MATHEMATICS OF OPERATIONS RESEARCH
Vol. 00, No. 00, xxx xxxx, pp. 1–30
Printed in U.S.A.

SOME ASYMPTOTIC RESULTS IN DISCOUNTED REPEATED GAMES OF ONE-SIDED INCOMPLETE INFORMATION

MARTIN W. CRIPPS AND JONATHAN P. THOMAS

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.

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 §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 a "fully revealing" equilibrium—one in which the informed player acts to reveal her information at the start of the game. In such an equilibrium, the play (probability distribution over paths) induced by the strategy of *each* type of Player 1 against Player 2's strategy must yield individually rational payoffs to Player 2. (The precise statement of this requires the use of Player 1's discount factor in the evaluation of Player 2's payoffs.) This is potentially a much stronger restriction on the set of equilibria than the condition that *average* play—averaging across Player 1's types using Player 2's prior beliefs—should satisfy individual rationality. This latter condition must hold in any equilibrium, and it also (trivially) holds in the complete information game between a particular type k and Player 2, where, combined with the condition that play must be individually rational for type k of Player 1, is essentially the only restriction on equilibrium play (by the Folk Theorem). Depending on the game, the former type-by-type condition can imply major restrictions on equilibrium payoffs of an incomplete information game relative to the corresponding (for each type) complete information game. This result implies a continuity result with the undiscounted case (the continuity property is not uniform with respect to initial beliefs). 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, because the latter has equilibrium payoff equivalence to fully revealing equilibria (Shalev 1994, see §3 for a precise statement). This contrasts with the Folk Theorem applied to the complete information game involving type k. (See, e.g., Forges 1992, for an example and precise statement of how perturbing an undiscounted complete information

Received August 1, 1998; revised June 1, 2000; July 1, 2001; and December 1, 2002.
MSC 2000 subject classification. Primary: 91A05, 91A20.
OR/MS subject classification. Primary: Games/group decisions.
Key words. Reputation, Folk Theorem, repeated games, incomplete information.

0364-765X/00/0000/0001/$05.00
1526-5471 electronic ISSN,

game by introducing a small probability of an alternative type of one of the players can lead to a large reduction in the set of payoffs that player can receive in equilibrium. A similar example is developed below in §4.)

In §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.

Very roughly, the difference between the two cases can be explained as follows. If the uninformed player is very patient relative to the informed, then the period of learning of the uninformed player will be unimportant in the calculation of the informed player's payoffs; from the point of view of the latter it is as if information is revealed early on in play and the equilibrium must approximately satisfy conditions of a fully revealing equilibrium. If the two players are equally patient, however, the period of learning can always be used, if necessary, to drive the payoff of one of the types of the informed player down towards her individually rational payoff, while rewarding Player 2 to avoid his individual rationality constraint from binding. (When there is no discounting, again, the period of learning has no effect on payoffs.) 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.

The situation where one or more players' preferences may be unknown to the opponent(s) has received 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 of incomplete information with known-own payoffs have, however, been studied in some depth (see §3). Some recent results exist for the discounted case, however. Kalai and Lehrer (1993) and Jordan (1995) have established that play, in a given state, must converge to Nash play of the complete information game played between the realized types in that state. Sorin (1999) provides a synthesis of a number of the results in this literature. 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 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. (This is a game of "known own payoffs.") 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 $\sigma_k = (\sigma_k^0, \sigma_k^1, \ldots)$, $\sigma_k^t \colon H^{t-1} \to \Delta^I$. We define $\sigma = (\sigma_k)_{k \in K}$. Likewise, a strategy for Player 2 is a sequence of maps $\tau = (\tau^0, \tau^1, \ldots)$, $\tau^t \colon H^{t-1} \to \Delta^J$. 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. 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 that 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 $\Gamma(\mathbf{p}, 1, 1)$. In Theorem 1 we shall show that essentially the same characterization as that of Shalev (1994) 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 $\mathbf{p} \in \Delta^K$, let $a(\mathbf{p})$ be Player 1's minmax payoff in the one-shot game with payoffs given by $\sum_{k \in K} p_k A_k(i, j)$. The set of payoffs $\{\mathbf{y} \in \Re^K \mid \mathbf{y} \leq \mathbf{x}\}$ is said to be *approachable* by Player 2 if and only if

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

Later, we use a sufficient condition for (1); this is $\mathbf{p} \cdot \mathbf{x} \geq \operatorname{Cav} a(\mathbf{p})$, where $\operatorname{Cav} a(\mathbf{p})$ is the (pointwise) smallest concave function $g(\mathbf{p})$ satisfying $g(\mathbf{p}) \geq a(\mathbf{p})$. 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 $\{\mathbf{y} \mid \mathbf{y} \leq \mathbf{x}\}$ 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 $\{\mathbf{y} \mid \mathbf{y} \leq \mathbf{x}\}$ 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

$$y \geq \hat{b}. \tag{2}$$

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 $\sum_{k\in K} p_k B(\pi_k) = b$. In other words, equilibria are payoff equivalent to equilibria in which Player 1 acts to reveal the true state at the start of the game. This requires that $B(\pi_k)$ is individually rational for Player 2 for each $k \in K$, as once Player 2 is aware of the state, play, as summarised by π_k, must yield Player 2 at least his minmax payoff, otherwise, he could profitably deviate.

We are now in a position to state Theorem 1: 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). We define

$$\mathbf{A}^* = \left\{(A_1(\pi_1), A_2(\pi_2), \ldots, A_K(\pi_K))\colon (\pi_k)_{k\in K} \in \Pi_0\right\}. \tag{3}$$

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*

$$\min_{\mathbf{x}\in\mathbf{A}^*} \|\mathbf{a} - \mathbf{x}\| < \epsilon. \tag{4}$$

The main ancillary result used to establish this 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. For a fixed equilibrium of $\Gamma(\mathbf{p}, \delta_1, \delta_2)$, we define the average frequencies over action profiles conditional on type k using discount factor δ as: $\pi_k^{ij}(\delta) = (1-\delta)E[\sum_{t=0}^{\infty} \delta^t \mathbf{1}\{i, j, t\} \mid k]$, 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 $E[\tilde{a}_k \mid k] = A_k(\pi_k(\delta_1))$ for each k and $E[\tilde{b}] = \sum_{k\in K} p_k B(\pi_k(\delta_2))$. 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 $y^N = ((i^{t+1}, j^{t+1}), \ldots, (i^{t+N}, j^{t+N}))$. For given equilibrium strategies (σ, τ), we let $\mathbf{q}^N(\cdot \mid h^t)$ be the distribution over these outcomes (i.e., $\mathbf{q}^N(y^N \mid h^t) = \text{prob}[h^{t+N} = (h^t, y^N) \mid h^t]$, using obvious notation; it is defined for h^t having positive probability) 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 $\hat{\mathbf{q}}^N$, $\|\mathbf{q}^N - \hat{\mathbf{q}}^N\| :=$

$\max_{y^N} |\mathbf{q}^N(y^N) - \hat{\mathbf{q}}^N(y^N)|$. 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 $\tilde{b}^{t+1} := (1-\delta_2)\cdot \sum_{r=t+1}^{\infty} \delta_2^{r-t-1} B(i^r, j^r)$. The proof of the following is straightforward and is omitted.

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*

$$E[\tilde{b}^{t+1} \mid h^t, k] \leq \hat{b} - \epsilon. \tag{5}$$

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

$$\|\mathbf{q}^N(\cdot \mid h^t) - \mathbf{q}^N(\cdot \mid h^t, k)\| > \eta. \tag{6}$$

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 $T(n, N) = \{n, n+N, n+2N, \dots\}$. 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) the probability, conditional on Player 1's true type being k, that there are more than m periods $t \in T(n, N)$ with

$$\|\mathbf{q}^N(\cdot \mid h^t) - \mathbf{q}^N(\cdot \mid h^t, k)\| > \psi \tag{7}$$

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*

$$B(\pi_k(\delta_1)) \geq \hat{b} - \phi. \tag{8}$$

PROOF. 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 (6) holds whenever (5) holds. Set $\psi = \eta$ in Result 1, take any integer n, $0 \leq n < N$, and set $\xi = (\phi/3N(\hat{b} - b_{\min}))$ (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 (6) holds more than m times in $T(n, N)$ is less than ξ, so the probability that inequality (5) 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

$$E[\tilde{b}^{t+1} \mid h^t, k] \leq \hat{b} - \frac{1}{3}\phi \tag{9}$$

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

$$(10) \qquad B(\pi_k(\delta_1)) = (1-\delta_1)\sum_{t=0}^{\infty} \delta_1^t E[B(i^t, j^t) \mid k] = \frac{1-\delta_1}{1-\delta_2}\sum_{t=0}^{\infty} \delta_1^t E[\tilde{b}^t - \delta_2 \tilde{b}^{t+1} \mid k]$$

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

Using the result on the number of times (9) holds, for $\delta_1 > \delta_2$ the random variable $\sum_{t=0}^{\infty} E[\delta_1^t(\delta_1 - \delta_2)\tilde{b}^{t+1} \mid h^t, k] \geq \{((\delta_1 - \delta_2)/(1-\delta_1))(\hat{b} - (\phi/3)) - (\delta_1 - \delta_2)(\hat{b} - b_{\min})Nm\}$ with probability at least $(1 - N\xi)$ conditional on k, where we are using the fact that in the event that (9) 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 $((\delta_1 - \delta_2)/(1-\delta_1))b_{\min}$, otherwise. Using this in (10) gives a lower bound, say $\Omega(\delta_1, \delta_2)$, so that $B(\pi_k(\delta_1)) \geq \Omega(\delta_1, \delta_2)$, and notice that $\Omega(\delta_1, \delta_2)$ is independent of the particular equilibrium studied. Next, taking the limit as $\delta_1 \to 1$ yields $\lim_{\delta_1 \to 1} \Omega(\delta_1, \delta_2) = (1 - N\xi)(\hat{b} - (\phi/3)) + N\xi b_{\min}$; hence, since $N\xi = (\phi/3(\hat{b} - b_{\min}))$, we get

$$(11) \qquad \lim_{\delta_1 \to 1} \Omega(\delta_1, \delta_2) = \hat{b} - \frac{\phi}{3} - \frac{\phi}{3(\hat{b} - b_{\min})}\left(\hat{b} - b_{\min} - \frac{\phi}{3}\right)$$

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

Choosing $\underline{\delta}_1^{(k)}$ such that $\Omega(\delta_1, \delta_2)$ is within $\phi/3$ 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. □

We are now in a position to prove Theorem 1.

