and $j \in J$, (where $\pi(k)_{ij}$ denotes the ij^{th} element of the correlated strategy $\pi(k)$). Choose the $K-1$ sequences so that the action pair (i, j) appears $T\pi(k')_{ij}$ times in the sequence $\{(i^s_{k'}, j^s_{k'})\}_{s=0}^{T-1}$. Continuity then ensures that there exists $\delta'(\epsilon)$ such that for all $\delta > \delta'(\epsilon)$ the result holds. *Q.E.D.*

We now prove our main result. It contains three main elements. The first element of the proof is an investigation of the two-type game where only type 1 and type k are given positive probability by player 2. We describe an equilibrium of this game where the combined actions of the players (i.e., using the priors over player 1's types) replicate the strategies $(\hat{\sigma}(N), \hat{\tau}(N))$, described in Lemma 4: type k repeatedly plays the finite sequence of Lemma 7, while type 1 occasionally randomizes. As there is strictly positive probability that this sequence is played out in full, provided the probability of type k is less than $\underline{r}$, it is possible for the combined actions of the types to replicate the strategy $\hat{\sigma}(N)$. And if the sequence is played out in full the players settle down at the equilibrium described in Lemma 5. In this construction we will use Lemma 6 to define punishments. By Lemma 4 we can therefore deduce that, provided type 1 is given sufficiently high probability, there is an equilibrium where type 1's payoff is arbitrarily close to any $a_1 \in [\underline{a}_1(0), \bar{a}_1(0)]$.

The second step is to add an initial random move by type 1 in the two-type game. At this random move type 1 reveals herself with a high probability and after this plays out an equilibrium of the full information game where player 2 receives the payoff b. Provided the probability of type 1 is sufficiently high, this allows us to find an equilibrium of the two-type game where, for given $\nu > 0$, and given any pair $(a_1, b) \in G_1(\nu)$, type 1's equilibrium payoff is close to a_1 and player 2's payoff is close to b. The final step in the construction is an initial signalling phase where the types $k' > 1$ of player 1 signal their type and type 1 randomly mimics one of the types $k' > 1$. This is not simple to implement, because type 1 must be made indifferent between mimicking all other types. To ensure her indifference it is necessary that player 2 randomizes in the period that type k signals and that the outcome of player 2's randomization determines the equilibrium of the two-type game that is subsequently played.

Theorem 3 *Assume A.1 and let $\nu > 0$ be given. Then there exists $\delta_\nu < 1$, $p_1^\nu < 1$ such* *that for all* $\mathbf{p}$ *with* $p_1 > p_1^\nu$ *and for all* $\delta > \delta_\nu$*, given any* $(a_1, b) \in G_1(\nu)$ *the game* $\Gamma(\mathbf{p}, \delta)$ *has an equilibrium with the payoffs* $((\alpha_1, ..., \alpha_K), \beta) \in \Re^{K+1}$ *which satisfy*

$$\|(\alpha_1, \beta) - (a_1, b)\| < \nu. \tag{30}$$

Proof: *Some definitions and notation:* Choose $Q > 0$ to be a linear upper bound on the difference between $\bar{a}_k(\epsilon)$ and $\bar{a}_k$ for all $\epsilon \in (0, \bar{\epsilon})$ and for all k (where $\bar{\epsilon}$ is defined in Assumption 1); in particular, choose Q so that

$$\bar{a}_k - \bar{a}_k(3\epsilon) + 3\epsilon/4 < Q\epsilon \qquad \forall k \in K, \quad 0 < \epsilon < \bar{\epsilon}. \tag{31}$$

(See, e.g., the argument for (23) in Lemma 5.) We will also define a non-negative constant R as follows (where $\underline{\pi}_k$ is defined in Lemma 6):

$$R := \max_k \left| \frac{\bar{a}_k - A_k(\underline{\pi}_k)}{\bar{a}_1 - A_1(\underline{\pi}_k)} \right|. \tag{32}$$

From Lemma 6(b) we have that

$$\frac{A_k(\underline{\pi}_k) - \bar{a}_k}{\bar{a}_1 - A_1(\underline{\pi}_k)} \geq \frac{A_k(\underline{\pi}_{k'}) - \bar{a}_k}{\bar{a}_1 - A_1(\underline{\pi}_{k'})}, \qquad \forall k, k' = 2, 3, ..., K. \tag{33}$$

We will begin by assuming that this inequality is strict when $k \neq k'$, that is,

$$\frac{A_k(\underline{\pi}_k) - \bar{a}_k}{\bar{a}_1 - A_1(\underline{\pi}_k)} > \frac{A_k(\underline{\pi}_{k'}) - \bar{a}_k}{\bar{a}_1 - A_1(\underline{\pi}_{k'})}, \qquad \forall k, k' = 2, 3, ..., K; \quad k \neq k'. \tag{34}$$

(We will deal with the case of $k \neq k'$ satisfying (33) with equality at the end of the proof.) Finally, Y is defined to be the slope (with 2's payoffs in the numerator) of $G_1(0)$ when this set is a line segment ($\operatorname{Int} G_1(0) = \emptyset$) and when $\operatorname{Int} G_1(0) \neq \emptyset$ we define $Y = 1$. Y is bounded above and strictly positive by Assumption A.1.

Let $\iota > 0$ be given, where $\iota < \bar{\epsilon}$. Choose $\epsilon > 0$ so that: (i) $3\epsilon < \iota$; (ii) for all $k, k' = 2, 3, ..., K$ with $k \neq k'$ it is true that for all $\delta > \delta'(\epsilon)$

$$\frac{\hat{A}_{k,k} - \bar{a}_k}{\bar{a}_1 - \hat{A}_{1,k}} > \frac{\hat{A}_{k,k'} - \bar{a}_k}{\bar{a}_1 - \hat{A}_{1,k'}} + (2 + R)\epsilon; \tag{35}$$

