Queen's Economics Department Working Paper No. 1032

Switching Costs in Infinitely Repeated Games

Barton Lipman
Boston University

Ruqu Wang
Queen's University

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

1-2006

Switching Costs in Infinitely Repeated Games[1]

Barton L. Lipman[2]
Boston University

Ruqu Wang[3]
Queen's University

Current Draft

January 2006

[1]The authors thank Ray Deneckere, and seminar audiences at BU, Colorado–Boulder, Rochester, Stanford, UBC, and Western Ontario for helpful comments. Lipman also thanks the National Science Foundation for support under Grant 0136887. Wang also thanks SSHRC of Canada for financial support.

[2]Email: blipman@bu.edu.

[3]Email: wangr@qed.econ.queensu.ca

Abstract

We show that small switching costs can have surprisingly dramatic effects in infinitely repeated games if these costs are large relative to payoffs in a single period. This shows that the results in Lipman and Wang [2000] do have analogs in the case of infinitely repeated games. We also discuss whether the results here or those in Lipman–Wang [2000] imply a discontinuity in the equilibrium outcome correspondence with respect to small switching costs. We conclude that there is not a discontinuity with respect to switching costs but that the switching costs do create a discontinuity with respect to the length of a period.

1 Introduction

Lipman and Wang [2000] showed that switching costs can have surprisingly strong effects in frequently but finitely repeated games. More specifically, suppose we have a finite stage game where the length of each period is Δ and the total length of time of play is equal to $\mathcal{L} = (T+1)\Delta$ for some integer T. Suppose the payoff to player i from the sequence of action profiles $(a^0, \ldots, a^T)$ is given by

$$\sum_{t=0}^{T} [\Delta u_i(a^t) - \varepsilon I_i(a^t, a^{t-1})] \tag{1}$$

where $I_i(a, a') = 0$ if $a_i = a'_i$ and 1 otherwise.[1] In other words, he receives the payoff associated with each action vector played times the length of time these actions are played, minus a cost for each time he himself changes actions. We showed some very unexpected behavior in such games for small ε and Δ as long as ε is sufficiently large relative to Δ. For example, in games like the Prisoners' Dilemma which have a unique subgame perfect equilibrium outcome without switching costs, we obtain multiple equilibrium outcomes. In other games, such as coordination games, which have multiple equilibria without switching costs, we showed that one can have a unique subgame perfect equilibrium with small switching costs.

The analysis used finite repetition in a critical way. We noted that if the switching cost is large relative to one period worth of payoff, then no player would find it worthwhile to change actions in the last period regardless of what actions were played in the preceding period. This causes the usual backward induction arguments to break down. The fact that actions must be fixed at the end can have large effects early in the game.

Here we consider the effect of switching costs in an infinitely repeated game for four reasons. First, given the way our earlier analysis exploited the finite horizon, it is not obvious whether similar effects could be obtained in an infinitely repeated game. Second, our earlier analysis had the drawback that it was impossible to give many general characterization results. A natural conjecture is that the simplicity of the infinite horizon may allow us to characterize the set of equilibrium payoffs, at least for "sufficiently patient" players.

Third, just as with our earlier paper, we seek to explore to what extent the standard analysis is robust with respect to modifications of the model which seem "small." A cost to changing actions from one period to the next seems natural for at least two reasons. First, it is a simple way of capturing a type of bounded rationality. Intuitively, it is easier to continue doing the same thing as in the past than it is to move to some new course of action. Second, in many economic settings, changing actions requires real costs.

[1] For simplicity, define $a_i^{-1} = a_i^0$ for all i.

For example, entering or exiting a market involves obvious costs. Changing prices often requires printing new menus or advertising the new prices in some fashion.

A subtle question surrounds when such modifications of the standard model are "small." Most models of dynamic oligopoly ignore menu costs, evidently under the hypothesis that such small costs are irrelevant. However, while the cost of changing prices presumably is small relative to the present value of the firm's profits, these costs may be quite large relative to a day's worth of profits. One implication of our analysis is that the standard Folk Theorem does not hold when costs are large relative to one period's worth of payoff, even if they are small relative to the present value of payoffs.

Finally, consideration of infinitely repeated games is necessary to determine whether our earlier results indicate a discontinuity in the equilibrium outcome correspondence. To understand this, note that our earlier results indicated that subgame perfect equilibria with small ε and Δ are quite different from equilibria of finitely repeated games with $\varepsilon = 0$. Does this mean that the equilibrium outcome is discontinuous in ε at $\varepsilon = 0$? The difficulty in answering this question comes from the fact that our results all require ε large relative to Δ. Hence if ε goes to zero, to maintain our results, we must take Δ to zero as well. However, we wish to keep the total length of the game fixed. Hence if the length of a period goes to 0, the number of periods must go to infinity. Hence we are forced to turn to infinitely repeated games to address the question.

Here we show that different but also surprising results are possible with switching costs in infinitely repeated games if the switching cost is large relative to one period's worth of payoff. That is, just as in our earlier analysis, we consider switching costs which are small in the sense that the cost of one change of action is small relative to total game payoffs that can be earned over the entire horizon. However, in the case of primary interest, this cost is large relative to the game payoff which can be earned in a single period. As in our previous work, we parameterize the length of a period and primarily focus on the case where the length of a period and the switching cost are both small but the latter is large relative to the former.

To be more precise, consider player i's payoff to an infinite sequence of action profiles $a^0, a^1, \ldots$. Suppose that, as in the finite horizon case discussed above, actions are changed only at intervals of length Δ and the stage game payoffs are flow rates. It seems natural to view the switching cost as an immediate payment, not a flow cost. Under these assumptions, the agent's payoff to this infinite sequence of actions is

$$\sum_{t=0}^{\infty} \int_{t\Delta}^{(t+1)\Delta} e^{-rs} u_i(a^t)\, ds - \sum_{t=0}^{\infty} e^{-rt\Delta} \varepsilon I_i(a^{t-1}, a^t)$$

where $I_i(a, a') = 0$ if $a_i = a_i'$ and 1 otherwise as before[2] and r is the (continuous time)

[2]Also, as before, let $a_i^{-1} = a_i^0$.

discount rate. If we carry out the integration, normalize $r = 1$, and set $\delta = e^{-\Delta}$, we get

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t u_i(a^t) - \sum_{t=0}^{\infty}\delta^t \varepsilon I_i(a^{t-1}, a^t). \tag{2}$$

Note that as the length of a period, Δ, gets small, δ approaches 1. Note that the cost of one change of action relative to one period worth of payoff is on the order of $\varepsilon/(1-\delta)$ and so becomes large as $\Delta \downarrow 0$ or $\delta \uparrow 1$. On the other hand, the cost of one change of action relative to all game payoffs earned is on the order of $\varepsilon/\left[(1-\delta)\sum_t \delta^t\right] = \varepsilon$. Hence this is not affected by δ and converges to 0 as $\varepsilon \downarrow 0$. Consequently, this formulation enables us to make the cost of switching large relative to one period worth of payoff while keeping it small relative to the whole repeated game's payoffs.

By contrast, consider instead a simple variation on the usual discounting formulation, where we evaluate paths of play by the discounted sum over periods of the payoff in a period minus a switching cost if incurred in that period. More specifically, suppose player i's payoff to a sequence of action profiles $a^0, a^1, \ldots$ is

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t[u_i(a^t) - \varepsilon I_i(a^{t-1}, a^t)]. \tag{3}$$

This formulation gives no obvious way to shrink the switching cost relative to the whole repeated game worth of payoff without shrinking it relative to the payoff in a single period. As before, period length can be thought of as affecting the discount rate δ. However, here δ affects game payoffs and switching costs in the same way. Hence if we reduce ε, we *must* reduce it relative to $u_i(a)$ and thus relative to one period worth of payoff. (As we explain below, there is a sense in which our formulation using equation (2) nests this alternative as a special case.)

We consider two different infinitely repeated games. In the first, each player i evaluates sequences of actions by the payoff criterion in equation (2). We denote this game $G(\varepsilon, \delta)$ where $\varepsilon \in [0, \infty)$ and $\delta \in [0, 1)$. The second infinitely repeated game we consider has game payoffs defined by the limit of means criterion instead of discounting. More precisely, we define the game $G_\infty(\varepsilon)$ to be the game where each player i evaluates sequences of actions by the payoff criterion

$$\liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} u_i(a^t) - \varepsilon \#\{t \mid a_i^t \neq a_i^{t-1}\}$$

where # denotes cardinality.[3] As we explain in more detail in Section 4, there is a natural sense in which this game is the limit of our finitely repeated game as $\Delta \downarrow 0$. Let

[3]To ensure that this is well–defined, we allow $-\infty$ as a payoff. In other words, we treat payoffs in the repeated game as elements of $\mathbf{R} \cup \{-\infty\}$.

$\mathcal{U}(\varepsilon, \delta)$ denote the set of equilibrium payoffs of $G(\varepsilon, \delta)$ and let $\mathcal{U}_\infty(\varepsilon)$ denote the set of equilibrium payoffs of $G_\infty(\varepsilon)$.

First, we consider $\mathcal{U}(\varepsilon, \delta)$. In line with the intuition suggested above, our results show that the set of equilibrium payoffs is exactly the usual Folk Theorem set if the switching cost is small relative to a period worth of payoff but differs from the usual set if the cost is large relative to one period of payoff. In other words, consider the limit of the set $\mathcal{U}(\varepsilon, \delta)$ as $(\varepsilon, \delta) \to (0, 1)$. This limit will depend on the particular (ε, δ) sequence chosen. We show that if we consider sequences such that $\varepsilon/(1-\delta) \to 0$, then the limiting set of payoffs is the same as the usual Folk Theorem set. That is, for such sequences, $\mathcal{U}(\varepsilon, \delta)$ converges to the set of feasible, individually rational payoffs. (These limits are defined more precisely in Section 2.) Note that such sequences can be thought of as including the formulation in equation (3) as a special case. If we use the payoff criterion in equation (2) but set $\varepsilon = \hat{\varepsilon}(1-\delta)$, we obtain equation (3) with $\hat{\varepsilon}$ replacing ε. If we take $(\varepsilon, \delta) \to (0, 1)$ in equation (2) in such a way that $\varepsilon/(1-\delta) \to 0$, this corresponds to taking $(\hat{\varepsilon}, \delta) \to (0, 1)$ in equation (3).

At the opposite extreme, if we consider a sequence such that $\varepsilon/(1-\delta)$ goes to infinity, we get a limiting set of payoffs which differs from the Folk Theorem set in two ways. First, the payoff a player can guarantee himself is smaller with switching costs. Intuitively, if a player needs to randomize to avoid punishment, the expected costs of switching actions makes this too costly. That is, we must appropriately redefine individual rationality. Second, the notion of feasibility changes as well since the switching costs can dissipate payoffs even in the limit as $\varepsilon \downarrow 0$. For example, in the coordination game

	a	b
a	$3, 3$	$0, 0$
b	$0, 0$	$1, 1$

the usual Folk Theorem set is all payoff vectors (u_1, u_2) where $u_1 = u_2$ and $.75 \leq u_i \leq 3$. By contrast, if $(\varepsilon, \delta) \to (0, 1)$ with $\varepsilon/(1-\delta) \to \infty$ along the sequence, the set of equilibrium payoffs converges to the set of all (u_1, u_2) such that $(0, 0) \leq (u_1, u_2) \leq (3, 3)$.

Of course, the requirement that $\varepsilon/(1-\delta) \to \infty$ is quite strong. We show that two results regarding intermediate values of $\varepsilon/(1-\delta)$. First, we show that for any of the payoffs we obtain when ε is becoming arbitrarily large relative to $1-\delta$, we can approximately achieve this payoff while keeping $\varepsilon/(1-\delta)$ bounded. To make the approximation arbitrarily accurate requires making $\varepsilon/(1-\delta)$ arbitrarily large.

We also show that the requirement that $\varepsilon/(1-\delta)$ becomes arbitrarily large is driven entirely by the change in feasibility, not the change in individual rationality. More specifically, if the switching cost is on the order of two periods worth of payoff, then every feasible (in the traditional sense) and individually rational (in our modified sense) payoff

is a limiting equilibrium payoff. Hence even intuitively small switching costs require us to modify the usual definition of individual rationality.

As we explain in Section 4, one way to understand these results is to note that if ε is not arbitrarily small relative to $1-\delta$, then there is a sense in which $G(\varepsilon,\delta)$ is bounded away from $G(0,\delta)$. In other words, when we consider a sequence of (ε,δ) such that

$$\lim_{\varepsilon\downarrow 0,\delta\uparrow 1}\mathcal{U}(\varepsilon,\delta)\neq\lim_{\delta\uparrow 1}\mathcal{U}(0,\delta),$$

we must also have

$$\lim_{\varepsilon\downarrow 0,\delta\uparrow 1}G(\varepsilon,\delta)\neq\lim_{\delta\uparrow 1}G(0,\delta).$$

(These limits are defined more precisely in Sections 2 and 4.) In this sense, these results do not indicate a discontinuity in the equilibrium outcome correspondence with respect to (ε,δ).

Next, we turn to the case where we use the limit of means to evaluate game payoffs. In this case, the switching cost is larger than the payoffs for *any* finite number of periods, so, naturally, we would expect the cost to have the largest effect here. In fact, Theorem 6 shows that both of the two earlier differences between the equilibrium payoff set and the usual Folk Theorem set remain and a third is added. This third difference is strikingly unusual: payoffs that are supported by putting some weight on payoff vectors that are not individually rational (in the modified sense appropriate for switching costs) cannot be obtained. For example, in the Prisoners' Dilemma,

	C	D
C	$3,3$	$0,4$
D	$4,0$	$2,2$

the usual Folk Theorem set is all feasible payoffs where each player gets at least 2. As $\varepsilon\downarrow 0$, the set of equilibrium payoffs of $G_\infty(\varepsilon)$ converges to the set of payoffs where each player gets at least 2 and neither gets more than 3. We get this result because in this game, we cannot put any "weight" on the $(4,0)$ or $(0,4)$ payoff vector. (Payoff vectors which are not convex combinations of $(3,3)$ and $(2,2)$ are obtained by players dissipating payoffs through the switching costs.) Intuitively, with this formulation, any path of play in the game must have the property that players change actions only finitely often with probability one. Hence any path eventually "absorbs" in the sense that at some point, actions never change again. It is obvious that we cannot have an equilibrium where the players know that actions will never change again from (C,D) or (D,C) since the player getting 0 will change actions. What is less obvious is why we cannot have some kind of randomization that "hides" from the players the fact that no further changes of action will occur. (We do allow the players to condition on public randomizing devices, so we give the maximum possible ability for the players to use such strategies.) We show that

players must eventually become sure enough that no change will occur that they will deviate from any such proposed equilibrium.