Proof of Theorem 1. 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,

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

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 $\boldsymbol{\sigma}^*$ the repeated game strategy in which Player 1 plays the mixed action f^* each period and independently of type k. Then $E_{\mathbf{p}, \sigma^*, \tau}[(1-\delta_1) \cdot \sum_{t=0}^{\infty} \delta_1^t \sum_k q_k^* A_k(i^t, j^t)] > \mathbf{q}^* \cdot \mathbf{a}$ (NB. k is not a random variable), so that

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

because 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 $E_{\mathbf{p}, \sigma^*, \tau}[\tilde{a}_k \mid k] > a_k$ 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

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

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

$$\Psi(\phi) = \left\{(\pi_k)_{k\in K} \mid B(\pi_k) \geq \hat{b} - \phi \text{ for all } k \in K\right\}.$$

Since Ψ is an upper hemicontinuous function of ϕ, it follows that the correspondence given by $\Psi \cap \widehat{\Pi}$, which is nonempty (Shalev 1994), is also upper hemicontinuous. 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 $\mathbf{A}[\Psi(\phi) \cap \widehat{\Pi}]$ is an upper hemicontinuous 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 $\mathbf{A}[\Psi(\phi) \cap \widehat{\Pi}]$ 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 (4) to hold. □

Theorem 1 developed necessary conditions which equilibrium payoffs must satisfy asymptotically. In the undiscounted model, the condition that play must correspond to a point in Π_0 is necessary *and sufficient* for equilibrium (Shalev 1994). Theorem 1 established that in the discounted game it is *necessary* that equilibrium play (averaged using δ_1) approximately satisfy the same condition when Player 1 is sufficiently patient. A partial converse is provided by the following, where it is assumed that the inequalities in the conditions of Definition 1 are assumed to hold strictly. We say that a payoff vector $\mathbf{a}$ is *strictly individually rational* for Player 1 if there exists some individually rational point $\mathbf{x}$ with $a_k > x_k$ for all k.

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) that 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 that arises is the punishment of Player 1; see §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 games in this class by $\Gamma(\mathbf{p}, \delta) := \Gamma(\mathbf{p}, \delta, \delta)$. In Theorem 3 we show 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, which we denote by $\Gamma_1(\delta)$, 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.

The definition in (1) of individual rationality given in §3 applies to Player 1's undiscounted payoffs and defined sets of approachable payoffs. In discounted games, as the players become more patient, Player 2 has a strategy that holds Player 1 to within $\epsilon > 0$ of a set of approachable payoffs.

DEFINITION 2. $\mathbf{x} = (x_k)_{k\in K} \in \Re^K$ is ϵ-individually rational (ϵ-IR) if the set $\{\mathbf{y} \in \Re^K \mid y_k + \epsilon \leq x_k \ \forall k \in K\}$ is approachable.

$\operatorname{Cav} a(\mathbf{p})$ 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). 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 $0 \leq v_\delta(\mathbf{p}) - \operatorname{Cav} a(\mathbf{p}) \leq M\sqrt{\{(K-1)(1-\delta)/(1+\delta)\}}$ (by Zamir 1992). 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).

RESULT 2. For any $\epsilon > 0$ there exists $\delta_\epsilon < 1$, so that for any $\delta > \delta_\epsilon$ Player 2 has a strategy that can hold Player 1 down any ϵ-IR payoff in $\Gamma(\mathbf{p}, \delta)$.

We shall assume that we can find strictly individually rational payoffs for the repeated game of incomplete information $\Gamma(\mathbf{p}, \delta)$.

ASSUMPTION 1. *There exists* $(\breve{\pi}_1, \breve{\pi}_2, \ldots, \breve{\pi}_K) \in (\Delta^{IJ})^K$ *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* $k \in K$.

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 Assumption 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 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. In the complete information games between type k and Player 2, we define $G_k(\epsilon)$ to be the set of feasible (uniformly *strictly* for $\epsilon > 0$) individually rational payoffs.

$$G_k(\epsilon) := \left\{(A_k(\pi), B(\pi)) \mid A_k(\pi) \geq \hat{a}_k + \epsilon, B(\pi) \geq \hat{b} + \epsilon, \pi \in \Delta^{IJ}\right\}, \quad k \in K. \tag{14}$$

It is now possible to state the main result of this section.

THEOREM 3. *Let Assumption* 1 *and* $\nu > 0$ *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* $(a_1, b) \in G_1(0)$ *the game* $\Gamma(\mathbf{p}, \delta)$ *has an equilibrium with the payoffs* $((\alpha_1, \ldots, \alpha_K), \beta) \in \Re^{K+1}$ *which satisfy*

$$\|(\alpha_1, \beta) - (a_1, b)\| < \nu. \tag{15}$$

4.0. Example. As an illustration, we consider an example. In this example, Shalev's (1994) results (discussed in §3) imply that there is a lower bound on type 1's equilibrium payoff in the *undiscounted* case strictly above her minmax payoff of 3/4 (see Forges 1992, Proposition 8.3, for a general statement of this result); 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 $A_1(\pi_1) \geq A_1(\pi_2) \geq 21/20$.

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

(A_1, B)

	L	R
T	1.2, 1	1, 0
B	0, 0	0, 3

(A_2, B)

Let $\epsilon > 0$ be given. In what follows, type 2 of Player 1 will play T on all equilibrium paths. Consider first the following (pooling) equilibrium of $\Gamma(\mathbf{p}, \delta)$: Both types of Player 1 play T and Player 2 plays L in every period, irrespective of past history. Player 1 gets $(3, 1.2)$ 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) is played to reduce type 1's payoff and, in general, will need to be replaced by a finite sequence). 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; 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 3 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, 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 (only) plays B with probability q such that

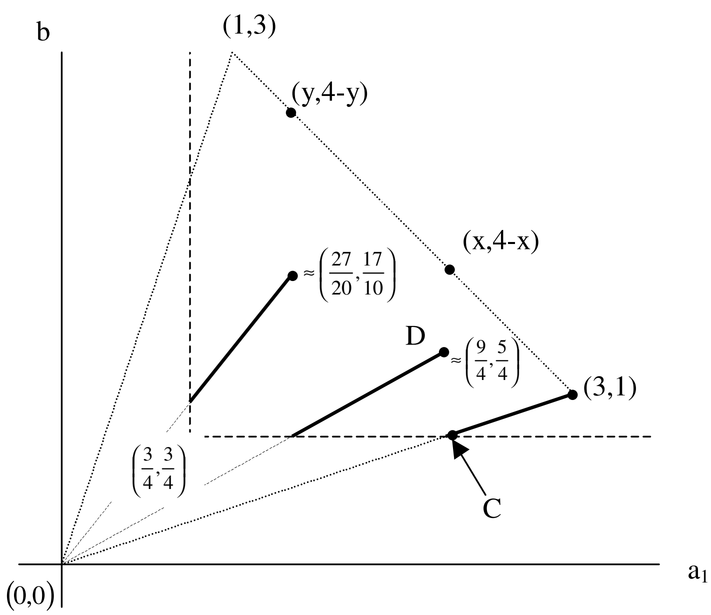

FIGURE 1. Payoffs to type 1 and Player 2.

$p_1 q = 0.5$, which is possible provided $p_1 > 0.5$, 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 is type 1, then from the start of the following period an equilibrium of the complete information game 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 individual rationality 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: $\delta^{N'}(2-\delta^N) \geq (1-\delta)3+\delta(3/4)$. Repeating the argument given earlier, for $\delta > \delta^*(\epsilon)$, $\delta^{N'}(2-\delta^N)$ is within $\epsilon/3$ of $3/4$, and type 1's payoff $3\delta^{N+N'} \leq 3(3/4+\epsilon/3)(3/5+32\epsilon/5(15-4\epsilon))$. 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 $\tilde{\epsilon} > 0$ such that for $\epsilon < \tilde{\epsilon}$ payoffs lie above the 45° line. Thus by choosing ϵ small enough, type 1 can be held as close to $27/20$ as desired provided $\delta > \delta^*(\epsilon)$. (It can easily be checked that there are no profitable deviations.) *This is strictly lower than the lowest payoff in the zero discounting game.*

A further repetition of the argument, so that another randomization (involving payoffs $(y, 4-y)$ in the figure) with more plays of (T, R) 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, and type 1 can be held as close to $3/4$ as desired provided $p_1 \geq 3/4$ (see Figure 1). 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 $\Gamma_1(\delta)$ gives Player 2 close to the desired payoffs. Provided type 1's probability is sufficiently close to 1, this will satisfy any desired degree of approximation.

In the generalization of the example which follows, we shall split the above construction into three 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 $\Gamma_1(\delta)$, the complete information game played by type 1 and Player 2. In Lemma 4, we construct a particular type of equilibrium where any feasible and strictly individually rational payoff to type 1 can be obtained as an equilibrium payoff. This will consist of a continuation equilibrium, in which type 1 receives a high payoff, preceded by play which yields type 1 a low payoff; by extending this latter phase of play, the overall payoff will be reduced toward any desired target payoff. It may be, however, that this process violates Player 2's individual rationality; each time this is threatened, a randomization by Player 1 is used to probabilistically reward Player 2 so the latter has sufficient incentive to stick to this path. In §4.2 we shall use these equilibrium strategies to construct an equilibrium of a two-type incomplete information game.

Some additional notation on payoffs is now necessary. Let M denote an upper bound on the absolute magnitude of the players' payoffs, that is, $M \geq |A_k(i, j)|, |B(i, j)|$, for all (i, j) and k. We define Player 1's largest and smallest payoffs in the sets of individually rational payoffs in (14):

$$\bar{a}_k(\epsilon) := \max_{(a_k, b)\in G_k(\epsilon)} a_k,$$

$$\underline{a}_k(\epsilon) := \min_{(a_k, b)\in G_k(\epsilon)} a_k,$$

$$\bar{\mathbf{a}} := (\bar{a}_1(0), \ldots, \bar{a}_K(0)).$$

Note that the function $\bar{a}_k(\cdot)$ (respectively, $\underline{a}_k(\cdot)$) maximizes (minimizes) a linear function on a set of linear inequalities that vary continuously in ϵ. $\bar{a}_k(\cdot)$ ($\underline{a}_k(\cdot)$) is, therefore, continuous in a neighborhood of zero. We will use $f\colon [\underline{a}_1(0), \bar{a}_1(0)] \to \Re$ to denote the maximum feasible payoff to Player 2

$$f(a_1) := \max\{b \mid (a_1, b) \in G_1(0)\}. \tag{16}$$

The function $f(\cdot)$ 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) also define $-s$ to be the greatest negative slope of $f(\cdot)$ when $f(\cdot)$ has a decreasing segment (so $s > 0$) and $s = 1$ otherwise.

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

RESULT 3. Let $\epsilon > 0$ be given. There is a $\hat{\delta}(\epsilon) < 1$ such that if $\delta > \hat{\delta}(\epsilon)$ and given any $\pi \in \Delta^{IJ}$, then there exists a sequence of actions $\{(i^t, j^t)\}_{t=0}^{\infty}$ such that: $A_k(\pi) = (1-\delta)\cdot \sum_{t=0}^{\infty} \delta^t A_k(i^t, j^t)$, for all $k \in K$, and $B(\pi) = (1-\delta)\sum_{t=0}^{\infty} \delta^t B(i^t, j^t)$; moreover

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

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

The proof of Result 3 can be adapted from the proof of Lemma 2 in Fudenberg and Maskin (1991). It follows immediately that $\underline{\delta}(\epsilon)$ exists in the following definition:

DEFINITION 3. Given $\epsilon > 0$, define $\underline{\delta}(\epsilon) \ge \hat{\delta}(\epsilon)$ to be such that any $(a_k, b) \in G_k(\epsilon)$ are sustainable as equilibrium payoffs for any $\delta > \underline{\delta}(\epsilon)$.

