Unique and Optimal Perfect Equilibrium

Prajit K. Dutta   Paolo Siconolfi
Columbia University

January 8, 2016

Abstract

This paper proves a First Welfare Theorem for Games - it shows that asynchronous dynamic games with voluntary one period ahead transfers have a unique optimal equilibrium. The equilibrium coincides with the Utilitarian Pareto Optimum and hence can be computed from a (simpler) programming problem (rather than as a fixed point). Whilst it is commonly thought that Folk Theorems are endemic in dynamic games, this paper shows that two simple changes in the standard model give us the other extreme - a unique optimal equilibrium. More broadly, this paper is a contribution towards uniting Mechanism Design and Repeated Games.

1 Introduction

This paper proves a First Welfare Theorem for Games - it shows that asynchronous dynamic games with voluntary one period ahead transfers have a unique optimal equilibrium. Whilst it is commonly thought that Folk Theorems are endemic in dynamic games, this paper shows that two simple changes in the standard model give us the other extreme - a unique optimal equilibrium. Put another way, Folk Theorems rely critically on two assumptions - simultaneous moves and no transfers.

Yet neither assumption is sacrosanct. An assumption of simultaneous moves may be no more "natural" than one of alternating moves. In an oligopoly context, say, it may be more unnatural to assume that firms literally move at the same time or that firms' moves do not become known to their competitors before the latter has to move. (See, for example, Maskin-Tirole (1988a, b).) Similarly, in many applications, a player can promise a conditional transfer to another player. This would be true between countries (foreign aid, investment, trade - see, for example, Klimenko, Ramey, Watson (.)), between firms (licensing or franchising fees - see, for example, Harrington-Skrypcaz (.)), between family members (gifts, alimony), etc.

Forcing moves to be simultaneous introduces an additional "coordination problem" (beyond that present in the payoffs). For example, multiple equilibria arise in, say, the Common Interest game only if players are forced to move simultaneously. Allowing players to alternate moves might reduce the number of equilibria; this is a first insight we explore. A second insight is that transfers can be a more efficient way to provide intertemporal incentives compared to the Abreu-Pearce-Stachetti method of variations in continuation payoffs (which is what is used in Repeated Games in the absence of transfers).

When we began research our hope was that, given alternating moves, we should have fewer equilibria and, given transfers, we should have more efficient equilibria. Our eventual result is that indeed there are fewer equilbria - being one! - and it is fully efficient.

Just having one of those two features is not enough - there is a Folk Theorem when moves are alternating but there are no transfers. Such a game is an example of a stochastic game and Dutta (1995) applies. There is also a Folk Theorem when transfers are allowed but players move simultaneously. For instance, see Dutta and Siconolfi (2016) or Goodlucke-Krantz (2012). Hence, we find minimally necessary conditions for our result.

More broadly, we view this paper as part of a research agenda which marries Mechanism Design to the theory of Repeated Games. Both fields have been hugely influential and yet they have kept their distance from each other. Mechanism Design rarely ventures into dynamic considerations. For its part, Repeated Games takes the basic model as given and analyzes equilibrium possibilities within that model. What this paper suggests is that there is a healthy and productive research agenda driven by the following question: is there way to set up the dynamic model such that all equilibria have desirable properties?

Before describing our model and result, it might be useful to review the reach and logic of Folk Theorems in Repeated Games, clearly the most celebrated result in that field. Folk Theorems show up in the canonical complete information model of infinitely and finitely repeated games (respectively, Fudenberg-Tirole (1986) and Benoit-Krishna (1985)) as also in variants of the canonical model - in the more general set-up of stochastic games, (Dutta (1995), Fudenberg-Yamamoto (2013)), and with overlapping generations of players (Smith (1997)). They also show up in informational variants - with imperfect public monitoring (Fudenberg-Levine-Maskin (1994)), and even with private monitoring (Horner-Olszewski (2006, 2009) and Sugaya (2012, 2013)).[1]

The driving force behind the Folk Theorem is the logic of reciprocity - that "players will do unto others what they expect others to do unto them". That logic delivers the result; every individually rational payoff is also an equilibrium payoff (for sufficiently patient players) or, colloquially, "anything can happen" in equilibrium. This conclusion is disappointing since it robs the theory of Repeated Games of any predictive power.

The model that we study in this paper is a variant on a repeated game. As in the latter, there is a fixed stage game that, say, two players play over and over again (over a finite or infinite horizon). In each period, the players get

[1] The literature on Folk Theorems is truly vast. Sugaya (2013) is a fairly up-to-date reference on informational variants (and contains about fifty citations). Pesci and Wiseman (2015) is a good reference for models that keep the assumption of (possibly imperfect) public monitoring but allow state variables (and this paper has another fifty citations).

payoffs and lifetime payoffs are simply the (discounted) sum of period payoffs. So far this is as in a repeated game.