where $\hat{A}_{k,k'}$ and $\delta'(\epsilon)$ are as defined in Lemma 7; (iii) $\lambda \in [0, 1]$ such that $\lambda \hat{a}_1 + (1 - \lambda)\bar{a}_1 > \hat{a}_1 + \iota - \epsilon/2$ implies $\lambda \mathbf{z} + (1 - \lambda)\bar{\mathbf{a}}$ is $(2 + (Q + 2)(R + 1))\epsilon$-IR; (iv) $\underline{a}_1(4\frac{1}{16}\epsilon) + \epsilon < \underline{a}_1(\iota) < \bar{a}_1(3\epsilon) - C\epsilon$ where C is defined in Lemma 5 ($\underline{a}_1(\iota) < \bar{a}_1(0)$, because $G_1(\bar{\epsilon})$ is non-empty by Assumption 1 and $\iota < \bar{\epsilon}$, so the last inequality holds for small ϵ); (v) $\iota > [8(9/8)^{K-2} - 7]\epsilon \max\{Y, 1\}$. ((ii) is possible as we have assumed (34) and the payoffs from playing out the action sequences can be made arbitrarily close to the payoff from playing the correlated strategies $\underline{\pi}_k$, $|\hat{A}_{k,k'} - A_k(\underline{\pi}_{k'})| < \epsilon/2$, by Lemma 7. (iii) is possible because the sets of ϵ-IR payoffs are convex and these sets converge to the set of IR payoffs as $\epsilon \to 0$. So (a) as the point $\bar{\mathbf{a}}$ is $(2+(Q+2)(R+1))\epsilon$-IR for ϵ sufficiently small, (b) the set of ϵ-IR payoffs is convex and converges to the set of IR payoffs as $\epsilon \to 0$, and (c) the point $\mathbf{z}$ is IR, the convex combination $(1-\lambda)\mathbf{z}+\lambda\bar{\mathbf{a}}$, for a given $\lambda < 1$ will be $(2+(Q+2)(R+1))\epsilon$-IR provided ϵ is sufficiently small.) Given this value for ϵ, let T and $\delta'(\epsilon)$ be as defined in Lemma 7, and setting $\delta^*(\epsilon) = \delta'(\epsilon)$, let $\tilde{\delta}(\epsilon)$ be as defined in Lemma 4 (each of the $K-1$ finite sequences specified in Lemma 7 satisfies the conditions of Lemma 4; $\tilde{\delta}(\epsilon)$ depends on them only through T). Choose $\delta_\iota = \max\{\tilde{\delta}(\epsilon), \delta_\epsilon, \bar{\delta}(\epsilon), (4M/(4M+\epsilon)^{1/K}\}$, where δ_ϵ is defined below Definition 2 and $\bar{\delta}(\epsilon)$ is defined in Lemma 5.

1. The Game with Two Types : arbitrary payoff for type 1

Let some type $k > 1$ be given. Recall that Lemma 4 defined an equilibrium $(\hat{\sigma}(N), \hat{\tau}(N))$ of the complete information game where, with occasional randomizations, type 1 and player 2 play out a finite sequence of actions N times and then settle on an equilibrium. Recall also that type 1's average payoff over the finite sequence of actions $\{(\hat{\imath}_k^s, \hat{\jmath}_k^s)\}_{s=0}^{T-1}$ (defined in Lemma 7) is not greater than $\hat{a}_1 + \epsilon$ for all $\delta > \delta'(\epsilon)$, and for all $\delta > \bar{\delta}(\epsilon)$ that the equilibrium defined in Lemma 5 has payoffs, $(\bar{\alpha}_1, \bar{\alpha}_2, ..., \bar{\alpha}_K, \bar{\beta})$, that satisfy $\bar{\beta} \geq \hat{b} + 3\epsilon$ and $\bar{a}_1(3\epsilon) - \epsilon/2 \geq \bar{\alpha}_1 > \bar{a}_1(3\epsilon) - C\epsilon$. Let $a_1' \in [\underline{a}_1(\iota), \bar{a}_1(3\epsilon) - C\epsilon]$ be given (this interval is non-empty by (iv) above); then by Lemma 4 with $(a_1^*, b) = (\bar{\alpha}_1, \bar{\beta})$, and by (iv), for all δ close to 1, there exists N and strategies which we denote as $(\hat{\sigma}(k; N), \hat{\tau}(k; N))$ which constitute an equilibrium of $\Phi_1(\delta)$, in which type 1 gets a payoff within $\frac{1}{32}\epsilon$ of a_1'. At this equilibrium the sequence $\{(\hat{\imath}_k^s, \hat{\jmath}_k^s)\}_{s=0}^{T-1}$ is played N times with occasional randomizations by type 1 and finally, if 1 has not deviated from the sequence, play settles on an equilibrium of $\Phi_1(\delta)$ where the players receive the payoffs $(\bar{\alpha}_1, \bar{\beta})$. By Lemma 4, there is a probability of at least $\underline{r}$, independent of δ, that type 1 ends up playing the equilibrium with payoffs $(\bar{\alpha}_1, \bar{\beta})$.

Let $\mathbf{p}$ with $0 < p_1 < \frac{1}{4}$ and $p_{k'} = 0$ for all $k' \neq 1, k$ be given. We will now show there exists a $\mathbf{p}'$, satisfying $p'_1 \geq p_1$, $p'_k \leq p_k$ and $p'_{k'} = 0$ for all $k' \neq 1, k$, such that the following strategies, or a slight modification explained below, are an equilibrium in the game $\Gamma(\mathbf{p}', \delta)$:

Type k plays out the finite sequence $\{(\hat{\imath}_k^s, \hat{\jmath}_k^s)\}_{s=0}^{T-1}$ N times and then plays out the strategy (for k) in the equilibrium of $\Gamma(\mathbf{p}, \delta)$ with the payoffs $(\bar{\alpha}_1, ..., \bar{\alpha}_K, \bar{\beta})$ given above. Deviations by player 2 from his equilibrium strategy are minmaxed.

Type 1 plays a strategy so that from player 2's perspective the combined actions of types 1 and k over the first TN periods replicate the strategy $\hat{\sigma}(k; N)$, defined above, and, after TN periods of playing the sequence, type 1 settles down to play the equilibrium of $\Gamma(\mathbf{p}, \delta)$ given above. Thus, in periods where $\hat{\sigma}(k; N)$ requires player 1 to randomize with probability 1/2, type 1 actually deviates from the sequence with probability more than 1/2 to compensate for the fact that type k never deviates from the sequence. If r (where $r > \underline{r}$) is the *total* probability that player 1 does not deviate from this sequence, then after TN periods player 2 has the prior $(r - (1 - p'_1))/r$ that player 1 is type 1. Provided we chose $\mathbf{p}'$ such that $p_1 = 1 - (1 - p'_1)/r$, or $p'_1 = 1 - r(1 - p_1)$, then playing the continuation equilibrium is feasible. Deviations by player 2 from his equilibrium strategy are minmaxed.