Now, we define the strategies $\hat{\sigma}(n; a_1^*, b^*, x, \hat{u})$ and $\hat{\tau}(n; a_1^*, b^*, x, \hat{u})$, which will be used to construct an equilibrium in which a single randomization occurs.

DEFINITION 4. Let: $\epsilon > 0$, a sequence $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$, an $(a_1^*, b^*) \in G_1(3\epsilon)$ and $(x, f(x)) \in G_1(2\epsilon)$ be given. Then, for $\delta > \underline{\delta}(\epsilon)$, $\hat{u} \in (0, 1]$, the strategy profile $(\hat{\sigma}(n; a_1^*, b^*, x, \hat{u}), \hat{\tau}(n; a_1^*, b^*, x, \hat{u}))$ is defined as follows (suppressing dependence on δ, $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$):

$\underline{\hat{\sigma}(n; a_1^*, b^*, x, \hat{u})}$: In period 0, play $\hat{\imath}^0$ with probability $\hat{u}$ and $\tilde{\imath} \ne \hat{\imath}^0$ with probability $1-\hat{u}$. 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 3, to get the payoffs $(x, f(x)) \in G_1(2\epsilon)$. Minmax Player 2, thereafter, if a zero probability action is taken.

$\underline{\hat{\tau}(n; a_1^*, b^*, x, \hat{u})}$: 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 3, to get the payoffs $(x, f(x)) \in G_1(2\epsilon)$. Minmax Player 1, thereafter, if a zero probability action is taken.

Define payoffs when there are n complete rounds of the sequence to be played as

$$a_1(n) := (1-\delta^{nT})\widehat{A}_1 + \delta^{nT} a_1^*; \qquad b(n) := (1-\delta^{nT})\widehat{B} + \delta^{nT} b^*.$$

LEMMA 3. *Let $\epsilon > 0$, $\hat{u} \in (0,1]$ be given; also let $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ and $\delta^*(\epsilon) < 1$ be so that $\widehat{A}_1 := ((1-\delta)/(1-\delta^T))\sum_{s=0}^{T-1}\delta^s A_1(\hat{\imath}^s, \hat{\jmath}^s) < \hat{a}_1 + \epsilon$ 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 $\delta > \max\{\underline{\delta}(\epsilon), \delta^*(\epsilon), [4M/(\epsilon+4M)]^{1/T}\}$ and $n \geq 1$ is the largest integer satisfying*

$$b(n) > \hat{b} + 2\epsilon, \tag{17}$$

$$a_1(n) > \underline{a}_1(2\epsilon) + \epsilon/2; \tag{18}$$

then there exists $(x, f(x)) \in G_1(2\epsilon)$ so that $(\hat{\sigma}(n; a_1^*, b^*, x, \hat{u}), \hat{\tau}(n; a_1^*, b^*, x, \hat{u}))$ is an equilibrium of $\Gamma_1(\delta)$.

PROOF. We will first show that $n \geq 1$. When $n = 1$ we have

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

For $n = 1$ (18) holds, because the bottom line is positive by our choice of a_1^* and δ which implies $(1-\delta^T)2M < (1/2)\epsilon$. A similar argument shows $b(1) > \hat{b} + 2\epsilon$.

The strategies are an equilibrium of $\Gamma_1(\delta)$ 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. 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

$$x = \frac{a_1(n)}{\delta} - \frac{(1-\delta)}{\delta} A_1(\tilde{\imath}, \hat{\jmath}^0). \tag{19}$$

However, (19) implies that $|a_1(n) - x| < 2M(1-\delta)/\delta < \epsilon/2$, where the last inequality follows from our choice of 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 (19).

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 asserted in the Lemma. A similar 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. □

In the next lemma, we start with an equilibrium of $\Gamma_1(\delta)$ 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 use Lemma 3 to find a new equilibrium with the payoffs $(a_1^{*1}, b^{*1}) := (a_1(n), \hat{u}b(n) + (1-\hat{u})[(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x)])$, where, by construction, $a_1^{*1} < a_1^*$, to find a further equilibrium of $\Gamma_1(\delta)$ where type 1 receives the payoff $a_1(n+n') < a_1^{*1} < a_1^*$. 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 $\Gamma_1(\delta)$ where type 1 receives even lower payoffs, and so on.

We now define the strategies $(\hat{\sigma}(N), \hat{\tau}(N))$. These are 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. There are (potentially) many periods in which the normal type randomizes and in the period in which a randomization occurs the continuation equilibrium is of the form described in Definition 4. In the very first period of play (the very last iteration of Lemma 3), there is no randomization, so at this point $\hat{u} = 1$.

Definition 5. $((\hat{\sigma}(N), \hat{\tau}(N)))$: Let ϵ, $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$, (a_1^*, b^*), δ, $\hat{u}$, be as in the statement of Lemma 3 (dependence of $(\hat{\sigma}(N), \hat{\tau}(N))$ on these variables is supressed). Denote $(a_1^{*(0)}, b^{*(0)}) := (a_1^*, b^*)$. Definition 4 defines a strategy profile $(\hat{\sigma}(1), \hat{\tau}(1)) := (\hat{\sigma}(n^{(1)}; a_1^{*(0)}, b^{*(0)}, x^{(1)}, \hat{u}), \hat{\tau}(n^{(1)}; a_1^{*(0)}, b^{*(0)}, x^{(1)}, \hat{u}))$ for some $(n^{(1)}, x^{(1)}) := (n, x)$ as given by Lemma 3. Let $(a_1^{*(1)}, b^{*(1)})$ denote the players' payoffs from playing these strategies. Repeat this for each $l < N-1$, that is, given the payoff profile $(a_1^{*(l)}, b^{*(l)})$ generated by the strategies $(\hat{\sigma}(l), \hat{\tau}(l))$ apply Definition 4 to define

$$(\hat{\sigma}(l+1), \hat{\tau}(l+1)) := (\hat{\sigma}(n^{(l+1)}; a_1^{*(l)}, b^{*(l)}, x^{(l+1)}, \hat{u}), \hat{\tau}(n^{(l+1)}; a_1^{*(l)}, b^{*(l)}, x^{(l+1)}, \hat{u})),$$

for $(n^{(l+1)}, x^{(l+1)})$ as given by Lemma 3. Finally, define the very last iteration (the first to be played) as the strategy profile without randomizations

$$(\hat{\sigma}(N), \hat{\tau}(N)) := (\hat{\sigma}(n^{(N)}; a_1^{*(N-1)}, b^{*(N-1)}, x^{(N)}, 1), \hat{\tau}(n^{(N)}; a_1^{*(N-1)}, b^{*(N-1)}, x^{(N)}, 1))$$

for $(n^{(N)}, x^{(N)})$ as given by Lemma 3.

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 $\underline{a}_1((65/16)\epsilon)$.)

Lemma 4. *Let $0 < \epsilon < s(\bar{a}_1(0) - \hat{a}_1)/(10+3s)$ and $C > 0$ be given and let $\{(\hat{\imath}^t, \hat{\jmath}^t)\}_{t=0}^{T-1}$ and $\delta^*(\epsilon) < 1$ be so that $\widehat{A}_1 := ((1-\delta)/(1-\delta^T))\sum_{s=0}^{T-1}\delta^s A_1(\hat{\imath}^s, \hat{\jmath}^s) < \hat{a}_1 + \epsilon$ 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 $\bar{a}_1(2\epsilon) - \epsilon/2 > a_1^* > \bar{a}_1(3\epsilon) - C\epsilon$, $a_1 \in [\underline{a}_1((65/16)\epsilon) + \epsilon, \bar{a}_1(3\epsilon) - C\epsilon]$, and $\delta > \tilde{\delta}(\epsilon)$, then there exists an N and a strategy $(\hat{\sigma}(N), \hat{\tau}(N))$ such that $(\hat{\sigma}(N), \hat{\tau}(N))$ is an equilibrium of $\Gamma_1(\delta)$ 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 $1-\underline{r}$.*

Proof. Let $\tilde{\delta}(\epsilon) := \max\{\underline{\delta}(\epsilon), \delta^*(\epsilon), [32M/(\epsilon+32M)]^{1/T}, 1-\epsilon/[32M(S+1)], 1-\epsilon(\bar{a}_1(0)-\hat{a}_1)/9M^2, 1-s\epsilon(\bar{a}_1(0)-\hat{a}_1)/(16M^2(1+s))\}$. This lower bound on δ implies that if x and y are any two feasible payoffs for Player i, then

$$\left|x - [(1-\delta^T)y + \delta^T x]\right| = (1-\delta^T)|x-y| < (1-\delta^T)2M < \frac{\epsilon}{16}. \tag{20}$$

We will first show that it is possible to choose $\hat{u} \leq 1/2$ strictly positive, independent of δ, such that the payoff to type 1 at the equilibrium $(\hat{\sigma}(N_{\max}), \hat{\tau}(N_{\max}))$ is no greater than $\underline{a}_1((65/16)\epsilon) + \epsilon$. 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 (where now $b(n)$ is defined using the strategies that iterate Lemma 3) and Lemma 3 cannot be reapplied because $(a_1(n), \hat{u}b(n) + (1-\hat{u})[(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x)]) \notin G_1(3\epsilon)$. There are now two separate cases to consider: (1) If $\hat{u}b(n) + (1-\hat{u})[(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x)] \geq \hat{b} + 3\epsilon$, but $(a_1(n), \hat{u}b(n)+(1-\hat{u})[(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f(x)]) \notin G_1(3\epsilon)$, then it must be that $a_1(n) < \underline{a}_1(3\epsilon)$. (2) If $\hat{u}b(n)+(1-\hat{u})[(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f(x)] < \hat{b}+3\epsilon$, then $b(n) > \hat{b}+2\epsilon$ implies

$$(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f(x) < \hat{b}+2\epsilon+\frac{\epsilon}{1-\hat{u}} \leq \hat{b}+4\epsilon \tag{21}$$

(which follows as $\hat{u} \leq 1/2$). 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 $(1/16)\epsilon$ of the point $(x, f(x))$, by (20). We know that $f(x) < \hat{b}+(65/16)\epsilon$, from (20) and (21). If $f(x)$ is nondecreasing, therefore, it follows that $x < \underline{a}_1((65/16)\epsilon)$. This and (20) applied again implies $a_1(n) < \underline{a}_1((65/16)\epsilon)+(1/16)\epsilon$. If, however, $f(x)$ is decreasing over part of its range, $f(x) < \hat{b}+(65/16)\epsilon$ can also imply that $x > \bar{a}_1((65/16)\epsilon)$ and $a_1(n) > \bar{a}_1((65/16)\epsilon)-(1/16)\epsilon$. We will now show that $\hat{u}$ can be chosen (independently of δ and ϵ) sufficiently small so that this second alternative cannot apply. To be precise we will show that we can choose $\hat{u} > 0$ sufficiently small (but independent of δ and ϵ) so that $\hat{u}b(n)+(1-\hat{u})[(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f(x)] > \hat{b}+3\epsilon$ whenever $x > \bar{a}_1((65/16)\epsilon)$. As $b(n) \geq \hat{b}+2\epsilon$, it is sufficient to show that there exists some $e > 0$ such that for all $0 < \epsilon < s(\bar{a}_1(0)-\hat{a}_1)/(10+3s)$ and $\delta > \tilde{\delta}(\epsilon)$ it is the case that $(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f(x) > \hat{b}+(3+e)\epsilon$. By (19), it is sufficient to show that

$$(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta f\left(\frac{a_1(n)}{\delta}-\frac{1-\delta}{\delta}A_1(\tilde{\imath}, \hat{\jmath}^0)\right) \geq \hat{b}+(3+e)\epsilon. \tag{22}$$

There must be at least one iteration of the strategies for the constraint to bind, so we will write $(a_1(n), b(n)) = (1-\delta^{Tn})(\widehat{A}_1, \widehat{B})+\delta^{Tn}(a_1^\dagger, b^\dagger)$ where $(a_1^\dagger, b^\dagger) \in G_1(3\epsilon)$ is the continuation equilibrium payoff after n iterations of the finite sequence. By construction $a_1^\dagger > a_1(n) > \bar{a}_1((65/16)\epsilon)-\epsilon/16$. If $x > \bar{a}_1((65/16)\epsilon)$ when $f(x) < \hat{b}+(65/16)\epsilon$, then $f(\cdot)$ contains linear segments with strictly negative slope. Recall that $-s$ is the largest strictly negative slope of $f(\cdot)$ (the flattest downward sloping segment). A line through $(a_1^\dagger, b^\dagger)$ with slope $-s$ will lie below $f(x')$ for $x' \in [\bar{a}_1((65/16)\epsilon), a_1^\dagger]$, that is, $b^\dagger - s(x'-a_1^\dagger) \leq f(x')$ for all $x' \in [\bar{a}_1((65/16)\epsilon), a_1^\dagger]$.