The departures are two-fold. First, players take turns moving. So, at any period t, the payoffs are derived from the action taken by the mover in period t as well as the (fixed) action taken by his opponent in $t-1$. Second, the mover also picks a non-negative transfer that she pays her opponent in the next period depending on her opponent's action.

To fix ideas, consider a stage game such as the Prisoners Dilemma:

$1\backslash 2$	$confess$	$not\ confess$
$confess$	$0,0$	c,b
$not\ confess$	b,c	d,d

with $confess$ as the dominant strategy and $not\ confess$, $not\ confess$ the Pareto dominant tuple ($c > d > 0 > b$, $2d > \max(b+c, 0)$).

To understand period t payoffs, suppose the $t-1$ period action was $not\ confess$ and the promised transfer schedule was $(0, \theta)$ implying a payment $\theta > 0$ would be made only if the t period mover also picked $not\ confess$. Then, if he did indeed do so, his payoff would be $d+\theta$ while the non-mover would get $d-\theta$ in that period. Alternatively, if he picked $confess$ then his payoff would be c while hers would be b.

As is well-known, in the standard repeated game model with simultaneous moves and no transfers, "$confess$, $confess$ always and after all histories" is a subgame perfect equilibrium (SPE) strategy (with 0,0 payoffs, the minmax payoffs). In turn, via the well-known grim trigger strategy, that SPE forms the punishment norm that supports any strictly individually rational (positive) payoff as a SPE payoff. This is the Folk Theorem construction for the (infinitely) repeated Prisoners Dilemma game.

Let us ask whether in our model "$confess$, $confess$ always and after all histories with no transfers" is a SPE? Without loss, we check for the viability of one-shot deviations. Suppose, again, that the fixed $t-1$ period action was $not\ confess$ and the promised transfer schedule was $(0, \theta)$. It is easy to see that the transfer θ needs to be large enough that the t period mover is (at least) indifferent between playing $confess$ and $not\ confess$, i.e., with δ as the discount factor,

$$d + \theta + \delta b = c$$

where the LHS is the lifetime payoff under this strategy from playing $not\ confess$ in period t: $(d+\theta)$ in that period plus the payoff in period $t+1$ (b) when the then mover will switch and play $confess$, followed by 0 payoffs every period thereafter. The RHS is the lifetime payoff under this strategy from playing $confess$ in period t; c followed by 0 every period after.

Should the period $t-1$ mover offer this transfer? She will if her lifetime payoff starting in period t is higher after paying this transfer, i.e. if,[2]

$$d - \theta + \delta c \geqslant b$$

which, after substituting the value of θ, simplifies and strengthens to

$$2d - (1 - \delta)(b + c) > 0$$

and that holds for all δ under the Pareto optimality hypothesis that $2d > \max(b + c, 0)$. In other words, the "bad equilibrium" of the standard model is not an equilibrium here. Whereas, in the standard model, one of the player switching to *not confess* is simply treated as a mistake here its logical implication is that she will thereafter provide a transfer that has the other player switch as well (at least for a period).

Of course, we still have a long way to go. Knowing that the (bad) *always confess* strategy is not an SPE simply rules out 0 as a SPE payoff. We do not yet know what positive payoffs *can* arise in equilibrium. What we will show as our main result is that **there is a unique SPE and it coincides with the utilitarian Pareto optimum solution (UPO)**, i.e, the sum of SPE payoffs (across players) must equal the maximum sum of player payoffs. One can intuit that result by the following re-write of the last inequality:

$$2d + \delta(b + c) > b + c$$

Notice that, once the $t - 1$ period mover has played *not confess*, the LHS is the sum of player payoffs to a "one-shot" deviation from the *always confess* strategy while the RHS is the sum of payoffs from continuing with that strategy. Evidently, maximizing the lifetime sum of payoffs is a Dynamic Programming problem. By a standard "Unimprovability" result in Dynamic Programming, a strategy is optimal iff it cannot be improved by a one-shot deviation. Since the *always confess* strategy can be improved, it is *not* UPO. Furthermore, the same deviation proves that the strategy is neither an SPE under transfers nor is it UPO which hints at a link between the two solutions.

In what follows, we generalize all this in two significant ways: we show that *any* strategy that is not UPO is subject to a profitable unilateral one-shot transfer deviation, i.e., is not an SPE. This proof requires a generalized version of the Abreu-Pearce-Stachetti (1986) construction of SPE in repeated games. We further show all that in the context of *any* two-player stage game (and not just the Prisoners Dilemma).

In this paper it is assumed that one-period ahead moves are contractible - or, equivalently, can be committed to. For actions, that assumption is embedded in the timing of alternating moves; a player is not allowed to change her action in the next period. Similarly, when she picks a transfer schedule it is assumed

[2] Note that these are the lifetime payoffs starting in period t, for the mover in period $t-1$. Hence, one way of thinking about the timing is that she has already decided to play *not confess* in period $t-1$ and is asking herself what transfer should she offer for the subsequent period t.

that she is committed to make the appropriate payment in the next period. This is a natural assumption in many contexts; for example, when money can be placed in an escrow account. Or when a country delivers foreign aid into the World Bank's CDMA account (to encourage adoption of clean technology). Or when firms agree to pay licensing fees. Note that the commitment is only for one period. Yet it implies an infinite period optimality. In Section 6 we discuss what happens when this assumption is not made.