Player 2 will play out the strategy $\hat{\tau}(k; N)$ on the equilibrium path over the first TN periods with the terminal equilibrium of $\Gamma(\mathbf{p}, \delta)$ given above being played thereafter, or one of the revealing equilibria if type 1 has revealed her type. However, if player 1 uses a pure action that deviates from her equilibrium strategy (i.e., a probability zero action), then player 2 responds in the following way. He first calculates type 1's expected payoff if she were to continue playing out her strategy (and player 2 plays the actions described above); call this c. Then he takes the convex combination $\lambda \mathbf{z} + (1 - \lambda)\bar{\mathbf{a}}$, of the point $\mathbf{z}$ (defined in (**??**)) and the point $\bar{\mathbf{a}}$ (defined in Lemma 6), that gives type 1 exactly the payoff c, that is, $\lambda = (\bar{a}_1 - c)/(\bar{a}_1 - \hat{a}_1)$. By the construction above (point (iii) below (35)), since $c > \hat{a}_1 + \iota - \epsilon/2$ then this convex combination is $(2 + (1 + R)(2 + Q))\epsilon$-IR.[5]

[5] At the equilibrium strategy for type 1 described above, type 1's payoff at the start of each finite sequence is a convex combination of $\hat{A}_{1,k}$ and the terminal equilibrium payoff $\bar{\alpha}_1 : (1 - \delta^{nT})\hat{A}_{1,k} + \delta^{nT}\bar{\alpha}_1$, for some integer $n \leq N$. The integer $n = N$ is chosen so that her equilibrium payoff (i.e., at the start of the first round of the finite sequence) is within $\epsilon/32$ of $a'_1 \geq \hat{a}_1 + \iota$, and hence at least $\hat{a}_1 + \iota - \epsilon/32$. The payoff $\bar{\alpha}_1$ is at least $\bar{a}_1(3\epsilon) - C\epsilon > \hat{a}_1 + \iota$ (by the assumption on ϵ). Allowing for the small integer effects which arise when playing out the finite sequence of actions, it is thus the case that her continuation payoff

That is, there exists a vector of IR payoffs $(\omega_1, ..., \omega_K) \in \Re^K$ such that

$$(\omega_1, ..., \omega_K) + (2 + (1 + R)(2 + Q))\epsilon \mathbf{1} \leq \lambda \mathbf{z} + (1 - \lambda)\bar{\mathbf{a}}$$

$$= \left(c, \bar{a}_2 - (\bar{a}_1 - c)\frac{\bar{a}_2 - A_2(\underline{\pi}_2)}{\bar{a}_1 - A_1(\underline{\pi}_2)}, ..., \bar{a}_K - (\bar{a}_1 - c)\frac{\bar{a}_K - A_K(\underline{\pi}_K)}{\bar{a}_1 - A_1(\underline{\pi}_K)}\right). \quad (36)$$

Player 2 responds to a deviation of player 1 by holding each type k to a payoff of at most $\omega_k + \epsilon$, which is possible as $\delta > \delta_\epsilon$.

To show that these strategies form an equilibrium of the game $\Gamma(\mathbf{p}', \delta)$ which gives positive probability only to types $\{1, k\}$, it is sufficient to show that type 1 and type k do not benefit by deviating from their equilibrium strategy.[6] Some deviations are not observed by player 2. We will first concern ourselves with deviations that are immediately detected by player 2. It will be convenient to let c (as above) and d denote, respectively, type 1 and type k's equilibrium continuation payoffs at the start of the period in which the observed deviation occurred. We will first show that type 1 does not benefit by deviating. By the construction above, if $\delta > \delta_\epsilon$ then type 1's expected payoff from deviation is at most $(1 - \delta)M + \delta(\omega_1 + \epsilon)$, whereas her expected payoff from continuing, c, satisfies $c > \omega_1 + 3\epsilon$; our assumption on δ is sufficient to ensure a deviation is suboptimal.

Next, we show that type k cannot profitably deviate from these strategies. Type k can make unobservable deviations from the equilibrium by mimicking type 1 revealing her type (by playing $\tilde{\imath}$ at a point of randomization), and then by continuing to mimic type 1, playing out an equilibrium of the game $\Phi_1(\delta)$. It is possible that such a deviation is profitable. A small re-working of the players' strategies gives an equilibrium with the same payoff to type 1 and a greater payoff to type k, if this is the case. Let t denote the first time at which this unobservable deviation is profitable for type k. Redefine the players' equilibrium strategies, so that before time t all players use exactly the same actions and at time t both types play $\tilde{\imath}$ (the revealing action) and play out the strategies of the equilibrium of the game $\Phi_1(\delta)$. (Player 2's strategy is exactly the same as before.) This does not change type 1's equilibrium payoff because she was indifferent at $\tilde{\imath}$. It raises type k's equilibrium payoff, because she prefers the deviation to the equilibrium. Player 2's payoffs also increase because the continuation equilibrium after $\tilde{\imath}$ was chosen to reward

c at any point always exceeds $\hat{a}_1 + \iota - \epsilon/16$.

[6]Lemma 4 guarantees that type 1 is indifferent between the positive probability actions in periods when she must randomize, and that player 2 is playing an optimal response to types 1 and k.

him for playing out the iterations of the finite sequence and he now receives this reward with higher probability. Finally, to verify that this is an equilibrium we must show that type k will not benefit from making an observable deviation at some later stage from the equilibrium of $\Phi_1(\delta)$. We will address this in the parentheses after case (b) below.

Now, we consider observable deviations by k from the equilibrium, which result in player 2 punishing player 1. By (36) there exists a vector of punishment payoffs ω such that

$$\begin{aligned}
&\omega_k + (2+(1+R)(2+Q))\epsilon \\
&\quad \le \bar{a}_k - (\bar{a}_1 - c)\frac{\bar{a}_k - A_k(\underline{\pi}_k)}{\bar{a}_1 - A_1(\underline{\pi}_k)} \\
&\quad = \{(1-\delta^{TN'})\hat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - d\} + \delta^{TN'}\{\bar{a}_k - \bar{\alpha}_k\} + (1-\delta^{TN'})\{A_k(\underline{\pi}_k) - \hat{A}_{kk}\} \\
&\qquad + \frac{\bar{a}_k - A_k(\underline{\pi}_k)}{\bar{a}_1 - A_1(\underline{\pi}_k)}\Big\{(1-\delta^{TN'})[\hat{A}_{1k} - A_1(\underline{\pi}_k)] + [c - (1-\delta^{TN'})\hat{A}_{1k} - \delta^{TN'}\bar{\alpha}_1] \\
&\qquad\qquad - \delta^{TN'}[\bar{a}_1 - \bar{\alpha}_1]\Big\} + d \\
&\quad < d + \{(1-\delta^{TN'})\hat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - d\} + Q\epsilon + \epsilon/2 \\
&\qquad + \frac{\bar{a}_k - A_k(\underline{\pi}_k)}{\bar{a}_1 - A_1(\underline{\pi}_k)}\Big\{\epsilon/2 - (1-\delta^{TN'})\hat{A}_{1k} - \delta^{TN'}\bar{\alpha}_1 + c + Q\epsilon\Big\}.
\end{aligned} \tag{37}$$