Now, we establish that $x < a_1^\dagger$. The constraint $b(n) > \hat{b}+2\epsilon$ binds and any further iterations of the finite sequence will violate the constraint, so from (20) it must be that $\hat{b}+(33/16)\epsilon > (1-\delta^{Tn})\widehat{B}+\delta^{Tn}b^\dagger > \hat{b}+2\epsilon$. This implies a lower bound on $1-\delta^{nT}$ and thus a lower bound on $a_1^\dagger - a_1(n)$ of $(b^\dagger-\hat{b}-(33/16)\epsilon)(a_1^\dagger-\widehat{A}_1)/(b^\dagger-\widehat{B})$. However, $b^\dagger \geq \hat{b}+3\epsilon$ and $b^\dagger-\widehat{B} < 2M$, so $a_1^\dagger - a_1(n) > ((15/(32M))\epsilon)(a_1^\dagger-\widehat{A}_1)$. The definition of $-s$ implies that $\bar{a}_1((65/16)\epsilon) > \bar{a}_1(0)-(65\epsilon)/(16s)$, so $a_1^\dagger-\widehat{A}_1 > \bar{a}_1((65/16)\epsilon)-(1/16)\epsilon-\hat{a}_1-\epsilon \geq \bar{a}_1(0)-\hat{a}_1-(\epsilon/16)((65/s)+17)$, where the first inequality follows from $a_1^\dagger > \bar{a}_1((65/16)\epsilon)-(\epsilon/16)$ and $\widehat{A}_1 < \hat{a}_1+\epsilon$. If this inequality is substituted into the earlier one we get

$$a_1^\dagger - a_1(n) > \frac{15\epsilon}{32M}\left[\bar{a}_1(0)-\hat{a}_1-\frac{\epsilon}{16}\left(\frac{65}{s}+17\right)\right] > \frac{15\epsilon}{64M}(\bar{a}_1(0)-\hat{a}_1). \tag{23}$$

The last inequality follows from the upper bound on ϵ. By construction $|a_1(n)-x| < (1-\delta)2M$, so the lower bound on δ $(\delta \geq 1-\epsilon(\bar{a}_1(0)-\hat{a}_1)/9M^2)$ ensures $|a_1(n)-x| < a_1^\dagger - a_1(n)$. This establishes that $x < a_1^\dagger$, and the construction at the end of the previous paragraph can be used. Therefore, a sufficient condition for (22) is

$$(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0)+\delta\left[b^\dagger - s\left(\frac{a_1(n)}{\delta}-\frac{1-\delta}{\delta}A_1(\tilde{\imath}, \hat{\jmath}^0)-a_1^\dagger\right)\right] > \hat{b}+(3+e)\epsilon.$$

Some rearranging of this condition gives

$$(1-\delta)\big[B(\tilde{\imath}, \hat{\jmath}^0)-b^\dagger+s(A_1(\tilde{\imath}, \hat{\jmath}^0)-a_1^\dagger)\big]+(b^\dagger-\hat{b}-3\epsilon)+s(a_1^\dagger-a_1(n)) > e\epsilon. \tag{24}$$

Replacing the first term by a lower bound, noting that the second term in (24) is nonnegative by construction, and using (23), a sufficient condition for this is

$$-(1-\delta)(1+s)2M + \frac{15s\epsilon}{64M}(\bar{a}_1(0) - \hat{a}_1) > e\epsilon. \tag{25}$$

The lower bound on δ ($\delta \geq 1 - s\epsilon(\bar{a}_1(0) - \hat{a}_1)/(16M^2(1+s))$) implies that the coefficient on ϵ on the left of (25) is at least $7s(\bar{a}_1(0) - \hat{a}_1)/64M$. As (25) is sufficient for $a_1^\dagger \geq x$, an e with the requisite properties exists, and we have completed this part of the proof.

The payoff to type 1 at the equilibrium $(\hat{\sigma}(N_{\max}), \hat{\tau}(N_{\max}))$ is thus no greater than $\underline{a}_1((65/16)\epsilon) + \epsilon$. Therefore, type 1's payoff at the equilibrium $(\hat{\sigma}(N), \hat{\tau}(N))$ ranges from less than $\underline{a}_1((65/16)\epsilon) + \epsilon$ (for N large) to $a_1^* > \bar{a}_1(3\epsilon) - C\epsilon$ (for $N = 0$). By (20), type 1's payoff at the equilibrium $(\hat{\sigma}(N), \hat{\tau}(N))$ decreases 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 $[\underline{a}_1((65/16)\epsilon) + \epsilon, \overline{a}_1(3\epsilon) - C\epsilon]$.

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 very first period of play, so $n + n' < N_{\max}$. At $N_{\max}$ the constraint (18) binds and at all other iterations constraint (17) binds. We must, therefore, have

$$b(n+n') = (1-\delta^{n'T})\widehat{B} + \delta^{n'T}\{\hat{u}b(n) + (1-\hat{u})[(1-\delta)B(\tilde{\imath}, \hat{\jmath}^0) + \delta f(x)]\} > \hat{b} + 2\epsilon, \tag{26}$$

where x is chosen as in (19). (If there are *any* randomizations, then $\widehat{B} < \hat{b} + 2\epsilon$, because otherwise, the constraint (17) will not bind.) By definition of there being a randomization at $n + n'$ the inequality in (26) must be violated for one more iteration of the finite sequence, that is, $n + n' + 1$ (since the constraint $a_1(n+n') > \underline{a}_1(2\epsilon) + (1/2)\epsilon$ 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 $n + n' = N_{\max}$). The inequality (26) is therefore reversed when n' is replaced by $n' + 1$. This gives an upper bound on $\delta^{(n'+1)T}$. δ^T is bounded below by the assumption $\delta > \tilde{\delta}$, so we then get an upper bound on $\delta^{Tn'}$:

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

However, $a_1(n+n') = (1-\delta^{Tn'})\widehat{A}_1 + \delta^{Tn'}a_1(n)$ and $\widehat{A}_1 < \hat{a}_1 + \epsilon$, so an upper bound on $\delta^{n'T}$ implies an upper bound on $a_1(n+n')$:

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

The above expression implies that $a_1(n) - \widehat{A}_1$ declines exponentially, at a rate independent of δ, if the term in braces is bounded below one. If this is the case, we will be able to show that a finite number of randomizations are needed for $a_1(N_{\max}) \leq \underline{a}_1(2\epsilon) + (1/2)\epsilon$.

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

$$(27)\quad 1+\frac{\epsilon}{32M}+\eta < \frac{\hat{u}b(n)+(1-\hat{u})[(1-\delta)B(\tilde{\imath},\hat{\jmath}^0)+\delta f(x)]-\widehat{B}}{\hat{b}+2\epsilon-\widehat{B}}, \quad \forall 1>\delta>\tilde{\delta}(\epsilon).$$

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

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

There is a randomization at the payoff $a_1(n)$, so by Equation (22) $\hat{u}b(n)+(1-\hat{u})[(1-\delta)\cdot B(\tilde{\imath},\hat{\jmath}^0)+\delta f(x)]$ is greater than $\hat{b}+3\epsilon$. Thus, this sufficient condition must hold. We have shown that after the first randomization the value $a_1(n)-\widehat{A}_1$ declines (at least) exponentially with each randomization at some constant rate, say $\psi<1$, independently of δ. That is, $a_1(n+n')-\widehat{A}_1 < \psi[a_1(n)-\widehat{A}_1]$ (where n and $(n+n')$ refer to consecutive randomizations, as before). Since $\widehat{A}_1 < \hat{a}_1+\epsilon$ 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 κ randomizations $a_1(n)-(\hat{a}_1+\epsilon) < \psi^{\kappa-1}[\bar{a}_1-(\hat{a}_1+\epsilon)]$. If κ^* satisfies $\psi^{\kappa^*-1} < \epsilon[\bar{a}_1-(\hat{a}_1+\epsilon)]^{-1}$, we can be certain that at most κ^* randomizations are required before $a_1(n) \le \underline{a}_1(2\epsilon)+(1/2)\epsilon$, and that there is a strictly positive lower bound (independent of δ) $\underline{r} \ge \hat{u}^{\kappa^*}$ on the probability of sticking to repeated play of the sequence $\{(\hat{\imath}^t,\hat{\jmath}^t)\}_{t=0}^{T-1}$. □

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. There are several lemmas needed before the proof of Theorem 3 can be given. Using Assumption 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. The terminal equilibrium will serve to describe the players' long-run behavior 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 Assumption* 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,* $((\bar{\alpha}_1,\dots,\bar{\alpha}_K),\bar{\beta})$, *that satisfy*:

(a) $\bar{a}_k(3\epsilon)-(1/2)\epsilon \ge \bar{\alpha}_k > \bar{a}_k(3\epsilon)-C\epsilon$ *for some constant* C, *independent of* ϵ *and* δ, *and for* $k=1,2,\dots,K$;

(b) $\bar{\beta} \ge \hat{b}+3\epsilon$.