Again, this result does not show a discontinuity in the equilibrium outcome correspondence with respect to ε. While

$$\lim_{\varepsilon \downarrow 0} \mathcal{U}_\infty(\varepsilon) \neq \mathcal{U}_\infty(0),$$

we will show in Section 4 that

$$\lim_{\varepsilon \downarrow 0} G_\infty(\varepsilon) \neq G_\infty(0).$$

Finally, we use Theorem 6 and results in Lipman–Wang [2000] to address whether our earlier results demonstrate a discontinuity in the equilibrium outcome correspondence. Let $G_f(\varepsilon, \Delta)$ be the game studied in Lipman–Wang [2000] using the payoff function (1) above and let $\mathcal{U}_f(\varepsilon, \Delta)$ denote the set of equilibrium payoffs. Analogously to the above, we show that when

$$\lim_{\varepsilon \downarrow 0, \Delta \downarrow 0} \mathcal{U}_f(\varepsilon, \Delta) \neq \lim_{\Delta \downarrow 0} \mathcal{U}_f(0, \Delta),$$

it is because

$$\lim_{\varepsilon \downarrow 0, \Delta \downarrow 0} G_f(\varepsilon, \Delta) \neq \lim_{\Delta \downarrow 0} G_f(0, \Delta).$$

Thus there is no discontinuity with respect to (ε, Δ). On the other hand, as above, there is in general a discontinuity with respect to Δ for any fixed $\varepsilon > 0$. Specifically,

$$\lim_{\Delta \downarrow 0} \mathcal{U}_f(\varepsilon, \Delta) \neq \mathcal{U}_\infty(\varepsilon)$$

even though

$$\lim_{\Delta \downarrow 0} G_f(\varepsilon, \Delta) = G_\infty(\varepsilon).$$

The possibility of such a discontinuity with $\varepsilon = 0$ is well known, but the discontinuity when $\varepsilon > 0$ is of a very different nature. The known discontinuity for $\varepsilon = 0$ is simply the difference between finitely and infinitely repeated games. However, this discontinuity is a failure of lower semicontinuity, not upper, as the limiting set of equilibrium payoffs is smaller than the set at the limit. The discontinuity with $\varepsilon > 0$ may have the limiting set larger than the set at the limit. Also, the discontinuity for $\varepsilon = 0$ occurs for a different set of games than the discontinuity for $\varepsilon > 0$.

Our results differ from those of Chakrabarti [1990] who considers a similar model. He analyzes infinitely repeated games with a more general "inertia cost" than we consider. His payoff criterion, however, does not fit into the class we consider. Specializing

his switching cost to our setting, he assumes players evaluate payoffs to the sequence $(a^0, a^1, \ldots)$ by

$$\liminf_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} [u_i(a^t) - \varepsilon I_i(a^{t-1}, a^t)],$$

the limit of means analog of (3) above. His formulation is a special case of a dynamic game as considered by Dutta [1995], while our formulation is not. Using the results of Dutta [1995], one can show that Chakrabarti's set of equilibrium payoffs differs from the usual Folk Theorem set in two ways, namely the two present in our Theorem 3. That is, both individual rationality and feasibility must be redefined to take account of the switching costs.[4] However, the third effect we obtain in Theorem 6 is not present. He does not discuss continuity issues.

In the next section, we state the model. In Section 3, we give our characterizations of equilibrium payoffs. In Section 4, we define a notion of closeness of games and use this to consider continuity of the equilibrium outcome correspondence. Proofs not in the text are contained in the Appendix.

2 Model

Fix a finite stage game $G = (A, u)$ where $A = A_1 \times \ldots \times A_I$, each A_i is finite and contains at least two elements, and where $u : A \to \mathbf{R}^I$. Let S_i denote the set of mixed stage game strategies — that is, S_i is the set of randomizations over A_i. We allow the players to use public randomizing devices, so a strategy for the repeated game can depend on the history of play as well as the outcome of the public randomization. For simplicity, we will suppose that there is an iid sequence of random variables, ξ_t, which are uniformly distributed on $[0, 1]$ which all players observe. A strategy for player i, then, is a function from the history of past actions and the realization of the randomizations (up to and including the current period) into S_i. That is, it is a function $\sigma_i : \cup_{t=0}^{\infty} A^t \times [0, 1]^t \to S_i$ where $A^0 \times [0, 1]^0$ is defined to be the singleton set containing the "empty history" e.

Remark 1 As shown by Fudenberg and Maskin [1991], the use of public randomization is purely a matter of convenience in the usual repeated game. More specifically, one can obtain the same characterization of equilibrium payoffs without public randomizations. However, the assumption is not as innocuous here. While it is not needed for any of the other results, the result of Theorem 6 is not true in general without public randomization. The reason is that the equilibrium construction which is typically used to replace public randomization requires numerous changes of action. Since such changes of action are

[4] Chakrabarti states his results in a different but equivalent way.

costly in this model, such behavior can be difficult to support as an equilibrium. On the other hand, the most interesting aspect of Theorem 6 is the payoffs which *cannot* be achieved. Since allowing public randomization can only increase the set of equilibrium payoffs, this is the most interesting case to consider for that result.

The payoffs in the game $G(\varepsilon, \delta)$, $\varepsilon \geq 0$, $\delta \in [0, 1)$, are defined as follows. Given a sequence of actions $(a^0, a^1, \ldots)$ where $a^t = (a_1^t, \ldots, a_I^t)$, i's payoff from this sequence is

$$(1-\delta)\sum_{t=0}^{\infty} \delta^t u_i(a^t) - \sum_{t=0}^{\infty} \delta^t \varepsilon I_i(a^{t-1}, a^t)$$

where $I_i(a^{t-1}, a^t) = 1$ if $a_i^{t-1} \neq a_i^t$ and 0 otherwise.[5] Let $\mathcal{U}(\varepsilon, \delta)$ denote the closure of the set of subgame perfect equilibrium payoffs in $G(\varepsilon, \delta)$.

Letting # denote cardinality, we define the payoff from this sequence in the game $G_\infty(\varepsilon)$ to be

$$\left[\liminf_{T\to\infty} \frac{1}{T}\sum_{t=0}^{T-1} u_i(a^t)\right] - \varepsilon \#\{t \mid a_i^t \neq a_i^{t-1}\}$$

if this is a real number and $-\infty$ otherwise. Note that the payoff is a well defined real number if i changes actions only finitely often. However, if i changes actions infinitely often, then the switching cost makes this payoff arbitrarily negative; hence we define the payoff to be $-\infty$. Let $\mathcal{U}_\infty(\varepsilon)$ denote the closure of the set of subgame perfect equilibrium payoffs in $G_\infty(\varepsilon)$.

We are interested in the set of $\mathcal{U}(\varepsilon, \delta)$ for ε very close to 0 and δ very close to 1. As we will see, this set will depend on the relationship of ε and δ. In addition, we are less interested in specific values of ε and δ than in general properties for (ε, δ) near $(0, 1)$. Consequently, it will prove most convenient to consider the set of limit points of $\mathcal{U}(\varepsilon_n, \delta_n)$ as $n \to \infty$ for various sequences $(\varepsilon_n, \delta_n)$.

In particular, for any $k \in [0, \infty]$, we define the set

$$\lim_{\varepsilon\downarrow 0, \delta\uparrow 1}^{k} \mathcal{U}(\varepsilon, \delta)$$

to be the set of $u \in \mathbf{R}^I$ such that there are sequences ε_n, δ_n, and u^n such that

$$\varepsilon_n > 0, \quad \delta_n \in [0, 1), \quad \text{and} \quad u^n \in \mathcal{U}(\varepsilon_n, \delta_n), \quad \forall n,$$

$$\lim_{n\to\infty} \frac{\varepsilon_n}{1-\delta_n} = k,$$

[5] We define a_i^{-1} to be equal to a_i^0 for any sequence of actions $(a^0, a^1, \ldots)$. In other words, there is no cost of "changing" actions in the first period regardless of the action played in that period.

and $(\varepsilon_n, \delta_n, u^n) \to (0, 1, u)$ as $n \to \infty$. Intuitively, then, k is a measure of how large ε is relative to $1 - \delta$ along the sequence. The case of $k = 0$ is effectively the situation where we take ε to zero first and then δ to 1; the case of $k = \infty$ is analogous to the reverse order of limits.

The issue of ε relative to $1 - \delta$ is irrelevant in the game $G_\infty(\varepsilon)$. Hence we simply define

$$\lim_{\varepsilon \downarrow 0} \mathcal{U}_\infty(\varepsilon)$$

to be the set of $u \in \mathbf{R}^I$ such that there exist sequences ε_n and u^n with

$$\varepsilon_n > 0 \text{ and } u^n \in \mathcal{U}_\infty(\varepsilon_n), \quad \forall n$$

with $(\varepsilon_n, u^n) \to (0, u)$ as $n \to \infty$.

The usual Folk Theorem sets have $\varepsilon = 0$ — that is, they are $\mathcal{U}_\infty(0)$ and $\lim_{\delta \uparrow 1} \mathcal{U}(0, \delta)$. We define the latter analogously to the approach used above. That is, $\lim_{\delta \uparrow 1} \mathcal{U}(0, \delta)$ is the set of u such that there is a sequence δ_n converging to 1 from below and a sequence u^n converging to u with $u^n \in \mathcal{U}(0, \delta_n)$ for all n.

Define i's *reservation payoff*, v_i, by

$$v_i = \min_{s_{\sim i} \in S_{\sim i}} \left[\max_{s_i \in S_i} u_i(s_i, s_{\sim i}) \right].$$

Let

$$R = \{u \in \mathbf{R}^I \mid u \geq v\}$$

denote the usual set of individually rational payoffs where $v = (v_1, \ldots, v_I)$.[6] In the case of two players,[7] the classic minmax theorem states that the order of the minimization and maximization don't matter. That is, in this case, v_i is also equal to $\max_{s_i \in S_i} [\min_{s_{\sim i} \in S_{\sim i}} u_i(s_i, s_{\sim i})]$. However, even in the two player case, if i is restricted to pure strategies, the order matters very much. We will see that the relevant reservation utility for i in the game with switching costs is what we will call i's *pure reservation payoff*, w_i, defined by

$$w_i = \max_{a_i \in A_i} \left[\min_{s_{\sim i} \in S_{\sim i}} u_i(a_i, s_{\sim i}) \right].$$

Let

$$W = \{u \in \mathbf{R}^I \mid u \geq w\}$$

denote what we will call the set of *weakly individually rational* payoffs, where $w = (w_1, \ldots, w_I)$. Note that $w_i \leq v_i$ for all i so $R \subseteq W$. Any action $a_i \in A_i$ such that

$$\min_{s_{\sim i} \in S_{\sim i}} u_i(a_i, s_{\sim i}) = w_i$$

[6] Given vectors x and y, we use $x \geq y$ to mean greater than or equal to in every component and $x \gg y$ to mean strictly larger in every component.

[7] Or if we allow the players other than i to use correlated strategies.

will be referred to as *a maxmin action for i*.

It is worth noting for future use, that

$$w_i = \max_{a_i \in A_i} \left[\min_{a_{\sim i} \in A_{\sim i}} u_i(a_i, a_{\sim i}) \right].$$

To see this, simply note that for any $a_i \in A_i$ and any $j \neq i$, $u_i(a_i, s_{\sim i})$ is linear in j's mixed strategy. Hence the value of this expression when minimized over s_j is unaffected by restricting j to pure actions. We exploit this fact in what follows.

For any set $B \subseteq \mathbf{R}^I$, let $\text{conv}(B)$ denote its convex hull. Let U denote the set of payoffs feasible from pure strategies and let F denote the usual set of feasible payoffs. That is,

$$U = \{u \in \mathbf{R}^I \mid u = u(a), \text{ for some } a \in A\}$$

and $F = \text{conv}(U)$. For comparison purposes, we first state the usual Folk Theorem.

We define a game to be *regular* if there is $u = (u_1, \ldots, u_I) \in F \cap R$ such that $u_i > v_i$ for all i.

Theorem 1 (The Folk Theorem) *For any regular game,*

A. $\mathcal{U}_\infty(0) = F \ \cap \ R$.

B. If, in addition, F *has dimension* I*,*

$$\lim_{\delta \uparrow 1} \mathcal{U}(0, \delta) = F \ \cap \ R.$$

This result is a trivial extension of theorems in Fudenberg and Tirole [1991], Chapter 5, and so we omit the proof.

Remark 2 The restriction to regular games in Theorem 1 is generally not stated but is often used in some form. For example, one typical version of the Folk Theorem is that the equilibrium payoff set includes all feasible payoffs where each player i receives strictly more than v_i. We use the assumption to be able to give an exact statement of equilibrium payoff sets without having to consider tedious boundary calculations. The additional assumption we use in Theorem 1.B is the most simply stated sufficient condition. It could be replaced by the weaker NEU condition of Abreu, Dutta, and Smith [1994].

3 Results

First, we consider the case of discounting. Recall that

$$\lim_{\varepsilon\downarrow 0,\delta\uparrow 1}^{k}\ \mathcal{U}(\varepsilon,\delta)$$

is the set of limiting equilibrium payoffs in $G(\varepsilon,\delta)$ for sequences (ε,δ) converging to $(0,1)$ such that $\varepsilon/(1-\delta)\to k$. As explained in the introduction, if the switching cost is small relative to one period worth of payoff, we expect to obtain the same results as in the usual analysis. That is, we expect the $\lim^k$ set to equal the usual Folk Theorem set if k is small. This intuition is confirmed by

Theorem 2 *For any regular game such that F has dimension I,*

$$\lim_{\varepsilon\downarrow 0,\delta\uparrow 1}^{0}\ \mathcal{U}(\varepsilon_n,\delta_n) = F\ \cap\ R.$$

That is, if $\varepsilon/(1-\delta)$ goes to 0 along the sequence, the limiting payoff set is the set of feasible, individually rational payoffs.

To see that the limiting set of payoffs is contained in $F\cap R$, suppose not. First, suppose that $u\in\lim^0_{\varepsilon\downarrow 0,\delta\uparrow 1}\mathcal{U}(\varepsilon_n,\delta_n)$, but $u\notin R$. Let $(\varepsilon_n,\delta_n,u^n)$ be the required sequence converging to $(0,1,u)$ for which $u^n\in\mathcal{U}(\varepsilon_n,\delta_n)$ and $\varepsilon_n/(1-\delta_n)\to 0$. Fix any i for whom $u_i<v_i$. Since $\varepsilon_n/(1-\delta_n)\to 0$ and $u_i^n\to u_i$, it must be true that

$$u_i^n < v_i - \frac{\varepsilon_n}{1-\delta_n}$$

for all n sufficiently large. But then even if it requires changing actions every period, i can switch to the strategy of choosing a myopic best reply in every period to the strategies of the opponents for that period and be better off, a contradiction.