The final inequality follows from (31), $A_k(\underline{\pi}_k) - \hat{A}_{kk} < \epsilon/2$ and $\hat{A}_{1k} - A_1(\underline{\pi}_k) < \epsilon/2$ (which follows from Lemma 7). Type 1's continuation payoff, c, is determined either by (a) continued playing out of the sequence $\{(\hat{\imath}_k^s, \hat{\jmath}_k^s)\}$ followed by the terminal equilibrium (in this case type k's deviation is detected immediately), or by (b) her payoff from continued playing out the revealing equilibrium (relevant when type k made an undetected deviation by playing $\tilde{\imath}$ and then later made an observable deviation). Let us deal first with a deviation by type k in case (a). If type 1 has N' complete repetitions of the sequence left to perform, then, analogously with the derivation of (17), type 1's payoff c satisfies $|(1-\delta^{TN'})\hat{A}_{1k} + \delta^{TN'}\bar{\alpha}_1 - c| \le \frac{\epsilon}{16}$ and type k's continuation payoff, d, satisfies $|(1-\delta^{TN'})\hat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - d| \le \frac{\epsilon}{16}$. These inequalities, and (32), substituted in (37), imply that $\omega_k + (3+R)\epsilon < d$; thus a deviation for type k is not profitable in this case (by the assumption on δ). Now let us consider case (b). Assume the observed deviation occurred t periods after $\tilde{\imath}$ was played, so an equilibrium of $\Phi_1(\delta)$ has been played for the last t periods. Let the sequence $\{(i^s, j^s)\}_{s=0}^{\infty}$ have as an initial point the move $(\tilde{\imath}, \hat{\jmath}^0)$ and then include the sequence of actions played by the two players at this equilibrium. Let

$\omega_2' = (1-\delta)a_k + (1-\delta)\omega_2$ denote k's payoff in the period she deviates and the subsequent payoffs from the punishment. Her continuation payoff from playing $\tilde{\imath}$ and then making an observable deviation satisfies

$$\begin{aligned}(1-\delta)\sum_{s=0}^{t-1}\delta^s A_k(i^s,j^s) + \delta^t\omega_2' &= (1-\delta^t)(1-\delta)\sum_{s=0}^{\infty}\delta^s A_k(i^s,j^s) + \delta^t\omega_2' \\ &\quad + \delta^t(1-\delta)[\sum_{s=0}^{\infty}\delta^s A_k(i^s,j^s) - \sum_{s=t}^{\infty}\delta^{s-t}A_k(i^s,j^s)].\end{aligned}$$

Let d' denote type k's continuation payoff from not playing $\tilde{\imath}$, but from abiding by her equilibrium strategy. The unobservable followed by the observable deviation is optimal only if $d' < (1-\delta)\sum_{s=0}^{t-1}\delta^s A_k(i^s,j^s) + \delta^t\omega_2'$. The above implies that this is equivalent to

$$d' - \omega_2' < \frac{1-\delta^t}{\delta^t}[(1-\delta)\sum_{s=0}^{\infty}\delta^s A_k(i^s,j^s) - d'] + (1-\delta)[\sum_{s=0}^{\infty}\delta^s A_k(i^s,j^s) - \sum_{s=t}^{\infty}\delta^{s-t}A_k(i^s,j^s)].$$

By the above construction of a pooling equilibrium, we can assume that the first term on the RHS is non-positive. The final term on the RHS is less that $(\frac{1}{16}+\frac{1}{2})\epsilon$, because the strategies $\hat{\sigma}(k;N)$, defined above Lemma 4, used Result 2 to ensure that play after $\tilde{\imath}$ gives all types within $\epsilon/2$ of their continuation payoff at $\tilde{\imath}$ at all future times and the playing of $\tilde{\imath}$ can change the payoff by at most $\frac{1}{16}\epsilon$. Thus, this condition can only be true if $d' < \omega_2' + \frac{9}{16}\epsilon$, or $d' < \omega_2 + (\frac{9}{16}+\frac{1}{16})\epsilon$ because of the assumption on δ. The punishment payoff, ω_2, is determined by (36) and (c,d) (the continuation payoffs at the point of the observed deviation by type k). Since: (c,d), ω_2 satisfy (37) and (c,d) is within $10\epsilon/16$ of the continuation payoffs (c',d') at the time of the unobserved deviation, we can deduce from (37) that $\omega_2 + \frac{1}{2}(5+R)\epsilon < d'$. This is a contradiction as $d' < \omega_2 + (10/16)\epsilon$. [In the pooling equilibrium, described in the previous paragraph, type k and type 1 each play out the equilibrium of $\Phi_1(\delta)$. Type k benefits by a subsequent observable deviation if $(1-\delta)\sum_{s=0}^{\infty}\delta^s A_k(i^s,j^s)$ is less than $(1-\delta)\sum_{s=0}^{t-1}\delta^s A_k(i^s,j^s) + \delta^t\omega_2'$. This implies $\omega_2' > (1-\delta)\sum_{s=t}^{\infty}\delta^s A_k(i^s,j^s) = d$. The above then implies $\omega_2 + \epsilon/16 > d' - \epsilon\frac{9}{16}$. We can then use (37) to get a contradiction again.]