The paper is organized as follows. In Section 2 we detail the model and in Section 3 prove the result for the infinite horizon case. Section 4 proves the analogous result for the finite horizon. Section 5 contains an example that illustrates some issues that arise in the presence of infinite transfer histories. In Section 6 we discuss extensions and generalizations while Section 7 concludes.

2 Model

In this section, we present a general model. Let G denote a two player stage game (in strategic form). (In Section 6 we extend the analysis to N players.) Denote player i's strategy set A_i and her payoff function π_i, where, as usual, $\pi_i : A \to \mathbb{R}$, and $A \equiv A_1 \times A_2$ is the set of strategy tuples for the players. Suppose that A_i is finite for every i.[3] Let i denote the generic player and let j denote the "other" player, $j \neq i$.

The timing structure is that of **alternating moves** with players moving sequentially. Hence, player i gets to move in period t followed by player j in period $t+1$ and so on. Whenever it is player $i's$ turn to move, she can choose an action a_{it} from the set of feasible actions, A_i - and this action then remains fixed till the next time i can change her action. Payoffs are ongoing. If $\overline{a}_j$ denotes the (fixed) action of the other player, then the period t stage game payoffs are $\pi_i(a_{it}, \overline{a}_j)$, $\pi_j(a_{it}, \overline{a}_j)$. There is an *initial action state* $\overline{a}$ for the game.[4] We will consider all possible initial action states, i.e., all possible fixed actions and both possible first movers.

Additionally, suppose that a player can also make a conditional **transfer** to the other player - conditional in that it can be tied to the next period action of the other player. For example, if in the previous period the then mover (player j) had played $\overline{a}_j$ and promised the current mover that she would be paid according to the schedule, $\theta_j(a_i|\overline{a}_j)$ then the total payoffs of player i inclusive of transfers will be $\pi_i(a_{it}, \overline{a}_j) + \theta_j(a_i|\overline{a}_j)$ while that of player j would be $\pi_j(a_{it}, \overline{a}_j) - \theta_j(a_i|\overline{a}_j)$.

One interpretation of these transfers is that they are escrow commitments; if player j promises according to the transfer schedule $\theta_j(.|\overline{a}_j)$ then an amount to cover the commitment is posted to an escrow account from which funds equal to $\theta_j(a_i^{'}|\overline{a}_j)$ are paid out upon play of the action $a_i^{'}$. How much information about the transfer schedules and the actual transfers paid out is publicly known will

[3] The analysis extends to the case of A_i compact.

[4] If, say, player i is the mover in the initial period 0, then the two players's payoffs in that initial period depend on the pair of actions $(a_{i0}, \overline{a})$.

depend on the escrow mechanism. We make a couple of different assumptions - detailed below - on how much is publicly known and publicly verifiable of past escrow commitments and actual transfers.

The horizon can be infinite or finite. In the latter case, lifetime payoffs will be the undiscounted sum. In the discounted case, the lifetime payoffs will be evaluated according to the discounted *average;* player $i's$ evaluation of the payoff will be given by

$$\begin{aligned}(1-\delta)\textstyle\sum_{t=0}^{\infty}\delta^t\{&[\pi_i(a_{it},a_{jt-1})+\theta_j(a_i|a_{jt-1}) \parallel i \; mover]+\\ &+[\pi_i(a_{it-1},a_{jt})-\theta_i(a_j|a_{it-1}) \parallel j \; mover]\}\end{aligned} \tag{1}$$

2.1 Equilibrium

Let us start with the assumption on what we take to be publicly observable/verifiable for all parties:

Informational Assumption - 1. *All stage game actions are observable.* Hence, an *action history* h_t^a at time t includes all the past actions - the actual history of play of the stage game, $h_t^a = (\overline{a},\, a_{j0}, a_{i1},\, ...a_{jt-1})$.

2. *Transfer/escrow commitments going back M periods are observable*, where $1 \leqslant M < \infty$. Hence, a *transfer history* h_t^θ at time t includes M past escrow commitments - $h_t^\theta = (\theta_{jt-M}(.), \theta_{it-M+1}(.),\, ...\theta_{jt-1}(.))$.

Note that the transfer history includes information about the entire transfer schedule and not just the actual transfer that was made in a previous period. Any results that we prove will continue to hold if we replaced that requirement with one that said only the actual transfers are observable. Of course, the entire schedule needs to be known for the immediately preceding period, i.e., for $\theta_{jt-1}(.)$, since the $t-th$ period action choice of player i will depend on that (payoff-relevant) schedule. Put another way, M must be at least $\geqslant 1$. The restriction that infinitely long transfer commitments are not publicly verifiable seems plausible. Typically, banks or other financial institutions, to which such an escrow commitment is posted, wipe out their records when a period of time has elapsed. We will have some results on the case $M = \infty$ as well.