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

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

D_0 has a nonempty interior, by Assumption 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 $D_\epsilon \neq \varnothing$ for all $\epsilon \leq \hat{\epsilon}$. For $k = 1, 2, \ldots, K$ and $\epsilon \leq \hat{\epsilon}$, choose $\pi_k^*(\epsilon)$ to maximize $A_k(\cdot)$ 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

$$C^o\epsilon > \bar{a}_k(3\epsilon) - A_k(\pi_k^*(\epsilon)), \quad \text{for } \epsilon \leq \tilde{\epsilon}, \forall k. \tag{28}$$

Let k be given. For $\lambda \in [0,1]$ define $\pi^\lambda := \lambda\pi^\dagger + (1-\lambda)\pi_k^*(0)$, where $\pi^\dagger \in D_{\tilde{\epsilon}}$. By linearity $B(\pi^\lambda) \geq \lambda(\hat{b}+4\tilde{\epsilon}) + (1-\lambda)\hat{b}$, so π^λ is a feasible solution to $\max\{A_k(\pi) \mid B(\pi) \geq \hat{b} + \lambda 4\tilde{\epsilon}\}$. Thus, $\bar{a}_k(\lambda\tilde{\epsilon}) \geq A_k(\pi^\lambda) = \lambda A_k(\pi^\dagger) + (1-\lambda)\bar{a}_k(0)$. Let $\lambda = \epsilon/\tilde{\epsilon}$ for $0 \leq \epsilon \leq \tilde{\epsilon}$; then this implies

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

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

$$\bar{a}_k(\epsilon) \geq \bar{a}_k(0) - C_k\epsilon, \tag{29}$$

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

$$(1-\lambda)\bar{a}_{k'}(0) + \lambda A_{k'}(\pi^\dagger) \leq \bar{a}_{k'}(0) - C_{k'}3\epsilon - \frac{3}{4}\epsilon \tag{30}$$

(it is sufficient since the LHS of (30) is an upper bound for $A_{k'}(\pi^\lambda)$, while the RHS is no greater than $\bar{a}_{k'}(3\epsilon) - (3/4)\epsilon$ by (29)). 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 (28) follows.

By Result 3, 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

$$\begin{aligned} A_{k'}(\pi_k^*(\epsilon)) &= (1-\delta)\sum_{s=0}^\infty \delta^t A_{k'}(i_k^t, j_k^t), \quad \forall k, k' \in K, \\ B(\pi_k^*(\epsilon)) &= (1-\delta)\sum_{s=0}^\infty \delta^t B(i_k^t, j_k^t), \quad \forall k \in K. \end{aligned} \tag{31}$$

By Result 3 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 $\breve{\omega}_k + 3\epsilon \leq A_k(\pi_k^*(\epsilon))$, 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. 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 $t = 0, 1, \dots, K-2$, 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 below (31)). 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 $\bar{a}_k(3\epsilon) - C^o\epsilon - (1/8)\epsilon$. (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 Player 2's strategy must give her a payoff, $\bar{\alpha}_k$, satisfying $\bar{\alpha}_k > \bar{a}_k(3\epsilon) - (C^o + 1/8)\epsilon$, 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 $C = C^o + 1/8$. 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 following the action sequence of 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 following the action sequence of type k' and then deviating in period s is thus bounded above by

$$(1-\delta^s)A_k(\pi_{k'}^*(\epsilon)) + \delta^s\epsilon/2 + (1-\delta)\delta^s M + \delta^{s+1}(\breve{\omega}_k + \epsilon). \tag{32}$$

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 $\{i_{k'}^t\}_{t=0}^\infty$ 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 (32) 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 $\breve{\omega}_k + \epsilon + (1/8)\epsilon$, 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 $(1-\delta^K)M + \delta^K(\bar{a}_k(3\epsilon) - ((3/4)\epsilon))$, by definition of $\pi_{k'}^*(\epsilon)$. Assumption (iii) on δ ensures that this is less than $\bar{a}_k(3\epsilon) - (1/2)\epsilon$. 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 $\{(i_k^t, j_k^t)\}$. 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. □

The next result determines $K-1$ correlated strategies $(\underline{\pi}_2, \dots, \underline{\pi}_K) \in (\Delta^{IJ})^{K-1}$. It shows that: (a) Each correlated strategy holds type 1 to at most her minmax level; (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 produce a convex combination which holds type 1 to her minmax level when type 1 uses the same correlated strategy $\underline{\pi}_k$, $\bar{a}_1 + \lambda_k(A_1(\underline{\pi}_k) - \bar{a}_1) = \hat{a}_1$. From (b) the $\underline{\pi}_k$ are chosen to maximise the rate at which type k acquires payoff relative to the rate at which type 1's falls. This will be shown to imply that given the choice between following the prescribed path for type k ($\underline{\pi}_k$) and deviating when type 1 has a given continuation payoff, and following the prescribed path for type k' ($\underline{\pi}_{k'}$) and deviating when type 1 has the same continuation payoff, type k would always prefer the former.

Lemma 6. *Given Assumption* 1 *there exist correlated strategies* $(\underline{\pi}_2, \dots, \underline{\pi}_K) \in (\Delta^{IJ})^{K-1}$ *such that*:

(a) $A_1(\underline{\pi}_k) \le \hat{a}_1$ *for all* $k = 2, 3, \dots, K$,

(b) $(A_k(\underline{\pi}_k) - \bar{a}_k(0))/(\bar{a}_1(0) - A_1(\underline{\pi}_k)) \ge (A_k(\underline{\pi}_{k'}) - \bar{a}_k(0))/(\bar{a}_1(0) - A_1(\underline{\pi}_{k'}))$ *for all* $k, k' = 2, 3, \dots, K$,

(c) $\mathbf{z}$ *is individually rational, where*

$$\mathbf{z} := \left(\hat{a}_1, \bar{a}_2(0) + \frac{\bar{a}_1(0) - \hat{a}_1}{\bar{a}_1 - A_1(\underline{\pi}_2)}(A_2(\underline{\pi}_2) - \bar{a}_2(0)), \dots, \right. \\ \left. \bar{a}_K(0) + \frac{\bar{a}_1(0) - \hat{a}_1}{\bar{a}_1(0) - A_1(\underline{\pi}_K)}(A_K(\underline{\pi}_K) - \bar{a}_K(0))\right). \tag{33}$$

Proof. Consider the constrained optimization

$$\max_{\pi \in \Delta^{IJ}} \frac{A_k(\pi) - \bar{a}_k}{\bar{a}_1(0) - A_1(\pi)}, \quad \text{subject to } A_1(\pi) \le \hat{a}_1. \tag{34}$$

As $\bar{a}_1(0) > \hat{a}_1$, by Assumption 1, the maximand is well defined. As the constraint set is nonempty (by the Minimax Theorem) and compact there is a solution $\underline{\pi}_k$ to the optimization for all $k > 1$.

We now show that $\mathbf{z}$, defined by (33), is individually rational, that is $\{x \mid x \le \mathbf{z}\}$ is approachable. By Zamir (1992), for example, it is sufficient to show that for any $\mathbf{q} \in \Re^K_+$ there exists a mixed action, g, for Player 2 such that

$$\mathbf{q}((A_1(i, g), \dots, A_K(i, g)) - \mathbf{z}) \le 0, \quad \forall i \in I. \tag{35}$$

Let $\hat{g}$ be a mixed strategy that ensures Player 2 receives his minmax level ($B(i, \hat{g}) \ge \hat{b}$ for all $i \in I$) and let $\hat{g}_1$ be a mixed strategy that minmaxes type 1 ($A_1(i, \hat{g}_1) \le \hat{a}_1$ for all $i \in I$). We will show that for any $\mathbf{q} \ge 0$ either $g = \hat{g}$ or $g = \hat{g}_1$ will ensure (35) holds. Suppose that for some $\mathbf{q} \ge 0$ (35) does not hold with $g = \hat{g}$; then there exists $i \in I$ such that $\mathbf{q}((A_1(i, \hat{g}), \dots, A_K(i, \hat{g})) - \mathbf{z}) > 0$. By the definition of $\bar{\mathbf{a}}$, $\bar{a}_k(0) \ge A_k(i, \hat{g})$, and together with the fact that $\mathbf{q} \ge 0$, this implies $\mathbf{q}(\bar{\mathbf{a}} - \mathbf{z}) > 0$. A substitution from the definition from (33) shows this is equivalent to

$$(\bar{a}_1(0) - \hat{a}_1)\left(q_1 + \sum_{k=2}^{K} q_k \frac{A_k(\underline{\pi}_k) - \bar{a}_k(0)}{A_1(\underline{\pi}_k) - \bar{a}_1(0)}\right) > 0. \tag{36}$$

We must show that if (36) holds, $\mathbf{q}((A_1(i,\hat{g}_1),\dots,A_K(i,\hat{g}_1))-\mathbf{z})\le 0$ for all $i\in I$. It is sufficient to show $\mathbf{q}((A_1(\pi),\dots,A_K(\pi))-\mathbf{z})\le 0$ for all π such that $A_1(\pi)\le\hat{a}_1$. A substitution for $\mathbf{z}$ then gives

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

The first inequality arises because π is replaced by $\underline{\pi}_k$ in $(A_k(\pi)-\bar{a}_k(0))/(\bar{a}_1-A_1(\pi))$ and this is therefore maximized on the set of π's with $A_1(\pi)\le\hat{a}_1$. The final inequality then follows from (36). Thus, if $\mathbf{q}((A_1(i,\hat{g}),\dots,A_K(i,\hat{g}))-\mathbf{z})>0$ it must be true that $\mathbf{q}((A_1(i,\hat{g}_1),\dots,A_K(i,\hat{g}_1))-\mathbf{z})\le 0$. We can conclude that $\mathbf{z}$ is individually rational. □

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

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* $k'=2,3,\dots,K$, *such that for all* $1>\delta>\delta'(\epsilon)$: (a) $|\widehat{A}_{k,k'}-A_k(\underline{\pi}_{k'})|<\epsilon/2$ *for* $k\in K$, $k'=2,3,\dots,K$; (b) $|\widehat{B}_{k'}-B(\underline{\pi}_{k'})|<\epsilon/2$ *for* $k'=2,3,\dots,K$; *where*

$$\widehat{A}_{k,k'}:=\frac{1-\delta}{1-\delta^T}\sum_{s=0}^{T-1}\delta^s A_k(\hat{\imath}^s_{k'},\hat{\jmath}^s_{k'}),\qquad \widehat{B}_{k'}:=\frac{1-\delta}{1-\delta^T}\sum_{s=0}^{T-1}\delta^s B(\hat{\imath}^s_{k'},\hat{\jmath}^s_{k'}).\tag{37}$$

PROOF. For $k'=2,3,\dots,K$, let $\pi(k')$ be a rational approximation to the correlated strategy $\underline{\pi}_{k'}$, such that $\|\underline{\pi}_{k'}-\pi(k')\|<\epsilon/4$ for $k'=2,3,\dots,K$. There exists a positive integer T such that $T\pi(k')_{ij}$ is an integer for all $k'=2,3,\dots,K$, $i\in I$ and $j\in J$, (where $\pi(k)_{ij}$ denotes the ijth 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. □

We now prove our main result. It contains two 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. 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. The second step in the construction is an initial signalling phase where each type $k>1$ of Player 1 sends a distinct signal, while type 1 randomly selects one of the signals. Assuming that p_1 is sufficiently high, after this signalling phase Player 2 assigns positive probability only to type 1 and one other $k>1$, with arbitrarily high probability on type 1. Consequently, the argument of the first part of the proof can be applied. Two main difficulties arise in the construction: First, ensuring the indifference of type 1 between each of the signals, which requires that Player 2 randomizes in the period that each type k signals and that the outcome of Player 2's randomization determines the equilibrium of the two-type game that is subsequently played. The second difficulty is checking that none of the types $k>1$ can profitably deviate by sending a signal other than the assigned one.