2. The game with two types : arbitrary payoff for type 1 and player 2

The strategies above are an equilibrium, so, given any $\delta > \delta_\iota$, $a_1' \in [\underline{a}_1(\iota), \bar{a}_1(3\epsilon) - C\epsilon]$ and terminal priors $\mathbf{p}$ satisfying $0 < p_1 < \frac{1}{4}$ and $p_{k'} = 0$ for all $k' \notin \{1,k\}$, there exists $\mathbf{p}'$ (with

$p'_1 = 1 - r(1-p_1))$ and an equilibrium of the game $\Gamma(\mathbf{p}', \delta)$ with the payoffs $(\tilde{\alpha}_1, \tilde{\beta})$ where type 1's payoff, $\tilde{\alpha}_1$, satisfies $|\tilde{\alpha}_1 - a'_1| < \frac{1}{32}\epsilon$. We use this result to show that there exists an $\underline{r}' > 0$ such that if $\delta > \delta_\iota$, $p''_1 > 1 - \underline{r}'$ and $p''_{k'} = 0$ for all $k' \notin \{1, k\}$, then $\Gamma(\mathbf{p}'', \delta)$ has an equilibrium with the payoffs (α^*_1, β^*) that satisfy $\|(\alpha^*_1, \beta^*) - (a_1, b)\| < \epsilon$ for any pair $(a_1, b) \in G_1(\iota)$ with $a_1 < \bar{a}_1(3\epsilon) - C\epsilon$. To do this it is necessary to alter the period zero strategies of the equilibrium described in part 1. Now type 1 randomizes in period zero — with probability $1-\mu$ she plays out the equilibrium just described where a'_1 is set equal to a_1, and with probability μ she reveals her type by playing $\tilde{\imath} \neq \hat{\imath}^0$, and play then follows an equilibrium of the complete information game in which first-period actions are $(\tilde{\imath}, \hat{\jmath}^0)$. As in the previous part, we can choose the equilibrium in the complete information game so that type 1 is indifferent between the two first-period actions $\tilde{\imath}$ and $\hat{\imath}^0$. Let $(\tilde{a}_1, \tilde{b}) \in G_1(\epsilon)$ denote the payoffs, discounted to period 0, type 1 and player 2 receive conditional on $\tilde{\imath}$ being played in the first period. As type 1 randomizes in the first period $\tilde{a}_1 = \tilde{\alpha}_1$, so $\tilde{a}_1$ is within $\frac{1}{32}\epsilon$ of a_1 and we can therefore also chose $\tilde{b}$ to be within $\frac{1}{32}\epsilon$ of b (since $(a_1, b) \in G_1(\iota)$ and $\epsilon < \iota$). The arguments above imply that this will also be an equilibrium for $\delta > \delta_\iota$, provided player 2 has the priors $\mathbf{p}'$ after $\hat{\imath}^0$ is observed in the first period. Type 1 and player 2's expected payoffs from these strategies are $(\alpha^*_1, \beta^*) = (\tilde{\alpha}_1, p''_1\mu\tilde{b} + (1 - p''_1\mu)\tilde{\beta})$, so

$$\begin{aligned} |\beta^* - b| &= |p''_1\mu\tilde{b} + (1 - p''_1\mu)\tilde{\beta} - \tilde{b} + \tilde{b} - b| \\ &\leq |\tilde{\beta} - \tilde{b}|(1 - p''_1\mu) + |\tilde{b} - b| \leq 2M(1 - p''_1\mu) + \frac{\epsilon}{32}. \end{aligned}$$

If μ can be chosen to satisfy $\mu \geq (1 - \epsilon/(6M))/p''_1$, we can ensure that β^* is within $\epsilon/2$ of b. If $\hat{\imath}^0$ is observed in the first period player 2's posterior for type k is $(1 - p''_1)/(1 - \mu p''_1)$, so to play the equilibrium of part 1, μ must also satisfy $1 - p'_1 = (1 - p''_1)/(1 - \mu p''_1)$. As $1 - p'_1 = r(1 - p_1)$ (where r is the probability that player 1 does not deviate from the fixed sequence in the equilibrium of part 1) we can re-write this condition as $1 - p''_1 = r(1 - p_1)(1 - \mu p''_1)$. For any $\mathbf{p}''$ and $\mu \in [0, 1]$ that satisfy $\mu \geq [1 - \epsilon/(6M)]/p''_1$ and $1 - p''_1 = r(1 - p_1)(1 - \mu p''_1)$, we have found an equilibrium where type 1 and player 2 get payoffs close to (a_1, b). Given a p''_1, a value for $\mu > 0$ can be found to satisfy these two conditions provided $1 - p''_1 < r(1 - p_1)\epsilon/6M$. We chose $p_1 < \frac{1}{4}$ and by Lemma 4, $r > \underline{r}$, where $\underline{r} > 0$ is independent of δ and a_1, so a sufficient condition for this is $1 - p''_1 < \underline{r}\frac{3}{4}\epsilon/6M$. Provided $p''_1 > 1 - \underline{r}'$ where $\underline{r}' := \underline{r}\frac{3}{4}\epsilon/6M$ we have found an equilibrium of $\Gamma(\mathbf{p}'', \delta)$ with the desired properties. (If type k prefers to mimic the revelation action of type 1 at date

0, then the strategies can be amended as in part 1 to re-establish equilibrium.) When $K = 2$ the choice of $\iota = \min\{\nu, \bar{\epsilon}/2\}$ proves the Theorem.

3. The game with many types $K > 2$

We now describe the players' strategies in the repeated game of incomplete information $\Gamma(\mathbf{p}, \delta)$ where all types are given positive probability, and show that these strategies are an equilibrium with payoffs satisfying (30). The play in the game is divided into a signalling phase, where all types are given positive probability, and a payoff phase where only two types of player 1 are given positive probability.

Periods t=0,1,...,K-3 : The Signalling Phase: The players use the following strategies: *Type* k, where $k = 2, 3, ..., K-1$, plays action $i^t = 1$ in periods $t = 0, 1, ..., k-3$ and in period $t = k - 2$ she plays action $i = 2$ to signal her type. *Type* K plays action $i^t = 1$ in periods $t = 0, 1, ..., K-3$. *Type 1* chooses a type $k = 2, 3, ..., K$ with probability ϕ_k and mimics her signalling strategy. (All of the types of player 1 minmax player 2 if she chooses a pure action that is not played with positive probability in the signalling phase.) *Player 2* plays action $j = 1$ with probability q^0 and action $j = 2$ with probability $1 - q^0$ in period zero. If, in period $t < K - 2$, player 1 used action $i = 1$ in all past periods, then player 2 plays action $j = 1$ with probability $q^t(h^{t-1})$ and action $j = 2$ with probability $1 - q^t(h^{t-1})$, where h^{t-1} is the history of player 2's past actions up to $t - 1$. (If player 2 observes a deviation in period $t \leq K - 3$ then he plays the punishments described above for the 2-type game with the types $\{1, t + 2\}$.)

After the signalling: At the end of the signalling phase only two types of player 1, $\{1, k\}$, will be given positive probability by player 2. The players then play an equilibrium described in part 2 of the proof; however, the equilibrium they play will depend on the entire sequence of actions player 2 plays during the signalling phase; h^{t-1}.