To summarize, a *history* h_t at time t is given by $h_t = (h_t^a, h_t^\theta)$. Our main theorem will apply to equilibria that condition on all histories h_t. (The build-up to that result will be one that allows only the currently outstanding transfer commitments to be observable, i.e., that preliminary result will assume that $M = 1$.)

Note also the following about the *payoff-relevant history.* For the $t-th$ period *action choice* a_{it}, the payoff relevant parts of the history are only the preceding period's choices: $a_{jt-1}, \theta_{jt-1}(.)$, the fixed action of the other player j and the transfer schedule that j has picked. For the $t-th$ period *transfer choice* $\theta_{it}(.)$, which is only paid in period $t+1$, there is *no* payoff relevant part of the history h_t since the $t+1$ period payoffs will depend on a_{it} and a_{jt+1} neither of which are in h_t. (The $t-th$ period payoffs will also, of course, depend on the yet to be made choice of $\theta_{it}(.)$). We will exploit this difference in payoff-relevance between a_{it} and θ_{jt-1} in the analysis that follows.

Let us though formally define strategies and equilibria first.

A $t-th$ period strategy σ_{it} for player i is an action choice a_{it} that maps from a history h_t and a transfer choice $\theta_{it}(.)$ that maps from (h_t, a_{it}). A strategy for player i in the game, σ_i, is a specification of a strategy σ_{it} for that player in every period that she is the mover. A strategy vector - one strategy for every player - defines in the usual way a (possibly probabilistic) history h_t. Denote the lifetime payoff of the mover at time t, v_i:

$$v_i(h_t) = \max_{a_i}\{(1-\delta)[\pi_i(a_i, a_{jt-1}) + \theta_{jt-1}(a_i)] + \delta w_i(h_t, a_i)\} \tag{2}$$

where w_i is the continuation for player i that follows from the optimal choice of a transfer schedule θ_{it}, given a_{it}, and anticipating the "Stackelberg follower" player $j's$ best response a_{jt+1} to that transfer, i.e.,

$$w_i(h_t, a_{it}) = \max_{\theta_i(.)}\{(1-\delta)[\pi_i(a_{it}, a_j(\theta_i)) - \theta_i(a_j)] + \delta v_i(h_{t+2}(a_{it}, \theta_{it}))\} \tag{3}$$

where $h_{t+2}(a_{it}, \theta_{it})$ is the history at period $t+2$ caused by player $i's$ period t actions. Hence, $h_{t+2}(a_{it}, \theta_{it}) = (h_t^a, a_{it}, a_{jt+1}(\theta_{it}); \theta_{jt-M+2}(.), ...\theta_{it}(.), \theta_{jt+1}(.))$, the entire action history and the last M periods of transfer schedules. In that history, the follower player $j's$ best response action in period $t+1$, $a_j(\theta_{it})$, comes from the $t+1$ period analog of Eq. 2:[5]

$$\begin{aligned} v_j(h_{t+1}(a_{it}, \theta_{it})) &= (1-\delta)[\pi_j(a_{it}, a_j(\theta_{it})) + \theta_{it}(a_j(\theta_{it}))] + \delta w_j(h_{t+1}, a_j(\theta_{it})) \\ &= \max_{a_j}\{(1-\delta)[\pi_j(a_{it}, a_j) + \theta_{it}(a_j)] + \delta w_j(h_{t+1}, a_j)\} \end{aligned} \tag{4}$$

and his optimal transfer schedule θ_{jt+1} must be a solution to the optimization problem

$$w_j(h_{t+1}, a_j) = \max_{\theta_j(.)}\{(1-\delta)[\pi_j(a_i(\theta_j)), a_j) - \theta_j(a_i)] + \delta v_j(h_{t+3}))\}$$

A *Subgame Perfect Equilibria* (SPE) is a pair of strategies that are best responses to each other.

Markov Strategies are choices that depend only on the payoff-relevant histories. A $t-th$ period Markov strategy for player i is hence an action choice a_{it} that maps from $a_{jt-1}, \theta_{jt-1}(.)$ and a constant function $\theta_i(.)$. A Markov Perfect Equilibria (MPE) is a pair of Markov strategies that are best responses to each other. As is well-known, MPE are also SPE.

3 Main Result: Equilibrium is Unique and Optimal

Let us start with a benchmark - the **Utilitarian Pareto Optimum**:

[5] Note that $h_{t+1}(a_{it}, \theta_{it}) = (h_t^a, a_{it}; \theta_{jt-M+1}(.), ... \theta_{it}(.))$.