Next, suppose $u\in\lim^0_{\varepsilon\downarrow 0,\delta\uparrow 1}\mathcal{U}(\varepsilon_n,\delta_n)$, but $u\notin F$. Again, fix the required sequence $(\varepsilon_n,\delta_n,u^n)$. Since $u^n\in\mathcal{U}(\varepsilon_n,\delta_n)$, there is some probability distribution, say q_n over A such that

$$u_i^n = \sum_{a\in A} q_n(a)u_i(a) - X_n$$

where X_n is the expected discounted switching costs in the equilibrium. We know that X_n must be bounded above by the cost of switching actions in every period so $X_n\le\varepsilon_n/(1-\delta_n)$. Hence $X_n\to 0$ as $n\to\infty$. So

$$u = \lim_{n\to\infty} u^n = \lim_{n\to\infty}\sum_{a\in A} q_n(a)u_i(a).$$

Hence $u \in F$.

So we see that

$$\lim_{\varepsilon\downarrow 0,\delta\uparrow 1}^{0} \mathcal{U}(\varepsilon_n, \delta_n) \subseteq F \cap R.$$

The proof that every payoff in $F \cap R$ is a limiting equilibrium payoff is a simple extension of arguments in Fudenberg and Tirole and so is omitted.[8]

As discussed in the introduction, we expect differences from the standard model when the switching cost is large relative to one period's worth of payoff. That is, we expect $\lim^{k}_{\varepsilon\downarrow 0,\delta\uparrow 1} \mathcal{U}(\varepsilon, \delta)$ to differ from the usual Folk Theorem set when k is sufficiently large. This intuition is confirmed by the next result.

First, we require a few definitions. Given a set $B \subseteq \mathbf{R}^I$, let $c(B)$ denote the comprehensive, convex hull of B. That is, $c(B)$ is the set of points less than or equal to a convex combination of points in B. Define $F^* = c(U)$. This is the feasible set of payoffs when we allow players the ability to "throw away" utility.

Define a game to be *weakly regular* if there is a payoff vector $u = (u_1, \ldots, u_I) \in F$ such that $u_i > w_i$ for all i. Since $v_i \geq w_i$, obviously, any regular game is weakly regular.[9]

While the following result is a corollary to a more general result below, we begin with it for the sake of clarity.

Theorem 3 *For any weakly regular game,*

$$\lim_{\varepsilon\downarrow 0,\delta\uparrow 1}^{\infty} \mathcal{U}(\varepsilon, \delta) = F^* \cap W.$$

That is, the limiting payoff set is the set of feasible payoffs (taking into account the ability to dissipate payoffs by switching actions) which are weakly individually rational.

Thus we have a simple characterization of equilibrium payoffs in the two extreme cases, where $\varepsilon/(1-\delta)$ converges to 0 and where it converges to ∞. As one might expect, the middle ground is more complex. Our next result shows that the transition between these extremes is gradual in the sense that as k increases, we gradually fill in all the payoffs in $F^* \cap W$ which are not in $F \cap R$. More precisely,

[8]The argument is to note that their proof for the observable mixed strategy case involves *strict* payoff comparisons. Hence small enough switching costs cannot affect the optimality of the strategies in question. Since the public randomization effectively creates observable mixed strategies, this completes the argument.

[9]Note that it would be equivalent to define weak regularity using F^* in place of F.

Theorem 4 *For any weakly regular game and any $\eta > 0$, there is a k_η such that for all $k \geq k_\eta$, for all $u \in F^* \cap W$, there is a u' within η of u with*

$$u' \in \lim_{\varepsilon \downarrow 0, \delta \uparrow 1}^{k} \mathcal{U}(\varepsilon, \delta).$$

As $\eta \downarrow 0$, $k_\eta \to \infty$.

Thus the set of limiting equilibrium payoffs when $\varepsilon/(1-\delta)$ converges to k approximates $F^* \cap W$ with the precision of the approximation improving as we increase k. However, to generate the entire set, we need $k \to \infty$ in general.

A natural question to ask is how large the limiting set of equilibria is for "moderate" values of k. If the set differs from the usual Folk Theorem set only when k is extremely large, the interest in Theorems 3 and 4 is somewhat limited. The next result shows that we only need the switching costs to be on the order of two periods worth of payoff to generate a very significant difference from the usual Folk Theorem. More specifically, we have

Theorem 5 *For any weakly regular game and any*

$$k > 2 \max_i \left[\max_{a \in A} u_i(a) - w_i \right],$$

we have

$$F \ \cap \ W \subseteq \lim_{\varepsilon \downarrow 0, \delta \uparrow 1}^{k} \mathcal{U}(\varepsilon, \delta).$$

That is, all feasible (in the traditional sense) and weakly individually rational payoffs are included in the limiting equilibrium set for "moderate" k.

To understand our interpretation of the condition on k, suppose we have a sequence $(\varepsilon_n, \delta_n)$ converging to $(0, 1)$ with $\varepsilon_n/(1-\delta_n)$ converging to a k satisfying the condition of Theorem 5. For large n, $\varepsilon_n/(1-\delta_n)$ is very close to k, so

$$\varepsilon_n > (1-\delta_n) 2 \max_i \left[\max_{a \in A} u_i(a) - w_i \right].$$

The left–hand side is the cost of a change of actions in the current period. The right–hand side is approximately the gain in payoff over two periods from moving from w_i to $\max_{a \in A} u_i(a)$ for some player i. As we will see shortly, w_i is the lowest payoff that can be imposed on a player, so this payoff gain is, roughly speaking, the large "plausible" payoff gain to a change in actions. In this sense, this condition says that the switching cost is bigger than the largest potentially relevant two–period payoff gain.

The proofs of these results are in the Appendix, but here we sketch the idea. First, it is obvious that for any k, the limiting payoff set is contained in $F^* \cap W$, so the critical issue is when these payoffs can be generated by some equilibrium.

First, we need to establish that W is the appropriate version of individual rationality. To see this, consider the payoff a player receives if the others are trying to minimize his payoff. If the other players continually move to the action which minimizes his payoff given the action he has most recently played, he will either stop changing actions and get his pure reservation payoff or change actions every period. If $\varepsilon/(1-\delta)$ converges to a large enough number, these switching costs become too large for this second option to be optimal. As it turns out, the bound on k given in Theorem 5 is sufficient to establish this. Hence when this condition holds, we can force a player down to his pure reservation payoff.

To see that F^* is the appropriate definition of feasibility is a little more complex. Suppose we wish to construct strategies generating a particular payoff vector u in F^*. Any such payoff can be written in the form

$$u = \sum_{a \in A} \alpha(a) u(a) - x$$

where α is a probability distribution over A and x is a vector of costs. It is tedious but not difficult to show that for large enough k, any such payoff can be approximately generated by constructing an appropriate cycle of actions. The cycle is chosen so that the relative frequencies of actions over the cycle approximates α and the relative frequency of changes of actions over the cycle generates x. To illustrate the latter, suppose, for example, that the cycle is of length N and that player i changes actions in the first N_i periods of the cycle only. Then his switching costs over the entire infinite horizon are

$$\varepsilon \left[\sum_{t=0}^{N_i-1} \delta^t\right]\left[\sum_{t=0}^{\infty} \delta^{Nt}\right] = \frac{\varepsilon}{1-\delta} \left[\sum_{t=0}^{N_i-1} \delta^t\right] \frac{1-\delta}{1-\delta^N} = \frac{\varepsilon}{1-\delta} \left[\frac{\sum_{t=0}^{N_i-1} \delta^t}{\sum_{t=0}^{N-1} \delta^t}\right].$$

The second term in the last expression converges to N_i/N as $\delta \to 1$. Hence if we take the limit as $(\varepsilon, \delta) \to (0,1)$ along a sequence for which $\varepsilon/(1-\delta) \to k$, we see that the switching costs converge to $k(N_i/N)$. Hence by setting the frequency of i's action changes over the cycle appropriately, we can generate whatever switching cost is needed. In short, this cycle can be chosen so that as $(\varepsilon, \delta) \to (0,1)$, the payoff converges to approximately u, where the approximation can be made arbitrarily close for large enough k. Thus any payoff in F^* is feasible even in the limit as $(\varepsilon, \delta) \to (0,1)$.

Given these two facts, the completion of the proofs of Theorems 3 and 4 is similar to a standard Folk Theorem construction.

To explain Theorem 5, we need to first clarify why k goes to infinity as our approximation error η goes to zero. To see this, consider the following game:

$$\begin{array}{cc} L & R \\ U \quad 0,-1 & 2,2 \\ D \quad -2,-2 & -1,0 \end{array}$$

It is not hard to see that $w_1 = w_2 = 0$. Hence $F^* \cap W$ is the set of (u_1, u_2) with $0 \leq u_i \leq 2$ for both i. Suppose we want to generate the payoff $(1, 2)$. The way we do this is to construct a cycle. More specifically, player 2 always plays R. In the cycle, player 1 changes actions N_1 times, then the players play (U, R) N_2 times. The frequency of play of the "wrong" action, (D, R), is approximately $(1/2)N_1/(N_1 + N_2)$. As explained above, player 1's switching cost converges as $(\varepsilon, \delta) \to (0, 1)$ to $kN_1/(N_1 + N_2)$. We need to choose N_1 and N_2 so that the frequency of (D, R) is close to 0 and so that player 1's switching cost over the cycle is close to 1. The former requires us to make N_1/N_2 very close to zero. Given this, the latter requires k very large. In particular, the closer we wish to approximate the payoff $(1, 2)$, the larger k will have to be. Hence $k \to \infty$ as $\eta \downarrow 0$.

On the other hand, this problem arises only when we want to construct certain payoffs in F^* which are not in F. If a payoff is in F, we do not need the initial switching phase to dissipate payoffs. In fact, for *any* k, we can find a cycle which approximately generates any payoff $u \in F$ arbitrarily well. This fact together with the observation above that the condition in Theorem 5 is sufficient to make w the relevant notion of individual rationality explains why Theorem 5 holds.

We obtain a more unusual characterization in the case of the limit of $\mathcal{U}_\infty(\varepsilon)$ as $\varepsilon \downarrow 0$. Let $U_\geq$ denote those points in U which are greater than w. That is,

$$U_\geq = \{u = (u_1, \ldots, u_I) \in U \mid u_i \geq w_i, \ \forall i\}.$$

For the next result, we need one additional assumption which we call *rewardability*. We say that a game satisfies rewardability if there is a payoff vector $u = (u_1, \ldots, u_I) \in U_\geq$ with $u_i > w_i$ for all i. It is worth emphasizing that this property is much stronger than regularity.

To see the idea behind the name, suppose this assumption does not hold. As mentioned in the introduction, in $G_\infty(\varepsilon)$, the only u vectors which can be achieved infinitely often with positive probability are those in $U_\geq$. If for some player i, all these vectors give him w_i, then he cannot be rewarded for aiding in the punishment of a deviator. This complication restricts the set of equilibria in a complex fashion as we explain in more detail below.

Theorem 6 *For any game satisfying rewardability,*

$$\lim_{\varepsilon \downarrow 0} \mathcal{U}_\infty(\varepsilon) = c(U_\geq) \ \cap \ W.$$

To see how this differs from the payoff set from Theorem 3, note that we can write that set as $c(U) \cap W$. In this form, the difference is obvious: the payoff set of Theorem 6 only puts weight on payoffs which are weakly individually rational, not all pure strategies.

The proof that any payoff in $c(U_\geq) \cap W$ is a limiting equilibrium payoff is similar to standard Folk Theorem arguments. The more unusual part of the proof is the demonstration that no payoff outside this set can be close to an equilibrium payoff. We sketch the idea in the context of the Prisoners' Dilemma we used in the introduction:

	C	D
C	$3,3$	$0,4$
D	$4,0$	$2,2$

Let $\mathcal{P}(C,C)$ denote the set of infinite sequences of actions which eventually "absorb" at (C,C) — that is, sequences with the property that for some T, the actions played at any $t \geq T$ are (C,C). Define $\mathcal{P}(C,D)$, etc., analogously. Note that any sequence of actions which is not in $\mathcal{P}(C,C)$, $\mathcal{P}(C,D)$, $\mathcal{P}(D,C)$, or $\mathcal{P}(D,D)$ has at least one player changing actions infinitely often. If any player has a positive probability of switching actions infinitely often, his expected payoff is $-\infty$ and so his strategy cannot be optimal. Hence any equilibrium has to put zero probability on such an event. That is, the sets $\mathcal{P}(C,C)$, $\mathcal{P}(C,D)$, $\mathcal{P}(D,C)$, and $\mathcal{P}(D,D)$ must have probability 1 in total. The main claim of Theorem 6 is that the sets $\mathcal{P}(C,D)$ and $\mathcal{P}(D,C)$ must have zero probability in equilibrium.

To see this, suppose, say, $\mathcal{P}(C,D)$ has probability $\mu > 0$. Clearly, it cannot have probability 1. If it did, player 1's payoff in equilibrium would be 0, while playing a constant action of D gives him a payoff of 2, a contradiction. Clearly, too, there can be no history with the property that the probability of $\mathcal{P}(C,D)$ conditional on this history is 1. If it were, then for any switching cost less than 2, player 1 could profitably deviate on that history to a constant action of D and be better off.

What is not so transparent is whether it is possible to construct the public randomizations in such a way that play does absorb at (C,D) but this fact is hidden from player 1. In other words, can we construct strategies with the property that there is a positive probability that (C,D) is played from a certain point onward and yet along this path of play, player 1 always believes there is a nontrivial probability that some other action will be played in the future?

In fact, the answer to this question is no. To see this, suppose (C,D) is played at period t and consider the probability player 1 gives to the event $\mathcal{P}(C,D)$ conditional

on this fact. Clearly, any path of play which absorbs at a different action profile at a period before t must have zero probability at this point. Hence for large t, the conditional probability that the play path is $\mathcal{P}(C,C)$, $\mathcal{P}(D,C)$, or $\mathcal{P}(D,D)$ must be getting small. At the same time, this fact that (C,D) is played at t cannot rule out the possibility that play has already absorbed at (C,D). Hence as t gets large, the conditional probability on $\mathcal{P}(C,D)$ must converge to 1. But once this conditional probability is large enough, player 1 will certainly deviate to D, a contradiction. Note that this argument actually implies that we cannot have a *Nash* equilibrium putting positive probability on $\mathcal{P}(C,D)$, much less a subgame perfect equilibrium.