We will begin by considering the case where $\operatorname{Int} G_1(0) \neq \emptyset$. Let (a_1, b) be a point in $G_1(\iota)$ that satisfies the condition $U[(a_1, b); \iota, \iota] \subset G_1(\iota) \cap \{(x, y) | x < \bar{a}_1(3\epsilon) - C\epsilon\}$ (ι will be chosen sufficiently small to ensure this is possible). Here we introduce notation for the open rectangle centred at the point (x, y) with width W and height H, that is,

$$U[(x_1, x_2); W, H] := \{ \ (x, y) \in \Re^2 \mid |x - x_1| < 0.5W, \ |y - y_1| < 0.5H \ \}.$$

We will show how the continuation equilibria after the signalling can be chosen to give the players incentives to randomize. We will also show that after the signalling phase player 2's posterior beliefs will still attach positive probability to type 1, and as $p_1 \to 1$ these posteriors give arbitrarily high probability to type 1. Thus, it is possible to choose p_1 sufficiently high for the equilibrium (described above) of the game with two types can be played after the signalling phase. We also show that the signalling strategies give the players payoffs close to (a_1, b).

Let $(\alpha_1^{k,j}, \beta^{k,j})$ denote the continuation equilibrium payoffs to type 1 and player 2 when player 1 signals type k and player 2 plays action j in the period the signal was sent. We will start in the final signalling period $t = K - 3$. We will choose the continuation equilibria in period $K - 3$ with payoffs that satisfy

$$(\alpha_1^{K,1}, \beta^{K,1}), (\alpha_1^{K-1,2}, \beta^{K-1,2}) \in U[(a_1^\dagger - \epsilon, b^\dagger - \epsilon); \epsilon, Y\epsilon], \tag{38}$$

$$(\alpha_1^{K,2}, \beta^{K,2}), (\alpha_1^{K-1,1}, \beta^{K-1,1}) \in U[(a_1^\dagger + \epsilon, b^\dagger + \epsilon); \epsilon, Y\epsilon], \tag{39}$$

where $(a_1^\dagger, b^\dagger)$ is chosen so that $U[(a_1^\dagger, b^\dagger); 3\epsilon, 3Y\epsilon] \subset U[(a_1, b); \iota, \iota]$. (Recall that $Y = 1$ when $\operatorname{Int} G_1(0) \neq \emptyset$, as assumed for the moment; however it will be convenient to retain the general notation for the case when $\operatorname{Int} G_1(0) = \emptyset$.) It is possible to choose such continuation equilibria, because the sets on the right of (38) and (39) are in $\operatorname{Int} G_1(\iota) \cap \{(x, y) | x < \bar{a}_1(3\epsilon) - C\epsilon\}$ and part 2 of the proof, therefore, applies. Continuation equilibria satisfying (38) and (39) can be found, because (by (17) and part 1) type 1's payoff can be approximated to within $\epsilon/16$ and by part 2 player 2's payoff can be approximated to within $\epsilon/2$. Given this choice of continuation equilibria in period $K - 3$ we will show that players' expected payoffs at the start of period $K - 3$ (potential continuation equilibria for period $K - 4$) lie in the set $U[(a_1^\dagger, b^\dagger); \epsilon\rho, Y\epsilon\rho]$, where $\rho = 1 + \frac{1}{8}$. This will furnish an inductive step. In period $K - 3$ type 1 randomizes between $i = 1$ and $i = 2$. Her payoffs from these actions are:

$$(i = 1) \qquad (1 - \delta) A_1(1, q^{K-3}) + \delta[q^{K-3} \alpha_1^{K,1} + (1 - q^{K-3}) \alpha_1^{K,2}],$$

$$(i = 2) \qquad (1 - \delta) A_1(2, q^{K-3}) + \delta[q^{K-3} \alpha_1^{K-1,1} + (1 - q^{K-3}) \alpha_1^{K-1,2}].$$

($A_1(i, q^{K-3})$ is an abuse that denotes type 1's stage-game payoff from action i when player 2 plays $(q^{K-3}, 1 - q^{K-3})$.) Player 1 is indifferent between these two actions if

$$\frac{1 - \delta}{\delta}[A_1(1, q^{K-3}) - A_1(2, q^{K-3})] = q^{K-3}[\alpha_1^{K-1,1} - \alpha_1^{K,1}] + (1 - q^{K-3})[\alpha_1^{K-1,2} - \alpha_1^{K,2}]. \tag{40}$$

Let $(\mu, 1-\mu)$ denote the probability player 1 plays $i=1$ and $i=2$ in period $K-3$ given the observed history. If we abuse our notation in a similar fashion as before, player 2 is indifferent between action $j=1$ and $j=2$ when

$$\frac{1-\delta}{\delta}[B(\mu,1)-B(\mu,2)] = \mu[\beta^{K,2}-\beta^{K,1}] + (1-\mu)[\beta^{K-1,2}-\beta^{K-1,1}]. \tag{41}$$

We can find $q^{K-3} \in [0,1]$ and $\mu \in [0,1]$ to make both players indifferent. First, the LHS of (40) is less than $\epsilon/16$ (by our assumption on δ) and the LHS of (41) is less than $Y\epsilon\frac{1}{16}$ in absolute value ($2M$ is the maximum variation in player 1's payoffs so $2YM$ is the maximum variation in player 2's). The assumption on the continuation equilibria implies that the RHS of (40) [respectively (41)] is a linear function of q^{K-3} [respectively μ] that includes in its range $-\epsilon$ [respectively $-Y\epsilon$] to ϵ [respectively $Y\epsilon$]. Thus there exist q^{K-3} and μ that solve (40) and (41). There are upper and lower bounds on the value of μ for which (41) holds. As the LHS is less than $Y\epsilon\frac{1}{16}$, the first square bracket on the RHS is in $(Y\epsilon, 3Y\epsilon)$ and the second is in the interval $(-3Y\epsilon, -Y\epsilon)$, we get $\frac{3}{4}+\frac{1}{64} > \mu > \frac{1}{4}-\frac{1}{64}$. Also, by taking the minimal and maximal continuation payoffs we can show that type 1's and player 2's expected payoffs at the start of $K-3$ lie in the set $U[(a_1^\dagger, b^\dagger); \epsilon\rho, Y\epsilon\rho]$, where $\rho = 1+\frac{1}{8}$.