Definition 1 *The Utilitarian Pareto optimum (UPO) problem is the maximization of the discounted sum of players payoffs subject to an initial condition specifying a fixed action of the first period non-mover:*

$$\max_{\{a_{it},a_{jt}\}_{t\geq 0}} (1-\delta)\sum_{t=0}^{\infty}\delta^t[\pi_i(a_{it},a_{jt})+\pi_j(a_{it},a_{jt})] \tag{5}$$

$$s.t.\ a_{i0} = \bar{a}_i,\ a_{it}=a_{it-1},\ t\geqslant 1,\ t\ odd,\ and\ a_{jt}=a_{jt-1},\ t\ \geqslant 0,\ t\ even$$

It is well known that the solution is Markovian, i.e., there is $\widehat{\beta}_i(a_j)$ - respectively, $\widehat{\beta}_j(a_i)$ - such that the Pareto optimal solution is for player i to play according to $\widehat{\beta}_i(a_j)$ - respectively for player j to play $\widehat{\beta}_j(a_i)$ - whenever it is her turn to move.

The **main theorem** that ties equilibrium behavior to the UPO is:

Theorem 2 *There is a unique SPE to the game. It coincides with the Pareto optimum solution in terms of actions - the equilibrium is for player i to play according to $\widehat{\beta}_i(a_j)$ - respectively for player j to play $\widehat{\beta}_j(a_i)$ - whenever it is her turn to move. The transfers that are required are also unique.*

Since the UPO is a Markov strategy, it follows from the theorem that not only is there a unique SPE but it is, in fact, a MPE.

3.1 The Proof When Only the Last Transfer is Observed

It is easier to break the proof into two parts, first when $M=1$ and then the general case. The proof will proceed by way of a modified Bellman-Abreu-Pearce-Stachetti (BAPS) argument. There will be a set of steps.

3.1.1 Definition of Two Operators and an Equivalence

Let V_i and W_j be correspondences with domain A_j (and range $\mathbb{R}$). These are to be thought of as potential SPE continuation payoffs starting at a period when player i is the mover, and a_j is fixed, with continuation payoff $v_i(.)$ and $w_j(.)$ respectively for players i and j (and the arguments in the continuations could be all or part of the history of play h_t). Respectively, let V_j and W_i be correspondences with domain A_i (and range $\mathbb{R}$) that are to be thought of as potential SPE payoffs starting at a period when player j is the mover.

Define the following operator that determines player $i's$ decision on transfer

$\theta_i(a_j|\overline{a}_i)$ if she were to, simultaneously, play the action $\overline{a}_i$.

$$\begin{aligned}
\Gamma(V_i, W_j)(\overline{a}_i) &= \{w_i, v_j : \exists a_j^*, \theta_i^*(a_j \mid \overline{a}_i),\ v_i(a_j \mid \overline{a}_i), w_j(a_j \mid \overline{a}_i) \in V_i, W_j : \\
a_j^* &\in \arg\max_{a_j}\{(1-\delta)[\pi_j(\overline{a}_i, a_j) + \theta_i^*(a_j \mid \overline{a}_i)] + \delta w_j(a_j \mid \overline{a}_i)\} \\
v_j &= (1-\delta)[\pi_j(\overline{a}_i, a_j^*) + \theta_i^*(a_j^* \mid \overline{a}_i)] + \delta w_j(a_j^* \mid \overline{a}_i) \\
\forall a_j^{'}, \theta_i^{'}(.),\ a_j^{'} &\in \arg\max_{a_j}\{(1-\delta)[\pi_j(\overline{a}_i, a_j) + \theta_i^{'}(a_j)] + \delta w_j(a_j \mid \overline{a}_i)\} \\
w_i &= (1-\delta)[\pi_i(\overline{a}_i, a_j^*) - \theta_i^*(a_j^* \mid \overline{a}_i)] + \delta v_i(a_j^* \mid \overline{a}_i\} \\
&\geqslant (1-\delta)[\pi_i(\overline{a}_i, a_j^{'}) - \theta_i^{'}(a_j^{'})] + \delta v_i(a_j^{'} \mid \overline{a}_i)\}
\end{aligned} \tag{6}$$

Put another way, the operator Γ specifies equilibrium payoffs to a family of "one-shot" Stackelberg games with player i as the first mover choosing a transfer schedule $\theta_i(.|\overline{a}_i)$ and player j as the follower picking action a_j. The payoffs in these one-shot games are defined by two parts - a) a stage game payoff augmented by the transfer determined by the schedule (respectively, $\pi_i(\overline{a}_i, a_j) - \theta_i(a_j)$ and $\pi_j(\overline{a}_i, a_j) + \theta_i(a_j)$), and b) a "termination" payoff that needs to be drawn from the given correspondences (V_i, W_j). Subject to that feasibility constraint, these termination payoffs can depend on the action chosen by player j, a_j (and hence are written, respectively, $v_i(. \mid \overline{a}_i), w_j(. \mid \overline{a}_i)$.[6]