Remark 3 To see what happens with games which violate rewardability, consider the following game:

$$\begin{array}{cc} L & R \\ U \quad 0,1 & 1,-2 \\ D \quad -2,0 & -1,-1 \end{array}$$

It is not hard to see that $w_1 = w_2 = 0$ so $U_{\geq} = \{(0,1)\}$. Hence rewardability fails because the only vector in $U_{\geq}$ gives player 1 his pure reservation payoff. For this game,

$$c(U_{\geq}) \cap W = \{(0,x) \mid 0 \leq x \leq 1\}.$$

However, the unique equilibrium payoff is $(0,1)$. Intuitively, this is because player 1 must get a payoff of 0 in any subgame perfect equilibrium. Hence he cannot be induced to change actions and so will not punish 2 for deviations. Hence 2 must receive a payoff of 1. It is not hard to see how one could give a characterization of the limiting equilibrium set without rewardability. Analogously to Wen [1994], one can explicitly work out the way in which punishment is constrained to give an exact characterization of the limiting equilibrium set. More specifically, if for some player i, every vector in $U_{\geq}$ gives him a payoff of w_i, then he will must play a fixed action at every history of every equilibrium. We can set this player to the constant action he must play and solve the "reduced game" among the remaining players, iterating this procedure as necessary.

Remark 4 It is worth noting that the proof of Theorem 6 also shows that the reduction in the set of payoffs is not entirely a "vanishing ε" phenomenon. More specifically, the proof shows that there is a $\bar{\varepsilon} > 0$ such that for all $\varepsilon \in (0,\bar{\varepsilon})$, $\mathcal{U}_\infty(\varepsilon)$ is contained in $c(U_{\geq}) \cap W$.

It is natural to wonder why we get such a dramatic difference between the discounting and limit of means cases. This is much more than the dimensionality issue that comes up in the analysis of repeated games without switching costs. The difference here hinges, as with most of our results, on the relationship between the switching cost and the length of

a period. To see the point, consider the discounting case and suppose we take the limit of $\mathcal{U}(\varepsilon, \delta)$ as $(\varepsilon, \delta) \to (0, 1)$ along a sequence where $\varepsilon/(1-\delta) \to \infty$. As argued above, we can think of this as making the period length short relative to the switching cost. However, this effect can be undone in equilibrium. To see the point, note that we could always construct equilibria in which the players act as if a block of k periods was only one period. That is, they only change actions at intervals of k periods.[10] By constructing such equilibria, we can effectively make the length of a period arbitrarily long relative to the switching cost.

For any $\delta < 1$, this matters. However, in the limit of means case, it does not. In $G_\infty(\varepsilon)$, only the number of times the players change actions matters, not the intervals at which these changes occur. Hence this is the only situation where the players cannot endogenously alter the relationship between switching costs and payoffs in "a period."

4 Continuity

To say whether our results indicate a discontinuity in the equilibrium outcome correspondence, we must first define a notion of convergence of games. We say that a sequence of games G^n converges to a game G if for every player and every sequence of action profiles, the payoff in G^n converges as $n \to \infty$ to the payoff in G. That is, we say that

$$\lim_{n\to\infty} G(\varepsilon_n, \delta_n) = G(\varepsilon, \delta)$$

iff for all i and all $(a^0, a^1, \ldots) \in A^\infty$

$$\liminf_{n\to\infty} \left[(1-\delta_n)\sum_{t=0}^{\infty} \delta_n^t u_i(a^t) - \sum_{t=0}^{\infty} \delta_n^t \varepsilon_n I_i(a^{t-1}, a^t)\right] = (1-\delta)\sum_{t=0}^{\infty} \delta^t u_i(a^t) - \sum_{t=0}^{\infty} \delta^t \varepsilon I_i(a^{t-1}, a^t).$$

We define convergence to $G_\infty(\varepsilon)$ or convergence of the sequence $G_\infty(\varepsilon_n)$ analogously. We emphasize that we allow convergence of a payoff to $-\infty$ in this definition. That is, we define the limit of a monotonically decreasing sequence with no lower bound as $-\infty$. In particular, any sequence of actions where some player changes actions every period will have a payoff of $-\infty$ for that player in $G_\infty(\varepsilon)$. Hence if some sequence of games is to converge to $G_\infty(\varepsilon)$, we must allow a sequence of payoffs to converge to $-\infty$.[11]

[10] Of course, one cannot prevent players from deviating from this and changing actions more frequently. However, the punishment for deviations can also come more quickly as well.

[11] Because of this, our definition of convergence of a sequence of games is not the same as that generated by defining the distance between two games to be the supremum payoff difference over players and sequences of action profiles. In particular, Lemma 3 below would not hold under this alternative definition. To see this, note that every payoff in $G_f(\varepsilon, \Delta)$ is finite. Fix any sequence of actions where player i changes actions infinitely often. Then the difference in the payoffs to i from this sequence between $G_f(\varepsilon, \Delta)$ and $G_\infty(\varepsilon)$ is ∞ for every $\varepsilon > 0$ and $\Delta > 0$. Hence as $\Delta \to 0$, the distance between $G_f(\varepsilon, \Delta)$ and $G_\infty(\varepsilon)$ according to this definition does not go to zero.

Using this, we have

Lemma 1 *For any sequence $(\varepsilon_n, \delta_n)$, we have*

$$\lim_{n\to\infty} G(\varepsilon_n, \delta_n) = \lim_{n\to\infty} G(0, \delta_n)$$

if and only if $\varepsilon_n/(1-\delta_n) \to 0$.

Proof. Fix any i and any sequence of actions $a^0, a^1, \ldots$ where i changes actions every period. Note that the payoff to i from this sequence in $G(0, \delta_n)$ is bounded between $\min_{a\in A} u_i(a)$ and $\max_{a\in A} u_i(a)$. Hence as $n \to \infty$, the payoff cannot converge to $-\infty$. Hence if the payoff in $G(0, \delta_n)$ has the same limit as $n \to \infty$ as the payoff in $G(\varepsilon_n, \delta_n)$, the latter must also be finite, so the difference in payoffs must converge to 0 as $n \to \infty$. Note that the payoff to i in $G(0, \delta_n)$ minus the payoff in $G(\varepsilon_n, \delta_n)$ is

$$\sum_{t=0}^{\infty} \delta_n^t \varepsilon_n = \frac{\varepsilon_n}{1-\delta_n}.$$

Hence if $\varepsilon_n/(1-\delta_n) \not\to 0$, $G(\varepsilon_n, \delta_n)$ and $G(0, \delta_n)$ cannot have the same limit.

For the converse, fix any i and any sequence of actions $a^0, a^1, \ldots$, not necessarily one where i changes actions every period. Then the payoff in $G(0, \delta_n)$ minus the payoff in $G(\varepsilon_n, \delta_n)$ is

$$\sum_{t=0}^{\infty} \delta_n^t \varepsilon_n I_i(a^t, a^{t-1}) \leq \frac{\varepsilon_n}{1-\delta_n}.$$

Hence if $\varepsilon_n/(1-\delta_n) \to 0$, $G(\varepsilon_n, \delta_n)$ and $G(0, \delta_n)$ do have the same limit. ∎

Given this result, we see that Theorems 3, 4, and 5 do not indicate a discontinuity at $(\varepsilon, \delta) = (0, 1)$ in the equilibrium outcome correspondence. In particular, in the class of games considered, the limiting set of payoffs in $G(\varepsilon_n, \delta_n)$ and $G(0, \delta_n)$ as $n \to \infty$ differ only if the limiting games differ.

That Theorem 6 does not imply a discontinuity at $\varepsilon = 0$ is a corollary to

Lemma 2 *Fix any strictly positive sequence ε_n. Then*

$$\lim_{n\to\infty} G_\infty(\varepsilon_n) \neq G_\infty(0).$$

Proof. Fix any i and any sequence of actions where i changes actions infinitely often. i's payoff in $G_\infty(\varepsilon_n)$ is $-\infty$ for every n, while his payoff in $G_\infty(0)$ is bounded from below by $\min_{a\in A} u_i(a)$. Hence i's payoff in $G_\infty(\varepsilon_n)$ does not converge to his payoff in $G_\infty(0)$ as $n \to \infty$. ∎

Remark 5 It is common to describe the infinitely repeated game with the limit of means criterion as the limit of the game with discounting as $\delta \to 1$. Our definition of convergence does not support this view. The reason is simply that, as is well–known, there are sequences of action profiles for which the limiting average payoff does not exist and the liminf of the average payoff is not equal the limiting discounting payoff as $\delta \to 1$. On the other hand, if we define the limit of means game to be the limit as $\delta \to 1$ of $G(\varepsilon, \delta)$, then our results imply a discontinuity in δ for a fixed $\varepsilon > 0$. Specifically, for $\hat{\varepsilon} > 0$ small enough, we have

$$\lim_{\delta \uparrow 1} \mathcal{U}(\hat{\varepsilon}, \delta) \neq \mathcal{U}_\infty(\hat{\varepsilon}).$$

This holds simply because for $\hat{\varepsilon}$ sufficiently small, the left–hand side is close to

$$\lim_{\varepsilon \downarrow 0, \delta \uparrow 1}^{\infty} \ \mathcal{U}(\varepsilon, \delta) = F^* \ \cap \ W,$$

while the right–hand side is close to

$$\lim_{\varepsilon \downarrow 0} \mathcal{U}_\infty(\varepsilon) = c(U_{\geq}) \ \cap \ W,$$

and these sets are far apart in general. So if we define the limit as $\delta \to 1$ of the discounted game as the limit of means game, this result says that there is a discontinuity at $\delta = 1$. It is well–known that there are stage games for which $\lim_{\delta \uparrow 1} \mathcal{U}(0, \delta) \neq \mathcal{U}_\infty(0)$. However, the discontinuity for $\varepsilon > 0$ differs from the known discontinuity for $\varepsilon = 0$ in two ways. First, the known discontinuity is ruled out by fairly weak conditions such as the dimensionality condition we used in Theorem 1.B. The discontinuity for $\varepsilon > 0$ is not ruled out by such conditions. Second, the known discontinuity is a failure of lower semicontinuity, not upper. That is, the usual discontinuity occurs when $\mathcal{U}_\infty(0)$ is strictly larger than $\lim_{\delta \uparrow 1} \mathcal{U}(0, 1)$. Under very weak conditions, $\mathcal{U}_\infty(0)$ equals $F \cap R$ and it is *always* true that $\mathcal{U}(0, \delta) \subseteq F \cap R$. Hence we typically get upper semicontinuity in δ at $(\varepsilon, \delta) = (0, 1)$. By contrast, when the discontinuity in δ for fixed $\varepsilon > 0$ occurs, it is because $\mathcal{U}_\infty(\varepsilon)$ is close to a set which is strictly *smaller* than a set close to $\lim_{\delta \uparrow 1} \mathcal{U}(\varepsilon, \delta)$. Hence we generally violate upper semicontinuity.

Finally, we can use the results here to determine whether the results in Lipman–Wang [2000] indicate a discontinuity in the equilibrium outcome correspondence. To do so, we embed our previous model into this class of games in order to define convergence of that model. To be more precise, let $G_f(\varepsilon, \Delta)$ denote the infinitely repeated stage game where player i evaluates a sequence of actions $a^0, a^1, \ldots$ by the criterion

$$\sum_{t=0}^{T(\Delta)} \left[\Delta u_i(a^t) - \varepsilon \# \{ t \leq T(\Delta) \mid a_i^t \neq a_i^{t-1} \} \right]$$

where $T(\Delta)$ is the largest integer satisfying $(T+1)\Delta \leq \mathcal{L}$. (Recall that the length of time the game is played is equal to $\mathcal{L}$.) In other words, while the game is infinitely repeated,

only actions played in the first $T(\Delta)+1$ periods matter.[12] Given this, we can define convergence of a sequence of such games just as before. In particular, it is easy to show that

Lemma 3 *For any $\varepsilon > 0$ and any sequence Δ_n with $\Delta_n > 0$ and converging to 0,*

$$\lim_{n\to\infty} G_f(\varepsilon, \Delta_n) = G_\infty(\varepsilon).$$

On the other hand, analogously to Lemma 1, we have

Lemma 4 *Given a strictly positive sequence $(\varepsilon_n, \Delta_n)$,*

$$\lim_{n\to\infty} G_f(\varepsilon_n, \Delta_n) = \lim_{n\to\infty} G_f(0, \Delta_n)$$

if and only if $\varepsilon_n/\Delta_n \to 0$.

Proof. Fix any i and any sequence of action profiles where i changes actions every period. Since i's payoff in $G_f(0,\Delta_n)$ is bounded from below by $\min_{a\in A} u_i(a)$, the limit of his payoff as $n\to\infty$ must be finite. Hence if the payoff in $G_f(0,\Delta_n)$ has the same limit as $n\to\infty$ as the payoff in $G_f(\varepsilon_n,\Delta_n)$, the latter must also be finite, so the difference in payoffs must converge to 0 as $n\to\infty$. Note that the payoff to i in $G_f(0,\Delta_n)$ minus the payoff in $G_f(\varepsilon_n,\Delta_n)$ is $\varepsilon_n T(\Delta_n)$. But

$$\lim_{n\to\infty} \varepsilon_n T(\Delta_n) = \lim_{n\to\infty} \frac{\varepsilon_n}{\Delta_n} \Delta_n T(\Delta_n).$$

By definition of $T(\Delta_n)$ and the fact that $\Delta_n \to 0$, we must have $\lim_{n\to\infty} \Delta_n T(\Delta_n) = M$. Hence $\lim_{n\to\infty} \varepsilon_n T(\Delta_n) = 0$ iff $\lim_{n\to\infty} \varepsilon_n/\Delta_n = 0$. So if $G_f(\varepsilon_n,\Delta_n)$ and $G_f(0,\Delta_n)$ have the same limit, it must be true that $\varepsilon_n/\Delta_n \to 0$.

For the converse, fix any i and any sequence of actions $a^0, a^1, \ldots$, not necessarily one where i changes actions every period. Then the payoff in $G(0,\Delta_n)$ minus the payoff in $G(\varepsilon_n,\Delta_n)$ is

$$\varepsilon_n \#\{t \le T(\Delta) \mid a_i^t \neq a_i^{t-1}\} \le \varepsilon_n T(\Delta_n).$$

If $\varepsilon_n/\Delta_n \to 0$, then $\varepsilon_n T(\Delta_n) \to 0$, so the payoff difference goes to zero. Hence if $\varepsilon/\Delta_n \to 0$, $G_f(\varepsilon_n,\Delta_n)$ and $G_f(0,\Delta_n)$ do have the same limit. ∎

[12] This method of embedding a finitely repeated game into the infinitely repeated one is similar to that used by Fudenberg and Levine [1983]. They used a fixed action after some period, an approach less convenient for our purposes.

Just as with $G(\varepsilon, \delta)$, this result implies that there is no discontinuity in ε. As discussed in Lipman–Wang [2000], $G_f(\varepsilon, \Delta)$ yields different results from the usual finitely repeated game only when $\Delta < K\varepsilon$ for some $K > 0$. In particular, if ε/Δ is small, the switching costs do not change the usual results. In other words, when $\mathcal{U}_f(\varepsilon, \Delta)$ for small ε is significantly different from $\mathcal{U}_f(0, \Delta)$, it must be true that Δ is small relative to ε and so $G_f(\varepsilon, \Delta)$ is far from $G_f(0, \Delta)$. Hence there is no discontinuity in ε at $\varepsilon = 0$.