The paragraph above describes potential continuation equilibria after period $K-4$ of the signalling phase (assuming type $K-2$ is not signalled). We will use this to describe an equilibrium for period $K-4$ onward with payoffs in $U[(a_1^\dagger, b^\dagger); \epsilon\rho^2, Y\epsilon\rho^2]$, provided

$$U[(a_1^\dagger, b^\dagger); (2+\rho+\rho^2)\epsilon, S(2+\rho+\rho^2)\epsilon] \subset U[(a_1, b); \iota, \iota]. \tag{42}$$

To build this equilibrium it is first necessary to describe behaviour in period $K-3$. Repeat the argument of the previous paragraph with the sets in (38) and (39) replaced by $U[(a_1^\dagger, b^\dagger) - (\epsilon\rho, Y\epsilon\rho) \pm (\epsilon, Y\epsilon); \epsilon, Y\epsilon]$, to find a period $K-3$ equilibrium with payoffs in $U[(a_1^\dagger, b^\dagger) - (\epsilon\rho, Y\epsilon\rho); \epsilon\rho, Y\epsilon\rho]$ ((42) is sufficient for this to be possible). This is the equilibrium played if $(i,j)=(1,1)$ in period $K-4$. A similar procedure can be followed to find a period $K-3$ equilibrium with payoffs in $U[(a_1^\dagger, b^\dagger) + (\epsilon\rho, Y\epsilon\rho); \epsilon\rho, Y\epsilon\rho]$ and again (42) is sufficient; this is played if $(i,j)=(1,2)$ in period $K-4$. If player 1 plays $i=2$ in period $K-4$ we can use the argument in part 2 and (42) to find two continuation equilibria of the game with the types $\{1, K-2\}$ with payoffs in $U[(a_1^\dagger, b^\dagger) - (\epsilon\rho, Y\epsilon\rho); \epsilon\rho, Y\epsilon\rho]$ and $U[(a_1^\dagger, b^\dagger) + (\epsilon\rho, Y\epsilon\rho); \epsilon\rho, Y\epsilon\rho]$, which are played when (i,j) equals respectively $(2,2)$ or

(2, 1) in period $K-4$. Now consider the randomizations in period $K-4$. We can apply the argument of the previous paragraph to show that the probability player 1 randomizes is again in $[\frac{1}{4}-\frac{1}{64}, \frac{3}{4}+\frac{1}{64}]$ and that type 1's and player 2's period $K-4$ expected equilibrium payoffs are in $U[(a_1^\dagger, b^\dagger); \epsilon\rho^2, Y\epsilon\rho^2]$. (It is necessary to replace ϵ by $\epsilon\rho$.)

Now we can iterate this argument working backwards to the first round of signalling at time zero — all the time getting bounds on player 1's randomization. When there are $K-2$ periods of signalling it is necessary to be able to find equilibria in period $K-3$ that lie in the sets $U[(a_1^\dagger, b^\dagger) \pm (1+\rho+...+\rho^{K-3})(\epsilon, Y\epsilon); \epsilon, Y\epsilon]$. This is possible if $(a_1, b) = (a_1^\dagger, b^\dagger)$, (v) holds and $U[(a_1, b); \iota, \iota] \subset G_1(\iota) \cap \{(x, y)|x < \bar{a}_1(3\epsilon) - C\epsilon\}$. The construction of the signalling phase ensures period zero's expected payoffs are in the interval $U[(a_1, b); \epsilon\rho^{K-2}, Y\epsilon\rho^{K-2}] \subset U[(a_1, b); \iota, \iota]$.

When $\text{Int}\, G_1(0) = \emptyset$ the above argument will work virtually unchanged, because of the inclusion of Y. However, it is necessary to replace the open rectangles $U[(a_1, b); x, Yx]$ with the open line segment between the points $(a_1, b) \pm 0.5(x, Yx)$ (this is the diagonal of the rectangle above). By the definition of Y, this lies in the feasible set and replaces the open rectangles as a measure of a neighbourhood in the one dimensional set.

The construction gives type 1 and player 2 period-zero expected payoffs in the set $U[(a_1, b); \iota, \iota]$. We must check that in all the continuation equilibria p_1 is sufficiently large. Given the lower bounds on player 1's probabilities derived above, each possible history of player 1's signalling-phase actions occurs with at least probability $(\frac{1}{4}+\frac{1}{64})^{K-1}$. Provided $p_k < \underline{r}'(\frac{1}{4}+\frac{1}{64})^{K-1}$ it is possible to apply Part 2 of the proof and play continuation equilibria satisfying (38) and (39). The required lower bound on p_1 is thus $1-\underline{r}'(\frac{1}{4}+\frac{1}{64})^{K-1}$ (since this implies $p_k < \underline{r}'(\frac{1}{4}+\frac{1}{64})^{K-1}$ for all $k > 1$).

We now show that no player wishes to deviate from her/his equilibrium strategies in the equilibrium with many types. As argued, under the assumption on δ and (a_1, b) player 2's continuation payoff is within ι of b during the entire signalling phase and hence greater than $\hat{b}+\iota$, whereas a deviation yields at most $\hat{b}+\epsilon/2$, which by $\epsilon < \iota/2$ is thus unprofitable. Thereafter, whichever types are signalled player 2 does not benefit from deviating by Lemma 4. A similar argument coupled with part 2 of this proof ensures that type 1 does not benefit by deviating from the strategies described above and neither does type k

benefit by deviating when she has signalled that she is type k, because the losses during the signalling phase are sufficiently small. The four possible extra deviations that can arise when there are many types are: type k mimics type k' (unobservable), type k mimics type k' and then deviates to take a punishment (unobservable then observable), type k mimics type k' and later she plays $\tilde{\imath}$ and then mimics type 1 at a revealing equilibrium (unobservable), or type k mimics type k', later she plays $\tilde{\imath}$ and then mimics type 1 before finally deviating from the revealing equilibrium to take a punishment (unobservable then observable). We will begin by showing that these deviations are not profitable when the strategy of type k' is to play the strategy described and then treat the case when the amended strategies are followed, as described in part 1 of the proof. Suppose type k sends the signal of type k' and then plays out her finite sequence N' times before settling at the equilibrium described in Lemma 5. From (29) her payoff from this, discounted to the period after the signalling is finished, is $(1-\delta^{TN'})\hat{A}_{k,k'} + \delta^{TN'}\bar{\alpha}_k$, whereas her payoff from playing her equilibrium strategy can be written as $(1-\delta^{TN})\hat{A}_{k,k} + \delta^{TN}\bar{\alpha}_k$. At an equilibrium type 1 will mimic type k and type k' with positive probability. Let c be type 1's expected equilibrium payoff from mimicking type k and c' be her expected payoff from mimicking type k', that is,

$$c = (1-\delta^{TN})\hat{A}_{1,k} + \delta^{TN}\bar{\alpha}_1 = (1-\delta^{TN})(\hat{A}_{1,k} - \bar{\alpha}_1) + \bar{\alpha}_1; \tag{43}$$

$$c' = (1-\delta^{TN'})\hat{A}_{1,k'} + \delta^{TN'}\bar{\alpha}_1 = (1-\delta^{TN'})(\hat{A}_{1,k'} - \bar{\alpha}_1) + \bar{\alpha}_1. \tag{44}$$