Consider a related - and seemingly simpler - operator T defined on the same domain as Γ:

$$\begin{aligned}
T(V_i, W_j)(\overline{a}_i) &= \{w_i, v_j : \exists \widehat{a}_j, v_i(a_j \mid \overline{a}_i), w_j(a_j \mid \overline{a}_i) \in V_i, W_j(a_j) : \\
\widehat{a}_j &\in \arg\max_{a_j}\{(1-\delta)\pi_j(\overline{a}_i, a_j) + \delta w_j(a_j \mid \overline{a}_i)\} \\
v_j &= (1-\delta)\pi_j(\overline{a}_i, \widehat{a}_j) + \delta w_j(\widehat{a}_j \mid \overline{a}_i) \\
w_i &= \max_{\theta_i, a_j}\{(1-\delta)[\pi_i(\overline{a}_i, a_j) - \theta_i(a_j)] + \delta v_i(a_j \mid \overline{a}_i)\} \\
\text{s.t. } v_j &= (1-\delta)[\pi_j(\overline{a}_i, a_j) + \theta_i(a_j)] + \delta w_j(a_j \mid \overline{a}_i),\ a_j \in A_j
\end{aligned} \tag{7}$$

As with the operator Γ, the simpler operator $T(V_i, W_j)$ also defines "continuation payoffs" starting at a period when player i is the mover. Note that in T, the transfer schedule $\theta_i(.)$ is implicitly defined given the continuation selections $v_i(. \mid .), w_j(. \mid .)$ whereas in Γ the transfer is seemingly determined alongside the continuations. As with all Bellman and APS constructions, the idea of these operators is to reduce the infinite horizon problem to a sequence of one period problems and use the Unimprovability Principle to show an equivalence between the solution to the one period problem and the (original) infinite horizon one.

We first show that two operators are equivalent:

Lemma 3 *Given correspondences, V_i and W_j, $\Gamma(V_i, W_j)(\overline{a}_i) = T(V_i, W_j)(\overline{a}_i)$, for all $\overline{a}_i$. Similarly, $\Gamma(V_j, W_i)(\overline{a}_j) = T(V_j, W_i)(\overline{a}_j)$, for all $\overline{a}_j$.*

[6] In the usual way, the entire action history h_t^a gets incorporated since the termination payoffs can be drawn differently for every h_t^a.

Proof. *We first show that* $T(V_i, W_j)(\overline{a}_i) \sqsubseteq \Gamma(V_i, W_j)(\overline{a}_i)$. *Given* $w_i, v_j \in T(V_i, W_j)(\overline{a}_i)$, *we have continuations* $v_i(. \mid .), w_j(. \mid .)$ *and an associated "no transfer" most preferred action* $\widehat{a}_j$ *that satisfy Eq. 7. From the last line of that equation, we can derive the transfer schedule* $\theta_i^*(. \mid .)$ *as:*

$$(1-\delta)\theta_i^*(a_j \mid \overline{a}_i) = (1-\delta)[\pi_j(\overline{a}_i, \widehat{a}_j) - \pi_j(\overline{a}_i, a_j)] + \delta[w_j(\widehat{a}_j \mid \overline{a}_i) - w_j(a_j \mid \overline{a}_i)]$$

Note that it follows that $\theta_i^*(\widehat{a}_j \mid \overline{a}_i) = 0$. *It further follows from the last line of Eq. 7 that all actions* a_j *yield the same "one-shot" payoff* $(1-\delta)[\pi_j(\overline{a}_i, a_j) + \theta_i^*(a_j)] + \delta w_j(a_j \mid \overline{a}_i)$ *to player* j. *Define* a_j^* *as*

$$a_j^* \in \arg\max_{a_j}\{(1-\delta)[\pi_i(\overline{a}_i, a_j) - \theta_i(a_j)] + \delta v_i(a_j \mid \overline{a}_i)\} \tag{8}$$

Consider an alternative transfer schedule $\theta_i^{'}(.)$ *and a best action choice* $a_j^{'}$, *i.e., suppose that* $a_j^{'} \in \arg\max_{a_j}\{(1-\delta)[\pi_j(\overline{a}_i, a_j) + \theta_i^{'}(a_j)] + \delta w_j(a_j \mid \overline{a}_i)\}$. *In particular, then*

$$\begin{aligned}(1-\delta)[\pi_j(\overline{a}_i, a_j^{'}) + \theta_i^{'}(a_j^{'})] + \delta w_j(a_j^{'} \mid \overline{a}_i) &\geqslant (1-\delta)[\pi_j(\overline{a}_i, \widehat{a}_j) + \theta_i^{'}(\widehat{a}_j)] + \delta w_j(\widehat{a}_j \mid \overline{a}_i) \\ &\geqslant (1-\delta)[\pi_j(\overline{a}_i, \widehat{a}_j) + \theta_i^{*}(\widehat{a}_j)] + \delta w_j(\widehat{a}_j \mid \overline{a}_i) \\ &= (1-\delta)[\pi_j(\overline{a}_i, a_j^{'}) + \theta_i^{*}(a_j^{'})] + \delta w_j(a_j^{'} \mid \overline{a}_i)\end{aligned}$$