PROOF OF THEOREM 3. *Some definitions and notation:* Choose $Q > 0$ so that

$$\bar{a}_k(0) - \bar{a}_k(3\epsilon) + 3\epsilon/4 < Q\epsilon \quad \forall k \in K,\ 0 < \epsilon < \bar{\epsilon}, \tag{38}$$

(where $\bar{\epsilon}$ is defined in Assumption 1). (See, e.g., the argument for (29) in Lemma 5.) We will also define $R \geq 0$ as follows:

$$R := \max_k \left| \frac{\bar{a}_k(0) - A_k(\underline{\pi}_k)}{\bar{a}_1(0) - A_1(\underline{\pi}_k)} \right|, \tag{39}$$

(where $\underline{\pi}_k$ is defined in Lemma 6). From Lemma 6(b) we have that

$$\frac{A_k(\underline{\pi}_k) - \bar{a}_k(0)}{\bar{a}_1(0) - A_1(\underline{\pi}_k)} \geq \frac{A_k(\underline{\pi}_{k'}) - \bar{a}_k(0)}{\bar{a}_1(0) - A_1(\underline{\pi}_{k'})}, \quad \forall k, k' = 2, 3, \ldots, K. \tag{40}$$

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

$$\frac{A_k(\underline{\pi}_k) - \bar{a}_k(0)}{\bar{a}_1(0) - A_1(\underline{\pi}_k)} > \frac{A_k(\underline{\pi}_{k'}) - \bar{a}_k(0)}{\bar{a}_1(0) - A_1(\underline{\pi}_{k'})}, \quad \forall k, k' = 2, 3, \ldots, K,\ k \neq k'. \tag{41}$$

(We will deal with the case of $k \neq k'$ satisfying (40) with equality at the end of the proof.) Finally, Y is defined to be the slope (with Player 2's payoffs in the numerator) of $G_1(0)$ when this set is a line segment ($\operatorname{Int} G_1(0) = \varnothing$) and when $\operatorname{Int} G_1(0) \neq \varnothing$ we define $Y = 1$. Y is bounded above and strictly positive by Assumption 1.

Let $\iota > 0$ be given, where $\iota < \min\{\bar{\epsilon}, \tilde{\epsilon}\}$ ($\bar{\epsilon}$ is defined in 1, $\tilde{\epsilon}$ in Lemma 5). Choose $0 < \epsilon < (\bar{a}_1(0) - \hat{a}_1)/3$ so that: (i) $3\epsilon < \iota$; (ii) for all $k, k' = 2, 3, \ldots, K$ with $k \neq k'$ it is true that for all $\delta > \delta'(\epsilon)$

$$\frac{\widehat{A}_{k,k} - x_k}{x_1 - \widehat{A}_{1,k}} > \frac{\widehat{A}_{k,k'} - x_k}{x_1 - \widehat{A}_{1,k'}} + (2+R)\epsilon, \qquad \frac{\widehat{A}_{k,k} - x_k}{x_1 - \widehat{A}_{1,k}} < R + 1 \tag{42}$$

for all $x_k \in (\bar{a}_k(3\epsilon) - C\epsilon, \bar{a}_k(3\epsilon) - (1/2)\epsilon]$ and all $x_1 \in (\bar{a}_1(3\epsilon) - C\epsilon, \bar{a}_1(3\epsilon) - (1/2)\epsilon]$, where $\widehat{A}_{k,k'}$ and $\delta'(\epsilon)$ are as defined in Lemma 7; (iii) $\lambda \in [0,1]$ such that $\lambda \hat{a}_1 + (1-\lambda)\bar{a}_1(0) > \hat{a}_1 + \iota - \epsilon/2$ implies $\lambda \mathbf{z} + (1-\lambda)\bar{\mathbf{a}}$ is $(2 + (Q+2)(R+1))\epsilon$-IR; (iv) $\underline{a}_1((65/16)\epsilon) + \epsilon < \underline{a}_1(\iota) < \bar{a}_1(3\epsilon) - C\epsilon$ where C is defined in Lemma 5 ($\underline{a}_1(\iota) < \bar{a}_1(0)$, because $G_1(\bar{\epsilon})$ is nonempty 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 because $\bar{a}_k(3\epsilon) - C\epsilon$ is continuous in ϵ at zero and $|\widehat{A}_{k,k'} - A_k(\underline{\pi}_{k'})| < \epsilon/2$ (by Lemma 7) and the strict inequality (41) holds; (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 $\delta_\iota = \max\{\tilde{\delta}(\epsilon), \delta_\epsilon, \bar{\delta}(\epsilon), (4M/(4M+\epsilon^2))^{1/K}\}$, where δ_ϵ is defined in Result 2 and $\bar{\delta}(\epsilon)$ is defined in Lemma 5.

4.2.1. The game with two types. 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 $\delta > \delta'(\epsilon)$. From Lemma 5, for all $\delta > \bar{\delta}(\epsilon)$ $\mathbf{p} \in \Delta^K$ the game $\Gamma(\mathbf{p}, \delta)$ has an equilibrium with payoffs, $(\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_K, \bar{\beta})$, 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 nonempty by (iv) in the preceding paragraph); 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 $\Gamma_1(\delta)$, in which type 1 gets a payoff within $(1/32)\epsilon$ of a'_1. (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 $0 < p_1 < 1/4$ 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 $\Gamma(\mathbf{p}', \delta)$. For convenience, define continuation payoffs for $k' = 1, k$ after history h^{t-1} given a strategy pair (σ, τ) as $c_{k'}(\sigma, \tau; h^{t-1}) := E_{\sigma,\tau}[(1-\delta)\sum_{t'=t}^{\infty} \delta^{t'-t} A_{k'}(i^{t'}, j^{t'}) \mid h^{t-1}]$.

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

Type 1 plays a strategy, which we denote as $\hat{\sigma}_1(k; N)$, so that from Player 2's perspective the combined actions of types 1 and k over the first TN periods replicate the strategy $\hat{\sigma}(k; N)$, 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, type 1 actually deviates from the sequence with probability more than $1 - \hat{u}$ 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 choose $\mathbf{p}'$ 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 equilibrium of $\Gamma(\mathbf{p}, \delta)$ with the payoffs $(\bar{\alpha}_1, \ldots, \bar{\alpha}_K, \bar{\beta})$ given above being played thereafter. However, if Player 1 uses a pure action at t that deviates from her equilibrium strategy (i.e., a probability zero action), then Player 2 responds in the following way. For any $\tilde{h}^{\tilde{t}} = (\tilde{h}^{\tilde{t}-1}, (i^{\tilde{t}}, j^{\tilde{t}}))$ satisfying $\Pr_{\hat{\sigma}(k;N), \hat{\tau}(k;N)}(\tilde{h}^{\tilde{t}-1}) > 0$, $\Pr_{\hat{\sigma}(k;N), \hat{\tau}(k;N)}(\tilde{h}^{\tilde{t}}) = 0$, let $c^*_1 := c_1(\hat{\sigma}_1(k; N), \hat{\tau}(k; N); \tilde{h}^{\tilde{t}-1})$ be type 1's equilibrium payoff from $\tilde{t}$. Then she takes the convex combination $\lambda\mathbf{z} + (1-\lambda)\bar{\mathbf{a}}$, of the point $\mathbf{z}$ (defined in (33)) and the point $\bar{\mathbf{a}}$, that gives type 1 exactly the payoff c^*_1, that is, $\lambda = (\bar{a}_1(0) - c^*_1)/(\bar{a}_1(0) - \hat{a}_1)$. By the construction above (point (iii) below (42)), because $c^*_1 > \hat{a}_1 + \iota - \epsilon/2$, then this convex combination is $(2 + (1+R)(2+Q))\epsilon$-IR: 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 $\widehat{A}_{1,k}$ and the terminal equilibrium payoff $\bar{\alpha}_1$: $(1 - \delta^{nT})\widehat{A}_{1,k} + \delta^{nT}\bar{\alpha}_1$, 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 $\hat{a}_1(0) + \iota - \epsilon/32$. 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 c at any point always exceeds $\hat{a}_1 + \iota - \epsilon/16$. Thus, there exists a vector of IR payoffs $(\omega_1, \ldots, \omega_K) \in \Re^K$ such that

$$
\begin{aligned}
(43) \qquad & (\omega_1, \ldots, \omega_K) + (2 + (1+R)(2+Q))\epsilon\mathbf{1} \\
& \leq \lambda\mathbf{z} + (1-\lambda)\bar{\mathbf{a}} \\
& = \Bigg(c^*_1, \bar{a}_2(0) - (\bar{a}_1(0) - c^*_1)\frac{\bar{a}_2(0) - A_2(\underline{\pi}_2)}{\bar{a}_1(0) - A_1(\underline{\pi}_2)}, \ldots, \\
& \qquad \bar{a}_K(0) - (\bar{a}_1(0) - c^*_1)\frac{\bar{a}_K(0) - A_K(\underline{\pi}_K)}{\bar{a}_1(0) - A_1(\underline{\pi}_K)}\Bigg).
\end{aligned}
$$

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 $\delta > \delta_\epsilon$.

It is sufficient to show that types 1 and k do not benefit by deviating from their equilibrium strategy by choosing a pure strategy that specifies an action that is assigned probability zero by their equilibrium strategy. (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.) For a pure strategy of Player 1, σ', let $\tilde{h}^t$ be the history at t, $t = 0, 1, 2, \dots$, induced by the play of σ' against $\hat{\tau}(k; N)$, and define $\tilde{t} := \min\{t \geq 0\colon \Pr_{\hat{\sigma}(k;N),\hat{\tau}(k;N)}(\tilde{h}^{t-1}) > 0$ and $\exists i \in I, \sigma'(\tilde{h}^{t-1})(i) = 1$ and $\hat{\sigma}(k; N)(\tilde{h}^{t-1})(i) = 0\}$ to be the first period in which an observable deviation occurs; this is well-defined, provided σ' implies an observable deviation at some date as $\hat{\tau}(k; N)$ is pure up to that point. Equilibrium continuation payoffs are $c^*_{k'} = c_{k'}(\hat{\sigma}_{k'}(k; N), \hat{\tau}(k; N); \tilde{h}^{\tilde{t}-1}), k' = 1, k$. Type 1's continuation expected payoff from σ', $c_1(\sigma', \hat{\tau}(k; N); \tilde{h}^{\tilde{t}-1}) \leq (1-\delta)M + \delta(\omega_1 + \epsilon) < \omega_1 + 3\epsilon < c^*_1$ (because $\delta > \delta_\epsilon$, and from (43)); a deviation is suboptimal.