The following will be a sufficient condition to rule out the first form of deviation described above:

$$(1-\delta^{TN})\hat{A}_{k,k} + \delta^{TN}\bar{\alpha}_k > (1-\delta^{TN'})\hat{A}_{k,k'} + \delta^{TN'}\bar{\alpha}_k + 2\epsilon,$$

or equivalently

$$(1-\delta^{TN})(\hat{A}_{k,k} - \bar{\alpha}_k) > (1-\delta^{TN'})(\hat{A}_{k,k'} - \bar{\alpha}_k) + 2\epsilon,$$

or

$$\frac{\hat{A}_{k,k} - \bar{\alpha}_k}{\bar{\alpha}_1 - \hat{A}_{1,k}}(\bar{\alpha}_1 - c) > \frac{\hat{A}_{k,k'} - \bar{\alpha}_k}{\bar{\alpha}_1 - \hat{A}_{1,k'}}(\bar{\alpha}_1 - c') + 2\epsilon,$$

where the last inequality follows from substitution for $(1-\delta^{TN})$ from (43) and for $(1-\delta^{TN'})$ from (44). Type 1 randomizes between mimicking type k and type k' in equilibrium. The signalling phase payoff plus c and the signalling phase payoff plus c' give type 1 identical payoffs. The signalling phase payoffs contribute at most $\epsilon/2$, so $|c - c'| < \epsilon$. Also (35) applies, so the above inequality holds and it is optimal for type k to play her equilibrium strategy. We can now consider the second form of deviation. Suppose that type k mimics type k' and then deviates (before N' iterations are played) when type 1's continuation payoff is c. The strategies described in part 1 of the proof impose the same punishment on type k as the punishment she would have received if she had truthfully signalled her type and then deviated when type 1's continuation payoff was c (she can get the same deviation payoff by signalling truthfully). A repetition of the above argument shows that this latter option is strictly preferred to the former, and hence *a fortiori* type k prefers to use her equilibrium strategy. If the third type of deviation gives type k more than her equilibrium payoff a small emendation of the above strategies restores an equilibrium. To do this replace type k's strategy with her mimicking player k' and then playing $\tilde{\imath}$ in this way and remove the stage of the signalling phase where type k is signalled. This new equilibrium increases player 2's expected payoff when type k' is signalled and so will increase his willingness to abide by his equilibrium strategy (if there are more than two types for which this deviation is profitable, each type can likewise play the signal which she prefers). If the fourth type of deviation is optimal then type k must benefit from an observable deviation from the equilibrium of the complete information game after $\tilde{\imath}$ was signalled. In this case the argument in parentheses at the end of part 1 of this proof applies *mutatis mutandis*.

Now we must deal with the amended strategies and consider what occurs if type k' at some point plays a pooling equilibrium with type 1, rather than continuing to reveal her type. (This change was introduced at the end of part 1 of the proof.) If type k' and type 1 play the pooling equilibrium, then the possible deviations available to type k mimicking type k' or type 1 were available to her above also. Thus the argument above applies to this case also.

Now we return to the condition (34), that has been assumed to hold. This condition guaranteed that the types $k > 1$ *strictly* preferred to play the iterations of their finite sequence, $\{(\hat{\imath}_k^s, \hat{\jmath}_k^s)\}$, rather than another type's sequence, before settling on the terminal equilibrium. (This condition will fail if, for example, the payoffs of type k are a linear transformation of the payoffs of type k' and so $\underline{\pi}_k = \underline{\pi}_{k'}$.) Suppose, now, that there exist

k and k' so that

$$\frac{A_k(\underline{\pi}_k) - \bar{a}_k}{\bar{a}_1 - A_1(\underline{\pi}_k)} = \frac{A_k(\underline{\pi}_{k'}) - \bar{a}_k}{\bar{a}_1 - A_1(\underline{\pi}_{k'})}. \tag{45}$$

In this case we can choose $\underline{\pi}_k = \underline{\pi}_{k'}$ and the sequence $\{(\hat{\imath}^s_k, \hat{\jmath}^s_k)\}$ to be the same as $\{(\hat{\imath}^s_{k'}, \hat{\jmath}^s_{k'})\}$. A small change to the above strategies restores an equilibrium. Change type k's equilibrium strategy so that she plays exactly the same actions as type k' until the final playing of the equilibrium described in Lemma 5, that is, so that both k and k' signal at the same time (and in the same way) and so that the period in the signalling phase where type k was signalled is removed. Note that conditions (a)-(c) of Lemma 6 still apply when $\underline{\pi}_k$ is replaced by $\underline{\pi}_{k'}$, so the previous argument can be repeated *mutatis mutandis.* Any remaining indifferences can be handled in exactly the same way.

Let $R(\iota)$ denote the set of points (a_1, b) in the relative interior of $G_1(\iota) \cap \{(x,y)|x < \bar{a}_1(3\epsilon) - C\epsilon\}$ that are distance at least ι from the boundary of the relative interior of $G_1(\iota) \cap \{(x,y)|x < \bar{a}_1(3\epsilon) - C\epsilon\}$. We have shown that there exists a $\delta_\iota < 1$ and $p_1^\iota < 1$ such that for all $\mathbf{p}$ with $p_1 > p_1^\iota$ and $\delta > \delta_\iota$, given any $(a_1, b) \in R(\iota)$ the game $\Gamma(\mathbf{p}, \delta)$ has an equilibrium with payoffs that satisfy $\|(\alpha_1, \beta) - (a_1, b)\| < \iota$. By choosing $\iota < \nu$ and sufficiently small the Theorem follows.

Q.E.D..

References

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.

Cripps, M. W., K. M. Schmidt, and J. P. Thomas (1996): "Reputation in Perturbed Repeated Games," *Journal of Economic Theory*, 69, 387–410.

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 (1992): "Maintaining a Reputation when Strategies are Imperfectly Observed," *Review of Economic Studies*, 59, 561–579.

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.

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.

Lehrer, E., and L. Yariv (1999): "Repeated Games with Incomplete Information on One Side: The Case of Different Discount Factors," *Mathematics of Operations Research,* 24, 204–218.

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. (1999): "Merging, Reputation and Repeated Games with Incomplete Information," *Games and Economic Behavior,* 29, 274–308.

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.