■

Proof. *the second-last inequality following from the fact that* $\theta_i^*(\widehat{a}_j \mid \overline{a}_i) = 0$ *(as established above) and the last equality from the fact that player* j *is indifferent across actions for the transfer schedule* θ_i^{*}. *Looking at the two outer terms in the full inequality we can then conclude that* $\theta_i^{'}(a_j^{'}) \geqslant \theta_i^{*}(a_j^{'})$. *From Eq. 8 above we know that*

$$\begin{aligned}w_i &= (1-\delta)[\pi_i(\overline{a}_i, a_j^*) - \theta_i^*(a_j^*)] + \delta v_i(a_j^* \mid \overline{a}_i) \\ &\geqslant (1-\delta)[\pi_i(\overline{a}_i, a_j^{'}) - \theta_i^*(a_j^{'})] + \delta v_i(a_j^{'} \mid \overline{a}_i) \\ &\geqslant (1-\delta)[\pi_i(\overline{a}_i, a_j^{'}) - \theta_i^{'}(a_j^{'})] + \delta v_i(a_j^{'} \mid \overline{a}_i)\end{aligned}$$

■

and the last inequality follows from $\theta_i^{'}(a_j^{'}) \geqslant \theta_i^{*}(a_j^{'})$. *Hence, all the requirements of Eq. 6 have been satisfied, i.e., we have established that* $w_i, v_j \in T(V_i, W_j)(\overline{a}_i)$ *imply that* $w_i, v_j \in \Gamma(V_i, W_j)(\overline{a}_i)$.

For the opposite inclusion, suppose that $w_i, v_j \in \Gamma(V_i, W_j)(\overline{a}_i)$. *Given* $w_i, v_j \in \Gamma(V_i, W_j)(\overline{a}_i)$, *we know that we have continuations* $v_i(. \mid .), w_j(. \mid .)$ *and we can define an associated "no transfer" most preferred action* $\widehat{a}_j$ *that satisfies Eq.* 7. *Let us first show that* $\theta_i^*(\widehat{a}_j \mid \overline{a}_i) = 0$. *Suppose instead that* $\theta_i^*(\widehat{a}_j \mid \overline{a}_i) = \zeta > 0$. *We are also given a transfer inclusive most preferred action* a_j^*. *It follows that* $\theta_i^*(\widehat{a}_j \mid \overline{a}_i) > \zeta$. *Consider an alternative transfer schedule* $\theta_i^{'}(.)$ *that reduces the transfers at* $\widehat{a}_j$ *and* a_j^* *by* ζ *and reduces transfers* *at other actions such that the lifetime payoffs everywhere is the same as at $\widehat{a}_j$, i.e.,*

$$(1-\delta)\theta_i^{'}(a_j \mid \overline{a}_i) = (1-\delta)[\pi_j(\overline{a}_i, \widehat{a}_j) - \pi_j(\overline{a}_i, a_j)] + \delta[w_j(\widehat{a}_j \mid \overline{a}_i) - w_j(a_j \mid \overline{a}_i)] \quad (9)$$

It is not difficult to see that in this transfer scheme a_j^* continues to be the most preferred action and yet it costs the giver, player i, strictly less. Hence the contradiction proves that $\theta_i^*(\widehat{a}_j \mid \overline{a}_i) = 0$. The same argument also establishes that the transfer at a_j^* must be such that player j is just indifferent between a_j^* and $\widehat{a}_j$. Hence the first two lines of Eq. 7 have been established. That a_j^* must be derived from a maximization problem follows, for example, by considering a transfer scheme that satisfies Eq. 9. The fact that, from Eq. 6 we know that $w_i = (1-\delta)[\pi_i(\overline{a}_i, a_j^*) - \theta_i^*(a_j^* \mid \overline{a}_i)] + \delta v_i(a_j^* \mid \overline{a}_i) \geqslant (1-\delta)[\pi_i(\overline{a}_i, a_j^{'}) - \theta_i^{'}(a_j^{'})]$ $+ \delta v_i(a_j^{'} \mid \overline{a}_i)$ is just a re-statement of the penultimate equation in Eq. 7. The lemma is proved.■

The operator T above can be simplified further. It is clear from the definition of v_j above that there is an equivalent way of writing the transfer constraint - the last line in Eq. 3:

$$(1-\delta)\theta_i(a_j) = (1-\delta)[\pi_j(\overline{a}_i, \widehat{a}_j) - \pi_j(\overline{a}_i, a_j)] + \delta[w_j(\widehat{a}_j) - w_j(a_j)]$$