Next, we show that type k cannot profitably deviate. Type k can make unobservable deviations from the equilibrium by playing the action type 1 uses to reveal her type (by playing $\tilde{\imath}$ at a point of randomization), and then by continuing to follow type 1's actions, playing out an equilibrium of the game $\Gamma_1(\delta)$. It is possible that such a deviation is profitable. A small reworking of the players' strategies gives a "semipooling" equilibrium (either type 1 reveals her type or both types end up following the same path) 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 $\Gamma_1(\delta)$. (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 original putative equilibrium. Player 2's payoffs remain individually rational at each date because the continuation equilibrium after $\tilde{\imath}$ yields a higher payoff than the payoff when $\tilde{\imath}$ is not played, and so Player 2's payoff increases. 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 $\Gamma_1(\delta)$. 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, assuming for the moment that the equilibrium is not the semipooling type just described. By (43), there exists a vector of punishment payoffs ω such that

$$
\begin{aligned}
(44)\quad & \omega_k + (2 + (1+R)(2+Q))\epsilon \\
& \leq \bar{a}_k(0) - (\bar{a}_1(0) - c^*_1)\frac{\bar{a}_k(0) - A_k(\underline{\pi}_k)}{\bar{a}_1(0) - A_1(\underline{\pi}_k)} \\
& = \{(1-\delta^{TN'})\widehat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - c^*_k\} + \delta^{TN'}\{\bar{a}_k(0) - \bar{\alpha}_k\} + (1-\delta^{TN'})\{A_k(\underline{\pi}_k) - \widehat{A}_{kk}\} \\
& \quad + \frac{\bar{a}_k(0) - A_k(\underline{\pi}_k)}{\bar{a}_1(0) - A_1(\underline{\pi}_k)}\big\{(1-\delta^{TN'})[\widehat{A}_{1k} - A_1(\underline{\pi}_k)] + [c^*_1 - (1-\delta^{TN'})\widehat{A}_{1k} - \delta^{TN'}\bar{\alpha}_1] \\
& \qquad\qquad - \delta^{TN'}[\bar{a}_1(0) - \bar{\alpha}_1]\big\} + c^*_k \\
& < c^*_k + \{(1-\delta^{TN'})\widehat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - c^*_k\} + Q\epsilon + \epsilon/2 \\
& \quad + \frac{\bar{a}_k(0) - A_k(\underline{\pi}_k)}{\bar{a}_1(0) - A_1(\underline{\pi}_k)}\{\epsilon/2 - (1-\delta^{TN'})\widehat{A}_{1k} - \delta^{TN'}\bar{\alpha}_1 + c^*_1 + Q\epsilon\}.
\end{aligned}
$$

The final inequality follows from (38), $A_k(\underline{\pi}_k) - \widehat{A}_{kk} < \epsilon/2$ and $\widehat{A}_{1k} - A_1(\underline{\pi}_k) < \epsilon/2$ (which follows from Lemma 7). Type 1's equilibrium continuation payoff, c^*_1, 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 of 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 (20), type 1's payoff c_1^* satisfies $|(1-\delta^{TN'})\widehat{A}_{1k} + \delta^{TN'}\bar{\alpha}_1 - c_1^*| \leq \epsilon/16$ and type k's continuation payoff, c_k^*, satisfies $|(1-\delta^{TN'})\widehat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - c_k^*| \leq \epsilon/16$. These inequalities, and (39), substituted in (44), imply that $\omega_k + (3+R)\epsilon < c_k^*$; 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 at τ, so an equilibrium of $\Gamma_1(\delta)$ 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 $\omega_k' = (1-\delta)a_k + \delta\omega_k$ 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

$$(1-\delta)\sum_{s=0}^{t-1}\delta^s A_k(i^s, j^s) + \delta^t\omega_k' = (1-\delta^t)(1-\delta)\sum_{s=0}^{\infty}\delta^s A_k(i^s, j^s) + \delta^t\omega_k'$$
$$+\delta^t(1-\delta)\left[\sum_{s=0}^{\infty}\delta^s A_k(i^s, j^s) - \sum_{s=t}^{\infty}\delta^{s-t}A_k(i^s, j^s)\right].$$

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

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

By assumption, k does not want to pool on the revealing equilibrium, so the first term on the RHS is nonpositive. The final term on the RHS is less that $(9/16)\epsilon$, because the strategies $\hat{\sigma}(k; N)$ used Result 3 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 $(1/16)\epsilon$. Thus, this condition can only be true if $d' < \omega_k' + (9/16)\epsilon$, or $d' < \omega_k + (10/16)\epsilon$ because of the assumption on δ. The punishment payoff, ω_k, is determined by (43) and c_1^* (the continuation payoff to type 1 at the point of the observed deviation by type k). Replacing c_k^* by d' in (44), letting N' be the number of plays of the sequence left at τ(c_k^* and N' are arbitrary in (44)), noting that c_1^* is within $\epsilon/2$ of the continuation payoff at τ to type 1, say c', and as above $|(1-\delta^{TN'})\widehat{A}_{1k} + \delta^{TN'}\bar{\alpha}_1 - c'| \leq (\epsilon/16)$ and also $|(1-\delta^{TN'})\widehat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - d'| \leq (\epsilon/16)$, we can deduce from (44) that $\omega_k + (3+(15/16)R)\epsilon < d'$. This is a contradiction as $d' < \omega_k + (10/16)\epsilon$. [In the semipooling equilibrium, described in the previous paragraph, type k and type 1 both play out an equilibrium of $\Gamma_1(\delta)$. Type k benefits by a subsequent observable deviation if her payoff from continued play of the equilibrium, $d' \equiv (1-\delta)\sum_{s=0}^{\infty}\delta^s A_k(i^s, j^s)$ (where d' is again k's payoff from sticking to her equilibrium strategy, computed at the start of the revealing equilibrium, but now k's strategy specifies that $\tilde{\imath}$ is played), is less than what she receives by deviation t periods after $\tilde{\imath}$ was played: $(1-\delta)\sum_{s=0}^{t-1}\delta^s A_k(i^s, j^s) + \delta^t\omega_k'$. This implies $\omega_k' > (1-\delta)\sum_{s=t}^{\infty}\delta^s A_k(i^s, j^s)$. But by $\omega_k' = (1-\delta)a_k + \delta\omega_k$ and Result 3, we have again $\omega_k + (1/16)\epsilon > d' - (9/16)\epsilon$. However, noting that in a semipooling equilibrium d' satisfies $d' \geq (1-\delta^{TN'})\widehat{A}_{kk} + \delta^{TN'}\bar{\alpha}_k - (\epsilon/16)$ (where N' again denotes the number of plays of the sequence left at the start of the revealing equilibrium), and c as in the above argument satisfies $|(1-\delta^{TN'})\widehat{A}_{1k}+\delta^{TN'}\tilde{\alpha}_1-c|\leq(9\epsilon/16)$, so (44) again implies $\omega_k+(3+(15/16)R)\epsilon<d'$, a contradiction.]

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 $0<p_1<1/4$ 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 $|\tilde{\alpha}_1-a_1'|<(1/32)\epsilon$. 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 for any pair $(a_1,b)\in G_1(\iota)$ with $a_1<\bar{a}_1(3\epsilon)-C\epsilon,\Gamma(\mathbf{p}'',\delta)$ has an equilibrium with the payoffs (α_1^*,β^*) that satisfy $\|(\alpha_1^*,\beta^*)-(a_1,b)\|<\epsilon$. To do this, it is necessary to alter the period zero strategies of the equilibrium described above. 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 equilibrium just constructed, 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 $(1/32)\epsilon$ of a_1 and we can therefore also choose $\tilde{b}$ to be within $(1/32)\epsilon$ of b (since $(a_1,b)\in G_1(\iota)$ and $\epsilon<\iota$). The arguments immediately 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 constructed above, 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 above) we can rewrite 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 $p_1<1/4$ 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 $1-p_1''<(\underline{r}(3/4\epsilon))/6M$. Provided $p_1''>1-\underline{r}'$ where $\underline{r}':=(\underline{r}(3/4)\epsilon)/6M$ 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 the semipooling equilibrium to reestablish equilibrium.) We have now shown that when $K=2$ and $(a_1,b)\in G_1(\iota)\cap\{(x,y)\mid x<\bar{a}_1(3\epsilon)-C\epsilon-\epsilon\}$ the game has an equilibrium and payoffs that satisfy $\|(\alpha_1,\beta)-(a_1,b)\|<\iota$. (The condition $a_1<\bar{a}_1(3\epsilon)-C\epsilon-\epsilon$ ensures there is at least one randomization by Player 1.) By choosing $\iota<\min\{\nu/2,\bar{\epsilon}/2\}$ sufficiently small then proves the theorem when $K=2$.

4.2.2. 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 (15). 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 $t = 0, 1, \ldots, K-3$: *The Signalling Phase*. The players use the following strategies: *Type* k, where $k = 2, 3, \ldots, K-1$, plays action $i^t = 1$ in periods $t = 0, 1, \ldots, k-3$ and in period $t = k-2$ she plays action $i = 2$ to signal her type. *Type* K plays action $i^t = 1$ in periods $t = 0, 1, \ldots, K-3$. The signalling phase ends the first time $i^t = 2$ or in period $K-1$, whichever happens the sooner. *Type* 1 chooses a type $k = 2, 3, \ldots, K$ with probability ϕ_k and sends the signal appropriate for that type. (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 $q^t(\hat{h}^{t-1})$ and action $j = 2$ with probability $1 - q^t(\hat{h}^{t-1})$, where $\hat{h}^{t-1}$ 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 two-type game with the types $\{1, t+2\}$.)

After the signalling. As soon as type k is identified, 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 1 of the proof; however, the equilibrium they play will depend on the sequence of actions Player 2 played in the signalling phase.

We will begin by considering the case where $\operatorname{Int} G_1(0) \neq \varnothing$. Let us denote

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

as the open rectangle centered at the point (x, y) with width W and height H. Let (a_1, b) be a point in $G_1(\iota)$ that satisfies the condition $U[(a_1, b); \iota, \iota] \subset G_1(\iota) \cap \{(x, y) \mid x < \bar{a}_1(3\epsilon) - C\epsilon\}$ (ι will be chosen sufficiently small to ensure this is possible). We will now show that: The continuation equilibria after the signalling can be chosen to give the players incentives to randomize. After the signalling phase Player 2's posterior beliefs will still attach arbitrarily high probability to type 1 as $p_1 \to 1$, so an equilibrium (of Part 1) can then be played. We also show that the signalling strategies are an equilibrium that 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 choose the continuation equilibria in period $K-3$ with payoffs that satisfy

$$(\alpha_1^{K,1}, \beta^{K,1}), (\alpha_1^{K-1,2}, \beta^{K-1,2}) \in U[(a_1^\dagger - \epsilon, b^\dagger - \epsilon); \epsilon, Y\epsilon], \tag{45}$$

$$(\alpha_1^{K,2}, \beta^{K,2}), (\alpha_1^{K-1,1}, \beta^{K-1,1}) \in U[(a_1^\dagger + \epsilon, b^\dagger + \epsilon); \epsilon, Y\epsilon], \tag{46}$$

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 $\operatorname{Int} G_1(0) \neq \varnothing$, as assumed for the moment; however it will be convenient to retain the general notation for the case when $\operatorname{Int} G_1(0) = \varnothing$.) It is possible to choose such continuation equilibria, because the sets on the right of (45) and (46) are in $\operatorname{Int} G_1(\iota) \cap \{(x, y) \mid x < \bar{a}_1(3\epsilon) - C\epsilon\}$ and part 1 of the proof, therefore, applies. Continuation equilibria satisfying (45) and (46) can be found, because (by (20) and part 1) type 1's payoff can be approximated to within $\epsilon/16$ and by 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 have an incentive to randomize and 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 $\rho = 1 + 1/8$. 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:

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

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

($A_1(i, q^{K-3})$ 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

$$\frac{1-\delta}{\delta}[A_1(1, q^{K-3})-A_1(2, q^{K-3})]=q^{K-3}[\alpha_1^{K-1,1}-\alpha_1^{K,1}]+(1-q^{K-3})[\alpha_1^{K-1,2}-\alpha_1^{K,2}]. \tag{47}$$

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

$$\frac{1-\delta}{\delta}[B(\mu, 1)-B(\mu, 2)]=\mu[\beta^{K,2}-\beta^{K,1}]+(1-\mu)[\beta^{K-1,2}-\beta^{K-1,1}]. \tag{48}$$

We can find $q^{K-3}\in[0, 1]$ and $\mu\in[0, 1]$ to make both players indifferent. First, the LHS of (47) is less than $\epsilon/16$ (by our assumption on δ) and the LHS of (48) is less than $Y\epsilon(1/16)$ 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 (47) (respectively, (48)) 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 (47) and (48). There are upper and lower bounds on the value of μ for which (48) holds. As the LHS is less than $Y\epsilon(1/16)$, 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 $(3/4+1/64)>\mu>(1/4-1/64)$. 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 $\rho=1+1/8$.