On the other hand, for any strictly positive ε, in general, $\mathcal{U}_f(\varepsilon, \Delta)$ is discontinuous in Δ at $\Delta = 0$. The easiest way to see the point is to return to the Prisoners' Dilemma game discussed earlier:

	C	D
C	$3, 3$	$0, 4$
D	$4, 0$	$2, 2$

As we noted, for any $\varepsilon > 0$ but small, it is impossible to sustain an equilibrium in $G_\infty(\varepsilon)$ which puts positive weight on $(0, 4)$ or $(4, 0)$. Hence the set of equilibrium payoffs is a subset of the (u_1, u_2) such that $2 \leq u_i \leq 3$ for both i. On the other hand, it is easy to use Theorem 1 of Lipman–Wang [2000] to construct equilibrium payoffs outside this set for $G_f(\varepsilon, \Delta)$ as $\Delta \downarrow 0$.[13]

This implies that for ε sufficiently small,

$$\lim_{\Delta \downarrow 0} \mathcal{U}_f(\varepsilon, \Delta) \neq \mathcal{U}_\infty(\varepsilon)$$

even though

$$\lim_{\Delta \downarrow 0} G_f(\varepsilon, \Delta) = G_\infty(\varepsilon)$$

by Lemma 3. Hence we have a discontinuity in Δ at $\Delta = 0$ for any sufficiently small $\varepsilon > 0$.

For the case of $\varepsilon = 0$, there is already a well–known discontinuity as $\Delta \downarrow 0$. Recall that $G_f(0, \Delta)$ is just the usual finitely repeated game where $T(\Delta)$ is the number of repetitions and $G_\infty(0)$ is the usual infinitely repeated game with the limit of means payoff criterion. So the well–known difference between finitely repeated and infinitely repeated games corresponds to a discontinuity in the equilibrium payoff correspondence at $\Delta = 0$.

There are two ways to see that the discontinuity in Δ for $\varepsilon > 0$ is fundamentally different from the known discontinuity for $\varepsilon = 0$. First, the known discontinuity is a

[13] More specifically, Theorem 1 of Lipman–Wang implies that for this game, there is a $K > 0$ such that for all sufficiently small ε and all $\Delta \in (0, K\varepsilon)$, there is a subgame perfect equilibrium in $G_f(\varepsilon, \Delta)$ where both players cooperate in every period. Given this, we can construct an equilibrium which begins with (C, D) played for some number of periods, followed by (C, C) for the rest of the game with play moving to (D, D) in the event of deviation. As long as the fraction of the time spent at (C, D) is small enough that player 1 gets a payoff of at least 2, this will be a subgame perfect equilibrium. In particular, then, the fraction of time spent at (C, D) does not have to converge to 0 as $\Delta \downarrow 0$.

failure of lower semicontinuity, not upper. That is, in the known discontinuity, the set of equilibrium payoffs at the limit is larger than the limiting set. Above, we showed that the limiting set of equilibrium payoffs contains points not in the set at the limit. Hence the discontinuity when $\varepsilon > 0$ shows a failure of upper semicontinuity, not lower.

Second, the discontinuity for $\varepsilon > 0$ occurs in some games where there is no discontinuity for $\varepsilon = 0$. For example, consider the coordination game:

	L	R
U	$4, 4$	$1, 0$
D	$0, 1$	$2, 2$

This game has no discontinuity in Δ at $\varepsilon = 0$ since it satisfies the Benoit–Krishna [1985] conditions for a finite repetition Folk Theorem. On the other hand, Theorem 4 of Lipman–Wang implies that for all sufficiently small ε,

$$\lim_{\Delta \downarrow 0} \mathcal{U}_f(\varepsilon, \Delta) = \{(4, 4)\},$$

while Theorem 6 above shows that $\mathcal{U}_\infty(\varepsilon)$ is close to the set of all (u_1, u_2) with $1 \leq u_i \leq 4$, $i = 1, 2$. Hence there is a discontinuity in Δ for ε small but positive. Note that this is a failure of lower semicontinuity, so the equilibrium outcome correspondence $\mathcal{U}_f(\varepsilon, \Delta)$ is neither upper nor lower semicontinuous in Δ at $\Delta = 0$ and $\varepsilon > 0$.

A Proof of Theorem 4

It is obvious that no $u \notin F^* \cap W$ can be an equilibrium payoff. Such a u is either infeasible or has some player with a lower payoff than what he could guarantee himself by a constant action. Hence we only need to show that all payoffs in $F^* \cap W$ can be close to equilibrium payoffs for k sufficiently large.

We begin by showing that every such payoff can be approximately generated by a cycle of actions.

Lemma 5 *Fix any $\eta > 0$. Then there exists $\bar{k}$ such that for all $u = (u_1, \ldots, u_I) \in F^* \cap W$ and all $k \geq \bar{k}$, there is a finite cycle of action profiles such that for any sequence $(\varepsilon_n, \delta_n) \to (0,1)$ with $\varepsilon_n/(1-\delta_n) \to k$, the payoff to i along this sequence converges as $n \to \infty$ to within η of u_i.*

Proof. Fix $\eta > 0$. For each i, let

$$d_i = \max_{a \in A} u_i(a) - w_i.$$

Let $\mathcal{A}$ denote the cardinality of A, let C be any integer satisfying

$$C \geq \max \left\{ \mathcal{A}^2, \ \frac{4(\mathcal{A}-1)\max_i \left[\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a)\right]}{\eta} \right\},$$

and let

$$B = 1 + \frac{4}{\eta} \sum_i d_i.$$

Let $\bar{k} = BC \sum_i d_i + 1$. Obviously, C, B, and therefore $\bar{k}$ converge to infinity as $\eta \downarrow 0$. Fix any $k \geq \bar{k}$.

Fix any $u \in F^* \cap W$. Since $u \in F^*$, there exists a probability distribution α over A and numbers $(x_1, \ldots, x_I) \geq (0, \ldots, 0)$ such that

$$u_i = \sum_{a \in A} \alpha(a) u_i(a) - x_i.$$

Also, $u \in W$, so

$$\max_{a \in A} u_i(a) - x_i \geq \sum_{a \in A} \alpha(a) u_i(a) - x_i \geq w_i.$$

By definition of d_i, then, $x_i \leq d_i$.

Obviously, we can approximately α arbitrarily closely by a probability distribution $\hat{\alpha}$ with the property that $\hat{\alpha}(a)$ is a strictly positive rational number for all $a \in A$. What we

show now is that this can be done in such a way that the C defined above is a common denominator for these probabilities, even though C is defined independently of u or α.

To see this, first, fix any a^* which maximizes $\alpha(a)$. Note that $\alpha(a^*) \geq 1/\mathcal{A}$. For every $a \neq a^*$, let c^a denote the unique integer such that

$$c^a - 1 \leq C\alpha(a) < c^a.$$

Note that this ensures that $c^a \geq 1$ for $a \neq a^*$. Let $c^{a^*} = C - \sum_{a \neq a^*} c^a$. To see that $c^{a^*} \geq 1$, note that $C\alpha(a) \geq c^a - 1$ for $a \neq a^*$ implies

$$C \sum_{a \neq a^*} \alpha(a) \geq \sum_{a \neq a^*} c^a - (\mathcal{A} - 1)$$

or

$$C[1 - \alpha(a^*)] \geq \sum_{a \neq a^*} c^a - \mathcal{A} + 1.$$

Hence

$$c^{a^*} = C - \sum_{a \neq a^*} c^a \geq C\alpha(a^*) - \mathcal{A} + 1.$$

So $c^{a^*} \geq 1$ if $C\alpha(a^*) - \mathcal{A} + 1 \geq 1$ or $C \geq \mathcal{A}/\alpha(a^*)$. Since $\alpha(a^*) \geq 1/\mathcal{A}$, we have $\mathcal{A}^2 \geq \mathcal{A}/\alpha(a^*)$. The choice of C ensures that $C \geq \mathcal{A}^2$, so $C \geq \mathcal{A}/\alpha(a^*)$ as required.

Consider

$$\left| \sum_{a \in A} \left(\alpha(a) - \frac{c^a}{C} \right) u_i(a) \right|.$$

Given the specification of the c^a's, the weight on each a in $\sum_a (c^a/C) u_i(a)$ differs from the weight in $\sum_a \alpha(a) u_i(a)$ by less than $1/C$. The worst possible effect this could have is if $\mathcal{A} - 1$ points give i payoff $\max_{a \in A} u_i(a)$, the remaining action vector gives $\min_{a \in A} u_i(a)$, and we shift $1/C$ from each of the first $\mathcal{A} - 1$ action vectors to the last one. In this case, we lower i's payoff by $(\mathcal{A} - 1)/C$ times $\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a)$. Hence

$$\max_i \left| \sum_{a \in A} \left(\alpha(a) - \frac{c^a}{C} \right) u_i(a) \right| \leq (\mathcal{A} - 1) \frac{1}{C} \max_i \left[\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a) \right].$$

By assumption,

$$C \geq \frac{4(\mathcal{A} - 1) \max_i \left[\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a) \right]}{\eta}.$$

Hence

$$\max_i \left| \sum_{a \in A} \left(\alpha(a) - \frac{c^a}{C} \right) u_i(a) \right| \leq (\mathcal{A} - 1) \frac{1}{C} \max_i \left[\max_{a \in A} u_i(a) - \min_{a \in A} u_i(a) \right] \leq \frac{\eta}{4}.$$

For the next step, define $X = \sum_i x_i$. Define L to be the unique integer such that

$$L \leq \frac{k}{CX} < L + 1.$$

Because $B \geq 1$, the fact that $k > BC \sum_i d_i$ implies $k > C \sum_i d_i$. But $d_i \geq x_i$, so this implies $k > CX$. Hence $L \geq 1$.

Let N be any integer satisfying

$$N > \frac{4}{\eta} \max \left\{ \left[\frac{k(I-1) \sum_i d_i}{k - C \sum_i d_i} \right], \frac{k\mathcal{A}}{LC} \right\}.$$

For $i \neq I$, define N_i to be the unique integer such that

$$N_i \leq x_i \frac{LCN}{k} < N_i + 1.$$

Clearly, $N_i \geq 0$ for all i. Define $N_I = N - \sum_{i \neq I} N_i$. To see that $N_I \geq 0$, note that the definition of N_i implies

$$\sum_{i \neq I} N_i \leq \frac{LCN}{k} \sum_{i \neq I} x_i,$$

so

$$N_I = N - \sum_{i \neq I} N_i \geq N \left[1 - \frac{LC}{k} \sum_{i \neq I} x_i \right] = N \left[1 - \frac{LC}{k} (X - x_I) \right].$$

Hence $N_I \geq 0$ if

$$X - x_I \leq \frac{k}{LC}$$

or $X - (k/LC) \leq x_I$. From the definition of L, however, $X \leq k/LC$, so $x_I \geq 0$ implies that this holds.

For $i \neq I$, the definition of N_i implies

$$\frac{N_i k}{LCN} \leq x_i < \frac{(N_i + 1)k}{LCN}. \tag{4}$$

Hence

$$\left| x_i - \frac{N_i k}{LCN} \right| \leq \frac{k}{LCN}.$$

For $i = I$, we have a more complicated bound. Summing equation (4) over $i \neq I$ yields

$$\frac{(N - N_I)k}{LCN} \leq X - x_I \leq \frac{(N - N_I + I - 1)k}{LCN}.$$

Hence

$$X - \frac{k}{LC} - \frac{k(I-1)}{LCN} \leq x_I - \frac{kN_I}{LCN} \leq X - \frac{k}{LC}.$$

From the definition of L,

$$\frac{k}{(L+1)C} < X \leq \frac{k}{LC}.$$

Hence

$$\frac{k}{(L+1)C} - \frac{k}{LC} < X - \frac{k}{LC} \leq 0.$$

So

$$\frac{k}{(L+1)C} - \frac{k}{LC} - \frac{k(I-1)}{LCN} \leq x_I - \frac{kN_I}{LCN} \leq 0.$$

Hence

$$\left| x_I - \frac{kN_I}{LCN} \right| \leq \frac{k}{LC} \left(\frac{1}{L+1} + \frac{I-1}{N} \right).$$

Since $I \geq 2$, the bound we have for our error on I must exceed the bound for any $i \neq I$. That is,

$$\left| x_i - \frac{kN_i}{LCN} \right| \leq \frac{k}{C} \left(\frac{1}{L(L+1)} + \frac{I-1}{LN} \right), \quad \forall i.$$

Note that the right–hand side is strictly decreasing in L. By definition, $L > (k/CX) - 1$. Hence for all i

$$\left| x_i - \frac{kN_i}{LCN} \right| < \frac{k}{C} \left(\frac{1}{[(k/CX)-1](k/CX)} + \frac{I-1}{[(k/CX)-1]N} \right).$$

The right–hand side is

$$= \left(\frac{X^2 C}{k - CX} + \frac{kX(I-1)}{[k-CX]N} \right). \tag{5}$$

The first term in (5) is strictly less than $\eta/4$ if

$$k > CX + X^2 C \frac{4}{\eta} = \left[1 + X \frac{4}{\eta} \right] CX. \tag{6}$$

Recall that $X \leq \sum_i d_i$. Using this and the definition of B, we have

$$BC \sum_i d_i = \left[1 + \frac{4}{\eta} \sum_i d_i \right] C \sum_i d_i \geq \left[1 + X \frac{4}{\eta} \right] CX.$$

By construction, $k > BC \sum_i d_i$, so (6) holds. Hence

$$\frac{X^2 C}{k - CX} < \frac{\eta}{4}.$$

Consider the second term in (5). Note that it is strictly increasing in X and recall that $X \leq \sum_i d_i$. Hence

$$\frac{kX(I-1)}{[k-CX]N} \leq \frac{1}{N} \left(\frac{k(I-1)\sum_i d_i}{k - C\sum_i d_i} \right).$$

The right-hand side is strictly less than $\eta/4$ if

$$N > \left[\frac{k(I-1)\sum_i d_i}{k - C\sum_i d_i}\right]\frac{4}{\eta}.$$

By our choice of N, this holds. Hence

$$\left|x_i - k\frac{N_i}{LCN}\right| < \frac{\eta}{2}.$$

Finally, note that

$$N > \frac{4}{\eta}\frac{k\mathcal{A}}{LC} \text{ implies } \frac{k\mathcal{A}}{LCN} < \frac{\eta}{4}.$$

We are now ready to construct the cycle of actions. After doing so, we use the facts established above to characterize the limiting payoffs along the cycle.

First, we have a *switching phase*. We begin at an arbitrary action vector a^0. Then player 1 changes actions back and forth N_1 times. Let a^1 denote the vector created when he changes actions. After this, depending on whether N_1 is even or odd, we are at either a^0 or a^1. At this point, player 2 changes actions N_2 times. Let a^2 denote the vector created when 2 changes. Continue similarly to construct action profiles $a^3, \ldots, a^I$. After player I carries out his N_I switches, we move on to the *payoff phase*. Fix any order of action profiles in A. In the payoff phase, the action profiles are played in this fixed order. When it is profile a's turn to be played, it is played $c^a LN$ times minus the number of times it was played (if at all) in the switching phase. Note that a given action profile must be played strictly fewer than N times in the switching phase (since this is the length of the phase). Hence $c^a LN$ minus the number of times a profile is played in the switching phase must be strictly positive since $c^a \geq 1$ for every $a \in A$ and $L \geq 1$. After the last action has its turn in the payoff phase, the cycle starts over.

Note that the total length of the cycle is $\sum_a c^a LN = CLN$. Profile a is played exactly $c^a LN$ times over the course of the cycle, so the frequency with which it is played is c^a/C.

How many times does player i change actions over the course of the cycle? By construction, he changes exactly N_i times in the switching phase. We cannot say how many times he changes in the payoff phase, but we know that it cannot be more times that the number of action profiles. Call Z_i the number of times i changes in the switching phase and recall that $\mathcal{A}$ is the number of action profiles. So $Z_i \leq \mathcal{A}$.

Fix any sequence $(\varepsilon_n, \delta_n)$ converging to $(0,1)$ with $\varepsilon_n/(1-\delta_n)$ converging to k. It is not hard to see that i's payoff to the sequence of actions constructed above converges as $n \to \infty$ to

$$\sum_{a \in A} \frac{c^a}{C} u_i(a) - k\frac{N_i + Z_i}{LCN}.$$

Note that

$$\max_i \left| u_i - \left(\sum_{a \in A} \frac{c^a}{C} u_i(a) - k \frac{N_i + Z_i}{LCN} \right) \right| = \max_i \left| \sum_{a \in A} \left(\alpha(a) - \frac{c^a}{C} \right) u_i(a) + x_i - k \frac{N_i + Z_i}{LCN} \right|.$$

But the right–hand side is

$$\leq \max_i \left| \sum_{a \in A} \left(\alpha(a) - \frac{c^a}{C} \right) u_i(a) \right| + \max_i \left| x_i - k \frac{N_i}{LCN} \right| + \frac{k\mathcal{A}}{LCN} < \frac{\eta}{4} + \frac{\eta}{2} + \frac{\eta}{4} = \eta,$$

completing the proof. ∎

Fix any $\eta > 0$. Let k_η be the larger of the $\bar{k}$ defined in the proof of Lemma 5 and

$$2 \max_i [\max_{a \in A} u_i(a) - w_i].$$

As noted in the proof of Lemma 5, $\bar{k} \to \infty$ as $\eta \downarrow 0$, so obviously $k_\eta \to \infty$ as well.

Fix any $k \geq k_\eta$ and any $\bar{u} \in F^* \cap W$. We will show that there is a u within η of $\bar{u}$ such that $u \in \lim^k \mathcal{U}(\varepsilon, \delta)$.

First, we show that without loss of generality, we can assume that $\bar{u}_i > w_i$ for all i. To see this, recall that the game is weakly regular. By definition, then, there is a $u' \in F$ (and hence in F^*) such that $u'_i > w_i$ for all i. Let $\bar{u}' = \lambda \bar{u} + (1 - \lambda) u'$ for $\lambda \in (0, 1)$. Since F^* is convex, this must be contained in F^*. Clearly, by making λ large enough, we can make $\bar{u}'$ arbitrarily close to $\bar{u}$. Also, since $\bar{u}_i \geq w_i$ and $u'_i > w_i$, we have $\bar{u}'_i > w_i$ for all i. Hence if $\bar{u}_i = w_i$ for some i, we can replace $\bar{u}$ with $\bar{u}'$.

Fix any sequence $(\varepsilon_n, \delta_n)$ converging to $(0, 1)$ with $\varepsilon_n / (1 - \delta_n) \to k$. By Lemma 5, there is a finite cycle of actions, independent of n, which generates a payoff vector, say u^n, such that $\lim_{n \to \infty} u^n$ is within η of $\bar{u}$. Let u denote this limit. By taking $\eta < \min_i \bar{u}_i - w_i$, we can ensure that $u_i > w_i$ for all i.

An important observation which we use repeatedly below is that i's continuation payoff at any point in this sequence of actions converges as $n \to \infty$ to u_i. That is, suppose that we consider i's payoff from some point within the cycle onward. Suppose that ℓ periods remain in the current cycle. Because the cycle is finite, ℓ is bounded. Then i's continuation payoff from this point forward must be at least

$$\left[\sum_{t=0}^{\ell - 1} \delta_n^t \right] \left[(1 - \delta_n) \min_{a \in A} u_i(a) - \varepsilon_n \right] + \delta_n^\ell [u_i(n) - \varepsilon_n]$$

and must be less than

$$\left[\sum_{t=0}^{\ell - 1} \delta_n^t \right] (1 - \delta_n) \max_{a \in A} u_i(a) + \delta_n^\ell u_i(n).$$

Both of these expressions converge as $n \to \infty$ to u_i.

We now show that every sufficiently large n, there is a subgame perfect equilibrium with an equilibrium path equal to this cycle of actions. The subgame perfect equilibrium strategies are independent of n. Fix any payoff vectors $\hat{u}^1, \ldots, \hat{u}^I \in F^* \cap W$ such that

$$\hat{u}_i^j > w_i, \ \forall i, j$$

$$\hat{u}_i^i < u_i, \ \forall i$$

$$\hat{u}_i^i < \hat{u}_i^j, \ \forall i, j.$$

It is not hard to see that such $\hat{u}^j$'s must exist since F^* is full dimensional. For example, we could set $\hat{u}^j$ equal to u minus a small amount for player j. For each $\hat{u}^j$, there is a finite cycle such that the payoffs along the sequence converges as $n \to \infty$ to within η of $\hat{u}^j$. Without loss of generality, we can assume, in fact, that for each $\hat{u}^j$, there is a finite cycle such that payoffs along the sequence converge as $n \to \infty$ to exactly $\hat{u}^j$. If not, we can simply replace $\hat{u}^j$ with the limiting payoff along this sequence. For η small enough, these limiting payoffs will satisfy the required properties.[14]

The strategies are similar to those used by Fudenberg and Tirole [1991] in their proof of their Theorem 5.4. Play begins in phase I. In phase I, the players follow the cycle of actions generating the payoff near u. As long as no player deviates, all follow these actions. If there is a unilateral deviation by player i in phase I, we move to phase II_i. In this phase, i plays any one of his maxmin actions for M_i periods (M_i characterized below). In each period of phase II_i, all players other than i play some vector of actions which minimizes i's payoff against his action in the preceding period. (Given that i follows his equilibrium strategy in this phase, these actions which be those which minimize i's payoff from his chosen maxmin action from the second period of phase II_i onward.) At the end of these M_i periods, we move to phase III_i. In this phase, the players follow the cycle of actions which generates payoffs $\hat{u}^i$. If player j, $j \neq i$, unilaterally deviates in phase II_i or III_i, we move to phase II_j and then III_j. If player i unilaterally deviates in phase II_i, we restart the phase. We treat unilateral deviations by player i in phase III_i slightly differently. In this case, we move to phase II_i as above, but afterward move back to where we left off in phase III_i rather than beginning that phase again. That is, if in the kth period of the cycle generating payoff $\hat{u}^i$, player i deviates, then when we have completed phase II_i, we return to the action profile that was supposed to be played in the kth period of the cycle and continue from there. We assign an arbitrary subgame perfect continuation if there are multiple simultaneous deviations in any phase.

[14]To be more precise, suppose we choose the $\hat{u}^j$'s so that the payoff differences stated above at all at least γ where $\gamma > 2\eta$. Given this, we can use the limiting payoffs and know that they satisfy the inequalities stated above.

M_i is set to any integer M such that

$$\hat{u}_i^i > \frac{2}{M} \max_{a \in A} u_i(a) + \frac{M-2}{M} w_i. \tag{7}$$

Obviously, the fact that $\hat{u}_i^i > w_i$ implies that such an M exists.

To show that these strategies form a subgame perfect equilibrium, we show that there is a finite set of inequalities, each of which holds for n sufficiently large, which ensure that no player has a profitable deviation on any history. Since there are only finitely many inequalities involved, we know that there is a finite $\bar{n}$ such that for all larger n, all inequalities hold. Hence for $n \geq \bar{n}$, these strategies form a subgame perfect equilibrium.

So consider any history such that the players are in phase I. Suppose there are ℓ periods left in the current cycle. As argued above, player i's continuation payoff from following the equilibrium strategy converges as $n \to \infty$ to u_i. If i deviates, his payoff in the period of deviation certainly cannot be larger than $\max_{a \in A} u_i(a)$. Given that i follows the equilibrium strategies after the period of deviation, his payoff from deviating cannot be larger than

$$(1-\delta_n)\left[\max_{a \in A} u_i(a)(1+\delta_n) + \sum_{t=2}^{M_i} \delta_n^t w_i\right] + \delta_n^{M_i+1} \hat{u}_i^i.$$

To understand this calculation, recall that in the first period of the punishment, i will play one of his maxmin actions, but the other players will be choosing actions which maximize his payoff against his *previous* action, not his current action. In general, therefore, the payoff he earns in this period will not be w_i. Obviously, though, it cannot exceed $\max_{a \in A} u_i(a)$. Also, this calculation presumes that i will not incur any switching costs during the punishment and will not have to change actions to begin phase III_i. Clearly, then, this is an upper bound on i's payoff to deviating. As $n \to \infty$, this payoff converges to $\hat{u}_i^i$. By assumption $u_i > \hat{u}_i^i$, so for all n sufficiently large, i does not gain by deviating in phase I. Note that this argument implicitly considers finitely many inequalities since we certainly do not need more than one inequality for each player, each of his finitely many possible deviation actions, and each of the finitely many periods of the cycle in which he might deviate.

So consider any history such that the players are in phase II_i. Recall that where we go from here depends on how this phase was reached. In particular, if we were in the middle of a cycle in phase III_i and player i deviated, then after completing this phase, we pick up from where we left off the cycle. Let $\hat{u}_j(n)$ denote j's continuation payoff from the end of this phase onward. If we entered phase II_i in any fashion other than a deviation from i in the middle of III_i, then this is $\hat{u}_j^i$; otherwise, it a little more complex but the distinction will not be important for this part of the proof. The only important fact to note is that as $n \to \infty$, $\hat{u}_j(n) \to \hat{u}_j^i$ for all j.

First consider player i. Suppose we are at the beginning of the phase II_i and that i must change actions to reach a maxmin action. Then if he follows the equilibrium strategies from this point, his payoff will be at least

$$-\varepsilon_n + (1-\delta_n)\sum_{t=0}^{M_i-1}\delta_n^t w_i + \delta_n^{M_i}\hat{u}_i(n). \tag{8}$$

To understand this, recall that in the first period of punishment, he will be playing one of his maxmin actions, but the other players will be choosing actions which might not minimize his payoff given that action. Hence his payoff in that period must be at least w_i. Of course, in the remainder of the phase, his payoff will be exactly w_i each period.

For the deviation payoff, consider two cases. First, suppose the deviation involves not changing actions in the first period. Recall that the other players are choosing actions which minimize i's payoff to his previous period's action. Hence the payoff i earns in the period of deviation must be less than w_i. After this, following the equilibrium strategies must earn exactly the same payoff as he would have gotten had he followed them from the outset. That is, if e_i is i's payoff to following the equilibrium, then his payoff to deviating by not changing actions and then following the equilibrium thereafter is no more than $w_i(1-\delta_n)+\delta_n e_i$. Clearly, i is better off following the equilibrium if $e_i \geq w_i(1=\delta_n)+\delta_n e_i$ or $e_i \geq w_i$. The lower bound above for i's continuation payoff from following the equilibrium converges to $\hat{u}_i^i$ as $n \to \infty$. Since $\hat{u}_i^i > w_i$, we know that $e_i > w_i$ for all n sufficiently large.

Next, suppose the deviation involves changing actions but to some action other than one of the maxmin actions. In this case, the payoff to deviating and then following the equilibrium strategies cannot exceed

$$-\varepsilon_n(1+\delta_n) + (1-\delta_n)(1+\delta_n)\max_{a\in A} u_i(a) + (1-\delta_n)\sum_{t=2}^{M_i}\delta_n^t w_i + \delta_n^{M_i+1}\hat{u}_i(n). \tag{9}$$

Hence the deviation is not profitable if the expression in (8) exceeds the above or

$$\delta_n\varepsilon_n+(1-\delta_n)(1+\delta_n)w_i+\delta_n^{M_i}\hat{u}_i(n) > (1-\delta_n)(1+\delta_n)\max_{a\in A} u_i(a)+(1-\delta_n)\delta_n^{M_i}w_i+\delta_n^{M_i+1}\hat{u}_i(n)$$

or

$$\delta_n\frac{\varepsilon_n}{1-\delta_n} + (1+\delta_n)w_i + \delta_n^{M_i}\hat{u}_i(n) > (1+\delta_n)\max_{a\in A} u_i(a) + \delta_n^{M_i}w_i.$$

Note that this holds in the limit as $n \to \infty$ if

$$k + 2w_i + \hat{u}_i^i > 2\max_{a\in A} u_i(a) + w_i.$$

Because $\hat{u}_i^i > w_i$, a sufficient condition is $k > 2[\max_{a\in A} u_i(a) - w_i]$ which holds by our assumption on k. Hence for n sufficiently large, again, the deviation is not profitable.

Now suppose that that there are ℓ periods left to the punishment phase, $\ell \leq M_i$, and that i played a maxmin action in the preceeding period. This would hold if we are in the midst of the punishment phase or if a deviation by i to this action is what initiated phase II_i. In this case, following the equilibrium strategies gives i a payoff of at least

$$(1-\delta_n)\sum_{t=0}^{\ell-1}\delta_n^t w_i + \delta_n^\ell \hat{u}_i(n).$$

Because $\hat{u}_i^i > w_i$, this is strictly decreasing in ℓ for n sufficiently large. Hence this is at least

$$(1-\delta_n)\sum_{t=0}^{M_i-1}\delta_n^t w_i + \delta_n^{M_i} \hat{u}_i(n).$$

Note that this is exactly ε_n larger than the expression in equation (8). If i deviated and then returned to the equilibrium strategies, his payoff would be no greater than the expression in equation (9). Since we already showed that the expression in (8) exceeds the expression in (9) for all sufficiently large n, obviously something which is ε_n larger will exceed it as well. Hence for all n sufficiently large, i has no profitable deviation on such a history. This concludes the analysis of player i in phase II_i.