We will use the continuation equilibria after period $K-4$ of the signalling phase (assuming types $1<k\leq K-2$ are not signalled), described above, to construct an equilibrium for period $K-4$ onward with payoffs in $U[(a_1^\dagger, b^\dagger); \epsilon\rho^2, Y\epsilon\rho^2]$, provided

$$U\big[(a_1^\dagger, b^\dagger);\ (2+\rho+\rho^2)\epsilon, S(2+\rho+\rho^2)\epsilon\big]\subset U[(a_1, b);\ \iota, \iota] \tag{49}$$

where S is defined below (16). Repeat the argument of the previous paragraph with the sets in (45) and (46) 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 $U[(a_1^\dagger, b^\dagger)-(\epsilon\rho, Y\epsilon\rho);\ \epsilon\rho, Y\epsilon\rho]$ ((49) 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 (49) 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 1 and (49) 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 $(1/4-1/64, 3/4+1/64)$ 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 $U[(a_1^\dagger, b^\dagger)\pm(1+\rho+\cdots+\rho^{K-3})(\epsilon, Y\epsilon);\ \epsilon, Y\epsilon]$. This is possible if $(a_1, b)=(a_1^\dagger, b^\dagger)$, (v) holds (see beginning of proof) and $U[(a_1, b); \iota, \iota]\subset G_1(\iota)\cap\{(x, y)\mid x<\bar{a}_1(3\epsilon)-C\epsilon\}$. 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 $\operatorname{Int} G_1(0)=\varnothing$, 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 neighborhood 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 $(1/4 - 1/64)^{K-1}$ (from the bound on μ above). Provided $p_k < \underline{r}'(1/4 - 1/64)^{K-1}$ we have $p_1'' \geq 1 - \underline{r}'$ and it is possible to apply part 1 of the proof and play continuation equilibria satisfying (45) and (46). The required lower bound on p_1 is, thus, $1 - \underline{r}'(1/4 - 1/64)^{K-1}$ (this implies $p_k < \underline{r}'(1/4 - 1/64)^{K-1}$ for all $k > 1$).

We now show that no player wishes to deviate from their 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 1 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 original strategy described and then treat the case described in part 1 when the semipooling strategies are followed. 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 (37), her payoff from this, discounted to the period after the signalling is finished, is $(1 - \delta^{TN'})\widehat{A}_{k,k'} + \delta^{TN'}\bar{\alpha}_k$, whereas her payoff from playing her equilibrium strategy can be written as $(1 - \delta^{TN})\widehat{A}_{k,k} + \delta^{TN}\bar{\alpha}_k$. At an equilibrium, type 1 will follow the action sequences of type k and type k' with positive probability. Let c be type 1's expected equilibrium payoff from type k's sequence and c' be her expected payoff from type k''s sequence, that is,

$$c = (1 - \delta^{TN})\widehat{A}_{1,k} + \delta^{TN}\bar{\alpha}_1 = (1 - \delta^{TN})(\widehat{A}_{1,k} - \bar{\alpha}_1) + \bar{\alpha}_1; \tag{50}$$

$$c' = (1 - \delta^{TN'})\widehat{A}_{1,k'} + \delta^{TN'}\bar{\alpha}_1 = (1 - \delta^{TN'})(\widehat{A}_{1,k'} - \bar{\alpha}_1) + \bar{\alpha}_1. \tag{51}$$

The following will be a sufficient condition to rule out the first form of deviation described above (since the signalling phase contributes at most $\epsilon^2/2$ to payoffs by our choice of δ):

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

or equivalently

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

or

$$\frac{\widehat{A}_{k,k} - \bar{\alpha}_k}{\bar{\alpha}_1 - \widehat{A}_{1,k}} - \frac{\widehat{A}_{k,k'} - \bar{\alpha}_k}{\bar{\alpha}_1 - \widehat{A}_{1,k'}} > \frac{\widehat{A}_{k,k} - \bar{\alpha}_k}{\bar{\alpha}_1 - \widehat{A}_{1,k}} \frac{c - c'}{\bar{\alpha}_1 - c'} + \frac{\epsilon^2}{\bar{\alpha}_1 - c'}, \tag{52}$$

where the last inequality follows from substitution for $(1-\delta^{TN})$ from (50) and for $(1-\delta^{TN'})$ from (51). By (42) the LHS above is greater than $(2+R)\epsilon$, so it is sufficient to show that the RHS is less than this. 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 $(1/2)\epsilon^2$, so $|c-c'|<\epsilon^2$. Also $c'<\bar{a}_1(3\epsilon)-C\epsilon-\epsilon$ so that there is at least one iteration of the finite sequence and $\bar{\alpha}_1>\bar{a}_1(3\epsilon)-C\epsilon$ so $\bar{\alpha}_1-c'$ (the denominator of the last term) is strictly bigger than ϵ. The last term is, therefore, strictly less than ϵ. Similarly, (42) implies the first term on the RHS is less than $(R+1)\epsilon$. So (52) 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 increases Player 2's payoff when k' is signalled so her payoffs remain individually rational throughout. (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 dealing with the semipooling equilibrium applies mutatis mutandis.

Now, we must deal with the amended strategies and consider what occurs if type k' at some point plays a semipooling equilibrium with type 1, rather than continuing to reveal her type. If type k' and type 1 play the semipooling 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 also applies to this case.

Now, we return to the condition (41), 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

$$\frac{A_k(\underline{\pi}_k)-\bar{a}_k(0)}{\bar{a}_1(0)-A_1(\underline{\pi}_k)}=\frac{A_k(\underline{\pi}_{k'})-\bar{a}_k(0)}{\bar{a}_1(0)-A_1(\underline{\pi}_{k'})}. \tag{53}$$

In this case, we can choose $\underline{\pi}_k=\underline{\pi}_{k'}$ and the sequence $\{(\hat{\imath}_k^s,\hat{\jmath}_k^s)\}$ to be the same as $\{(\hat{\imath}_{k'}^s,\hat{\jmath}_{k'}^s)\}$. 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 $\underline{\pi}_{k'}$ (since $\underline{\pi}_{k'}$ must also solve (34)), 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 $G_1(\iota)\cap\{(x,y)\mid x<\bar{a}_1(3\epsilon)-C\epsilon-\epsilon\}$ that are at a distance at least ι from the boundary of the relative interior of $G_1(\iota)\cap\{(x,y)\mid x<\bar{a}_1(3\epsilon)-C\epsilon-\epsilon\}$. We have shown that there exists a $\delta_\iota<1$ and $p_1^\iota<1$ such that for all p with $p_1>p_1^\iota$ and $\delta>\delta_\iota$, given any $(a_1,b)\in R(\iota)$ the game $\Gamma(p,\delta)$ has an equilibrium with payoffs that satisfy $\|(\alpha_1, \beta) - (a_1, b)\| < \iota$. By choosing $\iota < \nu/3$ and sufficiently small the Theorem follows. □

Acknowledgments. We are particularly grateful to three 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, and 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.

References

Blackwell, D. 1956. An analog of the minmax theorem for vector payoffs. *Pacific J. Math.* **65** 1–8.

Cripps, M. W., K. M. Schmidt, J. P. Thomas. 1996. Reputation in perturbed repeated games. *J. Econom. Theory* **69** 387–410.

Forges, F. 1992. Non-zero-sum repeated games of incomplete information. R. J. Aumann, S. Hart, eds. *Handbook of Game Theory*. North Holland, Amsterdam, The Netherlands.

Fudenberg, D., D. K. Levine. 1992. Maintaining a reputation when strategies are imperfectly observed. *Rev. Econom. Stud.* **59** 561–579.

———, E. Maskin. 1991. On the dispensability of public randomization in discounted repeated games. *J. Econom. Theory* **53** 428–438.

Hart, S. 1985. Nonzero-sum two-person repeated games with incomplete information. *Math. Oper. Res.* **10** 117–153.

Jordan, J. S. 1995. Bayesian learning in repeated games. *Games Econom. Behavior* **9** 8–20.

Kalai, E., 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., L. Yariv. 1999. Repeated games with incomplete information on one side: The case of different discount factors. *Math. Oper. Res.* **24** 204–218.

Shalev, J. 1994. Nonzero-sum two-person repeated games with incomplete information and known-own payoffs. *Games Econom. Behavior* **7** 246–259.

Sorin, S. 1999. Merging, reputation and repeated games with incomplete information. *Games Econom. Behavior* **29** 274–308.

Zamir, S. 1992. Repeated games of incomplete information: Zero sum. R. J. Aumann, S. Hart, eds. *Handbook of Game Theory*. North Holland, Amsterdam, The Netherlands, 119–126.

M. W. Cripps: John M. Olin School of Business, Washington University in St. Louis, 1 Brookings Drive, St. Louis, MO 63130-4899; e-mail: cripps@olin.wustl.edu

J. P. Thomas: University of Edinburgh, Edinburgh, Scotland, UK; e-mail: jps@st-and.ac.uk

Summary of Comments on MOOR0050

Page: 6

Sequence number: 1
Author:
Date: 7/9/2003 10:01:49 AM
Type: Note
Open Box for "end of proof" per journal style. OK? Please check placement of all boxes.

Page: 8

Sequence number: 1
Author:
Date: 7/9/2003 10:03:01 AM
Type: Note
This heading should be 4.0, OK?

Page: 10

Sequence number: 1
Author:
Date: 7/9/2003 10:04:07 AM
Type: Note
Call out for Figure 1, OK?

Page: 12

Sequence number: 1
Author:
Date: 7/9/2003 10:04:29 AM
Type: Note
Edit OK?

Page: 20

Sequence number: 1
Author:
Date: 7/9/2003 10:05:42 AM
Type: Note
Edit OK?

Page: 21

Sequence number: 1
Author:

Date: 7/9/2003 10:06:27 AM
Type: Note
Changed numbers of heads to avoid confusion. OK?

Page: 22

Sequence number: 1
Author:
Date: 7/9/2003 10:06:54 AM
Type: Note
Edit OK?

Page: 30

Sequence number: 1
Author:
Date: 7/9/2003 10:07:47 AM
Type: Note
Location?