So consider any player $j \neq i$ in phase II_i. Suppose that this player must switch actions in the current period and again in the next. If all players follow the equilibrium strategies, then this is the largest number of times j will have to switch actions during this phase, so this is the worst case scenario for j. In this case, following the equilibrium path gives j a payoff no worse than

$$-\varepsilon_n(1+\delta_n) + (1-\delta_n)\sum_{t=0}^{M_i-1}\delta_n^t \min_{a\in A} u_j(a) + \delta_n^{M_i}\hat{u}_j(n).$$

As $n \to \infty$, this converges to $\hat{u}_j^i$. If instead j were to deviate, his payoff would be no larger the bigger of

$$(1-\delta_n)\max_{a\in A} u_j(a) + (1-\delta_n)\sum_{t=1}^{M_j}\delta_n^t w_j + \delta_n^{M_j+1}\hat{u}_j^j$$

and

$$(1-\delta_n)(1+\delta_n)\max_{a\in A} u_j(a) - \delta_n\varepsilon_n + (1-\delta_n)\sum_{t=2}^{M_j}\delta_n^t w_j + \delta_n^{M_j+1}\hat{u}_j^j.$$

To see this, note that there are two possibilities after the deviation: either j's deviation is one of his maxmin actions or it is not. If it is one of his maxmin actions, then he will not have to change actions again, but will begin earning w_j in the next period. The first expression gives an upper bound on his payoff in this case. If the deviation action is not one of his maxmin actions, he will not get w_j until two periods later, but will have to

change actions in the following period. As $n \to \infty$, each expression converges to $\hat{u}_j^j$. By assumption $\hat{u}_j^i > \hat{u}_j^j$, so j does not wish to deviate for all n sufficiently large.

Note that our argument for all the phase II_i's together implicitly involves finitely many inequalities. To see this, note that we certainly do not need more inequalities than one for each player who is being punished, each phase in which the deviation leading to punishment occurred, each of the finitely many possible periods of the cycle at which it occurred, each of the finitely many actions the deviator may have chosen, each player who might now deviate, each of the finitely many periods of punishment in which he might deviate, and each of the finitely many deviation actions he might choose.

Now consider any history in phase III_i. Suppose there are k periods left in the cycle. Consider any player $j \neq i$. From the same reasoning as above, we know that as $n \to \infty$, j's payoff to following the equilibrium must converge to $\hat{u}_j^i$. The upper bound on j's payoff constructed in the analysis of deviations by j in phase II_i apply here as well, so we know that his deviation payoff converges as $n \to \infty$ to no more than $\hat{u}_j^j$. Because $\hat{u}_j^i > \hat{u}_j^j$, j does not gain by deviating for any sufficiently large n.

Finally, consider player i. Let $\hat{u}_i(n)$ denote i's continuation payoff to following the equilibrium. If i deviates, his payoff is less than or equal to

$$(1-\delta_n)(1+\delta_n)\max_{a\in A} u_i(a) + (1-\delta_n)\sum_{t=2}^{M_i}\delta_n^t w_i + \delta_n^{M_i+1}\hat{u}_i(n).$$

To see this, recall that a deviation by i within phase III_i leads to a move to phase II_i followed by a return to the previous point in the cycle. Note that the only way that i might have to switch in the current period on the equilibrium path but avoid the switch in returning to this point after phase II_i is if he is currently supposed to change to a maxmin action but does not. In this case, though, i will have to change actions after the deviation, a cost which is omitted from the above expression and is earlier (and so discounted less) than the one mistakenly attributed to him in $\hat{u}_i(n)$. Hence this expression is a valid lower bound even in that circumstance. Hence i does not have a profitable deviation if

$$\hat{u}_i(n)(1-\delta_n^{M_i+1}) > (1-\delta_n)(1+\delta_n)\max_{a\in A} u_i(a) + (1-\delta_n)\sum_{t=2}^{M_i}\delta_n^t w_i$$

or

$$\hat{u}_i(n) > \frac{1+\delta_n}{\sum_{t=0}^{M_i}\delta_n^t}\max_{a\in A} u_i(a) + \frac{\sum_{t=2}^{M_i}\delta_n^t}{\sum_{t=0}^{M_i}\delta_n^t} w_i.$$

This holds at the limit as $n \to \infty$ iff

$$\hat{u}_i^i > \frac{2}{M_i}\max_{a\in A} u_i(a) + \frac{M_i-2}{M_i} w_i,$$

which is required by our definition of M_i. Hence for all n sufficiently large, i does not have a profitable deviation.

Note that our argument for phase III_i implicitly involves finitely many inequalities, since we certainly do not need more than one for each of the phases, each player who might deviate, each of the finitely many periods in the cycle in which he might deviate, and each possible deviation action.

Summarizing, then, we have a finite set of inequalities. For each, there is an $\bar{n}$ such that the inequality holds for all $n \geq \bar{n}$. Letting $\hat{n}$ denote the largest of these finitely many $\bar{n}$'s, we see that all the inequalities hold for $n \geq \hat{n}$. Hence for $n \geq \hat{n}$, these strategies form a subgame perfect equilibrium. ▌

B Proof of Theorem 3

This result is a corollary to Theorem 4. To see this, fix any payoff $u \in F^* \cap W$. Fix a sequence of strictly positive numbers η_ℓ converging to 0. For each η_ℓ, Theorem 4 establishes that there is a k_ℓ and a $\bar{u}(\ell)$ within η_ℓ of u such that

$$\bar{u}(\ell) \in \lim_{\varepsilon \downarrow 0, \delta \uparrow 1}^{k_\ell} \mathcal{U}(\varepsilon, \delta).$$

In other words, for each ℓ, there is a sequence $(\varepsilon_n^\ell, \delta_n^\ell)$ converging to $(0,1)$ with $\varepsilon_n/(1-\delta_n) \to \bar{k}_\ell$ and a sequence $u^n(\ell)$ such that $u^n(\ell) \in \mathcal{U}(\varepsilon_n^\ell, \delta_n^\ell)$ and $u^n(\ell)$ converges to $\bar{u}(\ell)$ within η of u.

So define a new sequence $(\hat{\varepsilon}_n, \hat{\delta}_n)$ such that $(\hat{\varepsilon}_n, \hat{\delta}_n) = (\varepsilon_n^n, \delta_n^n)$. Let $\hat{u}_n = u_n(n)$. Then $(\hat{\varepsilon}_n, \hat{\delta}_n)$ converges to $(0,1)$, $\hat{\varepsilon}_n/(1-\hat{\delta}_n)$ converges to infinity, and $\hat{u}_n$ converges to u. Hence

$$u \in \lim_{\varepsilon \downarrow 0, \delta \uparrow 1}^{\infty} \mathcal{U}(\varepsilon, \delta). \blacksquare$$

C Proof of Theorem 5

The key step in the proof is

Lemma 6 *Fix any $\eta > 0$ and any k. Then for all $u = (u_1, \ldots, u_I) \in F \cap W$, there is a finite cycle of action profiles such that for any sequence $(\varepsilon_n, \delta_n) \to (0,1)$ with $\varepsilon_n/(1-\delta_n) \to k$, the payoff to i along this sequence converges as $n \to \infty$ to within η of u_i.*

Proof. The key difference between this lemma and Lemma 5 is that we require $u \in F \cap W$ instead of $F^* \cap W$ and allow for arbitrary k. To see that this is possible, fix any η and any k. Fix any $u \in F \cap W$. By definition, $u \in F$ implies that there exists a probability distribution α over A such that

$$u_i = \sum_{a \in A} \alpha(a) u_i(a).$$

It is easy to modify the argument in Lemma 5 to show that we can find integers $c^a \geq 0$ for all $a \in A$ such that

$$\max_i \left| \sum_{a \in A} \left(\alpha(a) - \frac{c^a}{C} \right) u_i(a) \right| < \frac{\eta}{2}$$

where $C = \sum_{a \in A} c^a$. Choose any integer L such that

$$L > \frac{2k\mathcal{A}}{\eta C},$$

where, as before, $\mathcal{A}$ is the cardinality of A. In other words,

$$k \frac{\mathcal{A}}{LC} < \frac{\eta}{2}.$$

Construct a cycle as follows. Fix any order over the action profiles a which have $c^a > 0$. The players then play the action profiles in this sequence where profile a is played for $c^a L$ periods. After this, the cycle starts over. The length of the cycle, then, is $\sum_a c^a L = CL$. Hence the relative frequency with which a is played is $c^a L / CL = c^a / C$. We cannot say exactly how many times player i changes actions over the course of the cycle. However, it is certainly fewer than $\mathcal{A}$ times since this is the number of action profiles. Let Z_i be the number of times i changes actions. Obviously, i's payoff over this infinite sequence of actions converges as $n \to \infty$ to

$$\sum_{a \in A} \frac{c^a}{C} u_i(a) - k \frac{Z_i}{CL}.$$

So

$$\max_i \left| u_i - \left[\sum_{a \in A} \frac{c^a}{C} u_i(a) - k \frac{Z_i}{CL} \right] \right| \leq \max_i \left| \sum_{a \in A} \left(\alpha(a) - \frac{c^a}{C} \right) u_i(a) \right| + k \frac{\mathcal{A}}{LC} < \frac{\eta}{2} + \frac{\eta}{2} = \eta. \blacksquare$$

To complete the proof, simply construct strategies exactly as in the proof of Theorem 4. The only condition on k used in that part of the proof of Theorem 4 was $k > 2 \max_i [\max_{a \in A} u_i(a) - w_i]$, which holds by assumption. $\blacksquare$

D Proof of Theorem 6

We first show that $\hat{u} \in c(U_{\geq}) \cap W \subseteq \lim_{\varepsilon \downarrow 0} \mathcal{U}_{\infty}(\varepsilon)$. So fix any $\hat{u} \in c(U_{\geq}) \cap W$. By definition, $\hat{u} \geq w$. Also, there are action profiles, say $a^1, \ldots, a^Z$, and strictly positive numbers $\alpha_1, \ldots, \alpha_Z$ such that $u(a^z) \geq w$ for all z, $\sum_{z=1}^{Z} \alpha_z = 1$, and $\hat{u} \leq \sum_z \alpha_z u(a^z)$. Without loss of generality, assume $Z = 1$. Once we prove the result for this case, the fact that we allow public randomizations extends the result to larger Z.

Fix an $\varepsilon > 0$, smaller than any possible payoff difference. That is, choose ε so that

$$\varepsilon < \min_i \min_{u,u' \in u(A), u_i \neq u'_i} |u_i - u'_i|. \tag{10}$$

Rewardability implies that there is no player whose payoff is constant over all $u \in u(A)$. Hence the right–hand side is strictly positive, so this is possible.

For each i, let c_i denote the largest nonnegative integer such that $u_i(a^1) - c_i\varepsilon \geq \hat{u}_i$. Since $u_i(a^1) \geq \hat{u}_i$, c_i is well–defined. Let u'_i denote $u_i(a^1) - c_i\varepsilon$ evaluated at this largest c_i. Note that c_i and thus u' are functions of ε, though we omit this dependence in the notation. Clearly, as $\varepsilon \downarrow 0$, $u' \to \hat{u}$. In light of this, we show that $u' \in \mathcal{U}(\varepsilon, 1)$ for all sufficiently small ε, thus demonstrating $\hat{u} \in \lim_{\varepsilon \downarrow 0} \mathcal{U}(\varepsilon, 1)$.

To show this, construct strategies as follows. In the first period, if c_i is even (where 0 is treated as even), player i plays a_i^1. Otherwise, he plays any other action. For the next several periods, each player changes between a_i^1 and any other action, concluding when he has changed actions c_i times. At this point, by construction, he will be back to a_i^1. Once all players have completed this phase, no player changes actions again, so a^1 is played forever after. It is easy to see that the payoffs if there are no deviations are u'.

To complete the specification of the strategies, we need to specify what happens in response to a deviation. Let $\bar{u}$ be the equally weighted average of the payoff vectors in $U_{\geq}$ — that is,

$$\bar{u} = \frac{1}{\#U_{\geq}} \sum_{u \in U_{\geq}} u.$$

Our rewardability assumption implies that $\bar{u} \gg w$.

For simplicity, we describe behavior at the out of equilibrium histories in terms of a number of different punishment modes. There is one punishment mode for each player and each action available to that player. So we refer to a typical punishment mode as the (i, a_i) punishment mode where $a_i \in A_i$. In punishment mode (i, a_i), i is the target of the punishment and a_i is the action he played which started this punishment mode.

We go to punishment mode (i, a_i) if i is the first player to deviate from the equilibrium

play above and deviates by playing a_i when he is supposed to play something different. (We ignore multiple simultaneous deviations throughout. Any specification of a subgame equilibrium will suffice for these histories.) If some player j (which may equal i) deviates while we are in punishment mode for i by playing action a_j when he is not supposed to, we move to punishment mode (j, a_j).

In the first period of punishment mode (i, a_i), all players other than i (the target) go to the actions which minimize the target's payoff under the hypothesis that the target plays the same action he played in the previous period. There is a number κ_i of times that the target is supposed to change actions, independent of a_i. As long as the target continues to change actions, the other players do not change their actions. If the target stops changing before he has changed κ_i times and the last action played is a_i', we move to punishment mode (i, a_i'). In particular, then, the other players move to the actions which minimize the target's payoff from a_i' and the changes of action must begin again. The exact sequence of actions used by the target while changing actions is unimportant with two exceptions. First, the target's strategy is to change actions κ_i times without stopping. Second, κ_i will be even and the sequence must have the property that the target concludes the sequence by returning to a_i. Any action which does not deviate from these requirements does not count as a deviation.

Punishment mode (i, a_i) is completed once the target has changed actions κ_i times. The continuation is then determined by the outcome of a publicly observed randomization to pick a vector from $U_{\geq}$. For any player i, let $\underline{u}^i$ denote any $u \in U_{\geq}$ which minimizes u_i subject over $U_{\geq}$. The public randomization after punishment mode (i, a_i) puts probability $q_i(a_i)$ on $\underline{u}^i$ and with probability $1 - q_i(a_i)$ chooses uniformly from (all of) $U_{\geq}$. When the outcome of this randomization is observed, all agents change actions (if need be) to move to any action profile generating the selected payoff vector and never change actions again.

For the computation of κ_i and $q_i(a_i)$, we need some more notation. Let $p_i(a_i)$ denote the probability that i will have to change actions again when the randomization is observed given that the randomization is uniform on $U_{\geq}$. That is, $p_i(a_i)$ is the probability that a_i is different from the action i plays in a uniformly drawn profile from $U_{\geq}$. (Recall that κ_i is even and the target must end up at a_i at the end of his κ_i changes of action.) Also, let $I_i(a_i) = 0$ if a_i is one of i's maxmin actions and 1 otherwise. Similarly, let $\underline{I}_i(a_i)$ be 0 if a_i is the same action i plays at $\underline{u}^i$ in equilibrium and 1 otherwise.

Let β_i be the smallest integer b such that $b\varepsilon \geq \underline{u}_i^i - w_i$. Note that $\underline{u}_i^i \geq w_i$, so this is well defined. Let κ_i equal the smallest even integer greater than or equal to

$$\beta_i + 1 + 2 \max_{i,j|i\neq j} \frac{\bar{u}_i - \underline{u}_i^i + 1}{\bar{u}_j - w_j}.$$

Note that $\bar{u} \gg w$ implies that κ_i is well defined. Set $q_i(a_i)$ so that

$$1 - q_i(a_i) = \frac{\kappa_i \varepsilon + w_i - \underline{u}_i^i + [\underline{I}_i(a_i) - I_i(a_i)]\varepsilon}{\bar{u}_i - \underline{u}_i^i + [\underline{I}_i(a_i) - p_i(a_i)]\varepsilon}.$$

By construction, $\kappa_i \varepsilon > \underline{u}_i^i - w_i + \varepsilon$, so the numerator is at least

$$\varepsilon + [\underline{I}_i(a_i) - I_i(a_i)]\varepsilon \geq 0.$$

For ε sufficiently small, the denominator must be strictly positive as well since $\bar{u}_i > \underline{u}_i^i$. Finally, as ε goes to zero, the fraction converges to 0, so it must be less than 1 for small enough ε. Hence $q_i(a_i)$ is well defined for ε sufficiently small.

The key fact to note about this choice of $q_i(a_i)$ is that it ensures that the target is indifferent between following the equilibrium punishment and not. To see this, note that the target's expected payoff to following the equilibrium punishment is

$$q_i(a_i)[\underline{u}_i^i - \underline{I}_i(a_i)\varepsilon] + [1 - q_i(a_i)][\bar{u}_i - p_i(a_i)\varepsilon] - \kappa_i \varepsilon.$$

Rearranging, this is

$$\underline{u}_i^i - \underline{I}_i(a_i)\varepsilon - \kappa_i \varepsilon + [1 - q_i(a_i)] \left\{ \bar{u}_i - \underline{u}_i^i + [\underline{I}_i(a_i) - p_i(a_i)]\varepsilon \right\}.$$

Substituting for $1 - q_i(a_i)$ from the above gives

$$\underline{u}_i^i - \underline{I}_i(a_i)\varepsilon - \kappa_i \varepsilon + \kappa_i \varepsilon + w_i - \underline{u}_i^i + [\underline{I}_i(a_i) - I_i(a_i)]\varepsilon = w_i - I_i(a_i)\varepsilon.$$

Suppose that i does not follow the equilibrium punishment. What is the best alternative? Clearly, i can either not change actions ever again or change to one of his maxmin actions (if he is not already playing one) and never change again. If a_i is a maxmin action, staying at this action is the best alternative to following the equilibrium punishment. This would give him a payoff of w_i. So suppose a_i is not one of i's maxmin actions. Let z_i denote i's second best payoff when the others are trying to minmax him. In other words, letting A_i^* denote i's set of maxmin action, define

$$z_i = \max_{a_i \notin A_i^*} \left[\min_{a_{\sim i} \in A_{\sim i}} u_i(a_i, a_{\sim i}) \right].$$

By hypothesis, $a_i \in A_i \setminus A_i^*$, so $A_i^* \neq A_i$. Hence z_i is well–defined. Clearly, $w_i > z_i$. Given that ε is chosen to satisfy (10), $w_i - z_i > \varepsilon$. If i never changes actions again, his payoff is z_i at best. If he changes to one of his maxmin actions, his payoff is $w_i - \varepsilon$. Clearly, then, it is optimal for him to change to one of his maxmin actions. In short, i's payoff if he does not follow the equilibrium punishment is w_i minus the switching cost if he is not already playing one of his maxmin actions, or $w_i - I_i(a_i)\varepsilon$. So i is indifferent between

following the equilibrium punishment and not. In short, the target of a punishment has no incentive to deviate prior to the random determination of u.

To complete the proof that this is a subgame perfect equilibrium, consider any history for which there has been no deviation and any player i. If i does not deviate, his payoff will be at least u_i' (more if he has already carried out some changes of action). If i deviates, his expected payoff will be w_i at best. Since $u' \geq u \geq w$, i has no incentive to deviate.

Consider any history which puts us in punishment mode (i, a_i) and any $j \neq i$. Does j have an incentive to deviate prior to the realization of the public randomization determining u? If j does not deviate, his payoff is at least

$$q_i(a_i)\underline{u}_j^i + [1 - q_i(a_i)]\bar{u}_j - 2\varepsilon.$$

(This would be the case if j has to switch actions at this point to punish the target and will have to switch again once u is realized.) If j deviates, his payoff will be w_j at best. Because $\underline{u}_j^i \geq w_j$, a sufficient condition for j to not deviate is

$$q_i(a_i)w_j + [1 - q_i(a_i)]\bar{u}_j - 2\varepsilon \geq w_j$$

or $[\bar{u}_j - w_j][1 - q_i(a_i)] \geq 2\varepsilon$. From the definition of $q_i(a_i)$ above, we see that a sufficient condition for this for $\varepsilon \leq 1$ is

$$[\bar{u}_j - w_j]\frac{\kappa_i\varepsilon + w_i - \underline{u}_i^i - \varepsilon}{\bar{u}_i - \underline{u}_i^i + 1} \geq 2\varepsilon.$$

From the definition of κ_i,

$$\kappa_i\varepsilon \geq \underline{u}_i^i - w_i + \varepsilon + 2\varepsilon\frac{\bar{u}_i - \underline{u}_i^i + 1}{\bar{u}_j - w_j}.$$

Hence a sufficient condition is

$$\frac{\bar{u}_j - w_j}{\bar{u}_i - \underline{u}_i^i + 1}\ 2\varepsilon\ \frac{\bar{u}_i - \underline{u}_i^i + 1}{\bar{u}_j - w_j} \geq 2\varepsilon,$$

which is obviously true. Hence no player, the target or otherwise, will deviate from a punishment prior to the realization of the public randomization.

Hence it only remains to show that no player will deviate after the realization of the randomization. Consider any player i who played a_i in the period before the realization and suppose $u = (u_1, \ldots, u_I)$ is the realization of the randomization. If i deviates from the equilibrium by playing $\hat{a}_i$, we move into an $(i, \hat{a}_i)$ punishment mode and i's payoff is $w_i - I_i(\hat{a}_i)\varepsilon$ minus ε if $\hat{a}_i \neq a_i$. If, instead, i follows the equilibrium, his payoff is u_i minus

ε if he must switch actions. Recall that $u_i \geq w_i$. First, suppose $u_i > w_i$. By (10), then, $u_i - \varepsilon > w_i$. Hence i has no incentive to deviate since the worst he could do by following the equilibrium is strictly better than the best he could do by deviating. Suppose, then, that $u_i = w_i$. Then i's payoff from following the equilibrium is $w_i - I_i(a_i)\varepsilon$. If $\hat{a}_i = a_i$, this is exactly what i would get if he deviated. If $\hat{a}_i \neq a_i$, then i's payoff to deviating is

$$-\varepsilon + w_i - I_i(\hat{a}_i)\varepsilon \leq w_i - \varepsilon \leq w_i - I_i(a_i)\varepsilon.$$

Hence either way, i has no incentive to deviate.

This demonstrates that $c(U_{\geq}) \cap W \subseteq \lim_{\varepsilon \downarrow 0} \mathcal{U}_\infty(\varepsilon)$. To complete the proof, then, suppose $u \notin c(U_{\geq}) \cap W$. We now show that there is no equilibrium payoff nearby. Since $c(U_{\geq}) \cap W$ is closed and does not contain u, for every sufficiently small $\varepsilon > 0$ and every u' within ε of u, we have $u' \notin c(U_{\geq}) \cap W$. Choose any such ε which is small enough that

$$\varepsilon < w_i - u_i(a)$$

for all a and i such that $u_i(a) < w_i$. Suppose, contrary to our claim, that there is a u' within ε of u with $u' \in \mathcal{U}_\infty(\varepsilon)$. Fix such a u' and the equilibrium generating this payoff.

Obviously, we must have $u' \in W$ as any player i can guarantee himself a payoff of w_i from a constant action. Hence the fact that $u' \notin c(U_{\geq}) \cap W$ implies that $u' \notin c(U_{\geq})$. Hence there is some action profile $\hat{a}$ with $u_i(\hat{a}) < w_i$ which is played infinitely often with strictly positive probability.

Let $\mathcal{P}$ denote the set of paths (infinite sequences of action profiles) in the support of the equilibrium which generates payoff u'. Let μ denote the probability distribution over $\mathcal{P}$ induced by the equilibrium (including the effect of the public randomizations if strategies are based on these). Let $\mathcal{P}^*(a)$ denote the set of paths in $\mathcal{P}$ which eventually absorb at action profile a — that is,

$$\mathcal{P}^*(a) = \{(a^1, a^2, \ldots) \in \mathcal{P} \mid \exists T \text{ such that } a^t = a,\ \forall t \geq T\}.$$

Similarly, let $\mathcal{P}_d$ denote the set of paths in $\mathcal{P}$ which do not absorb — that is, $\mathcal{P}_d = \mathcal{P} \setminus \cup_{a \in A} \mathcal{P}^*(a)$. It is not hard to see that $\mu(\mathcal{P}_d) = 0$. If this were not zero, then the expected total switching costs of the players would necessarily be infinite, meaning that some player's payoff is $-\infty$, so obviously his strategy cannot be optimal.

Our selection of $\hat{a}$ implies that $\mu(\mathcal{P}^*(\hat{a})) > 0$. Recall that we chose $\hat{a}$ to be some profile played infinitely often with strictly positive probability. Since no path in $\mathcal{P}^*(a)$ for $a \neq \hat{a}$ has this property and since $\sum_{a \in A} \mu(\mathcal{P}^*(a)) = 1$, we must have $\mu(\mathcal{P}^*(\hat{a})) > 0$.

Let $\mathcal{P}^t(a)$ denote the set of paths in $\mathcal{P}$ with $a^t = a$. For any path $p \in \mathcal{P}^*(a)$, $a \neq \hat{a}$, there is a T such that $p \notin \mathcal{P}^t(\hat{a})$ for all $t \geq T$. That is, if a path eventually stays at $a \neq \hat{a}$

forever, it must visit $\hat{a}$ for a last time at some finite date. Hence for all $a \neq \hat{a}$,

$$\mathcal{P}^*(a) \cap \mathcal{P}^t(\hat{a}) \to \emptyset$$

as $t \to \infty$. On the other hand, consider any $p \in \mathcal{P}^*(\hat{a})$. By definition, there is a T such that this path has $a^t = \hat{a}$ for all $t \geq T$. Hence there is a T such that this path is in $\mathcal{P}^t(\hat{a})$ for all $t \geq T$. Hence

$$\mathcal{P}^*(\hat{a}) \cap \mathcal{P}^t(\hat{a}) \to \mathcal{P}^*(\hat{a})$$

as $t \to \infty$.

In light of this, consider

$$\mu(\mathcal{P}^*(\hat{a}) \mid \mathcal{P}^t(\hat{a})) = \frac{\mu(\mathcal{P}^*(\hat{a}) \cap \mathcal{P}^t(\hat{a}))}{\mu(\mathcal{P}^t(\hat{a}))}.$$

Since all the $\mathcal{P}^*(a)$ sets are disjoint and their union has probability 1, we can rewrite this as

$$\frac{\mu(\mathcal{P}^*(\hat{a}) \cap \mathcal{P}^t(\hat{a}))}{\sum_{a \in A} \mu(\mathcal{P}^*(a) \cap \mathcal{P}^t(\hat{a}))}.$$

Clearly, as $t \to \infty$, this converges to

$$\frac{\mu(\mathcal{P}^*(\hat{a}))}{\mu(\mathcal{P}^*(\hat{a}))} = 1.$$

(Note that this is well defined since $\mu(\mathcal{P}^*(\hat{a})) > 0$.) In short, $\mu(\mathcal{P}^*(\hat{a}) \mid \mathcal{P}^t(\hat{a})) \to 1$ as $t \to \infty$.

Consider player i, the player for whom $u_i(\hat{a}) < w_i$. Fix any t for which $\hat{a}$ is played at t with positive probability in the equilibrium. Consider the following strategy for player i: follow the equilibrium strategy until the equilibrium strategies call for $\hat{a}$ to be played at period t. Then deviate to one of the maxmin actions forever after. Clearly, since the original strategies form an equilibrium, this alternative strategy cannot be better for i for any choice of t. In comparing i's payoff in the equilibrium to i's payoff from the alternative strategy, obviously, we can condition on the set of paths for which $\hat{a}$ is played at time t — for any other paths, the payoff difference is zero. If player i deviates at time t as specified, his expected payoff from that point onward is at least $w_i - \varepsilon$. But as $t \to \infty$, player i's expected continuation payoff if he does not deviate is converging to $u_i(\hat{a})$. Since $u_i(\hat{a}) < w_i - \varepsilon$, there is some large t for which i strictly prefers the deviation, a contradiction. ▌

References

[1] D. Abreu, P. Dutta, and L. Smith, "The Folk Theorem for Repeated Games: A NEU Condition," *Econometrica*, **62**, July 1994, 939–948.

[2] J.-P. Benoit and V. Krishna, "Finitely Repeated Games," *Econometrica*, **53**, 1985, 890–904.

[3] S. Chakrabarti, "Characterizations of the Equilibrium Payoffs of Inertia Supergames," *Journal of Economic Theory*, **51**, 1990, 171–83.

[4] P. Dutta, "A Folk Theorem for Stochastic Games," *Journal of Economic Theory* **66**, 1995, 1–32.

[5] P. Dutta, "Collusion, Discounting, and Dynamic Games," *Journal of Economic Theory* **66**, 1995, 289–306.

[6] D. Fudenberg and D. Levine, "Subgame–Perfect Equilibria of Finite and Infinite Horizon Games," *Journal of Economic Theory*, **31**, 1983, 227–256.

[7] D. Fudenberg and E. Maskin, "On the Dispensability of Public Randomization in Discounted Repeated Games," *Journal of Economic Theory*, **53**, April 1991, 428–438.

[8] D. Fudenberg and J. Tirole, *Game Theory*, Cambridge: MIT Press, 1991.

[9] B. Lipman and R. Wang, "Switching Costs in Frequently Repeated Games," *Journal of Economic Theory*, **93**, August 2000, 149–190.

[10] Q. Wen, "The 'Folk Theorem' for Repeated Games with Complete Information," *Econometrica*, **62**, July 1994, 949–954.