Substituting that into the definition of w_i, the second-last line in Eq. 3, we have

$$w_i = \max_{a_j}\{(1-\delta)[\pi_i(\overline{a}_i, a_j) + \pi_j(\overline{a}_i, a_j)] + \delta[v_i(a_j) + w_j(a_j)]\} - v_j$$

This allows a more compact definition of the operator T with the transfer terms completely removed:

$$\begin{aligned} T(V_i, W_j)(\overline{a}_i) &= \{w_i, v_j : \exists \widehat{a}_j, v_i(a_j), w_j(a_j) \in V_i, W_j(a_j) : \\ \widehat{a}_j &\in \arg\max_{a_j}\{(1-\delta)\pi_j(\overline{a}_i, a_j) + \delta w_j(a_j)\} \\ v_j &= (1-\delta)\pi_j(\overline{a}_i, \widehat{a}_j) + \delta w_j(\widehat{a}_j) \\ w_i &= \max_{a_j}\{(1-\delta)[\pi_i(\overline{a}_i, a_j) + \pi_j(\overline{a}_i, a_j)] + \delta[v_i(a_j) + w_j(a_j)]\} - v_j \end{aligned} \quad (10)$$

Similarly, define the operator $T(V_j, W_i)(\overline{a}_j)$ by transposing the indices i and j. We will now establish some properties of the operator T.

Lemma 4 *If V_i, W_j, V_j, W_i are non-empty compact-valued correspondences then so are $T(V_i, W_j)$ and $T(V_j, W_i)$. Furthermore, they are monotone, i.e., $T(V_i, W_j) \supseteq T(V_i^{'}, W_j^{'})$ if $(V_i, W_j) \supseteq (V_i^{'}, W_j^{'})$. Similarly for $T(V_j, W_i)$.*

Proof: Follows from the definitions.■

3.1.2 Fixed Point of Operator and a Uniqueness Property

Let a sequence of correspondences be defined as follows:

Definition 5 *Consider the set of all feasible payoffs in the stage game G inclusive of transfers. Call that set of pairs, one payoff for each player, F. Define*

$$V_i^0, W_j^0 = F = V_j^0, W_i^0$$

Then, inductively define

$$\begin{aligned} V_i^{n+1}, W_j^{n+1} &= T(V_j^n, W_i^n) \\ V_j^{n+1}, W_i^{n+1} &= T(V_i^n, W_j^n) \end{aligned} \tag{11}$$

We now show, by a standard argument that uses the lemma above, that the sequence has a fixed point:

Lemma 6 *The pairs V_j^n, W_i^n and V_i^n, W_j^n converge to a non-empty compact-valued correspondence as $n \to \infty$. Call the limit correspondence V_j^∞, W_i^∞ and V_i^∞, W_j^∞.*

Proof. *From the fact that the operator preserves compact-valuedness and non-emptiness it follows that each correspondence has that property along the sequence. From the monotonicity property of the operator it follows that $V_j^n, W_i^n(a_j)$ and $V_i^n, W_j^n(a_i)$ are monotonically decreasing non-empty compact sets at each a_i and a_j and hence that the limit correspondence V_j^∞, W_i^∞ and V_i^∞, W_j^∞ are non-empty compact-valued correspondences.* ■

The next Lemma shows that $V_j^\infty, W_i^\infty(a_i)$ and $V_i^\infty, W_j^\infty(a_j)$ are singlentons, for all (a_i, a_j).

Lemma 7 *The correspondences V_2^∞, W_1^∞ and V_1^∞, W_2^∞ are single valued for all $a_1 \in A_1$ and $a_2 \in A_2$.*

Proof By the definition of the operator T and of the correspondences V_j^∞, W_i^∞, $i \neq j$, $i = 1, 2$, a vector

$$(\bar{v}_{20}, \bar{w}_{10}; \bar{v}_{10}, \bar{w}_{20}) \equiv \{[\bar{v}_{20}(a_1), \bar{w}_{10}(a_1)]_{a_1 \in A_1}; [\bar{v}_{10}(a_2), \bar{w}_{20}(a_2)]_{a_2 \in A_2}\}$$

is an element of the product correspondence

$$\times_{a_1 \in A_1} V_2^\infty, W_1^\infty(a_1) \times \{\times_{a_{2 \in A_2}} V_1^\infty, W_2^\infty(a_2)\}$$

if there exists a sequence

$$\{v_{2t}, w_{1t}; v_{1t}, w_{2t}\}_{t \geq 0} \equiv \{[v_{2t}(a_1), w_{1t}(a_1)]_{a_1 \in A_1}; [v_{1t}(a_2), w_{2t}(a_2)]_{a_2 \in A_2}\}_{t \geq 0}$$

with $(v_{20}, w_{10}; v_{10}, w_{20}) = (\bar{v}_{20}, \bar{w}_{10}; \bar{v}_{10}, \bar{w}_{20})$ such that $v_{it}(a_j), w_{jt}(a_j) \in (V_i^\infty, W_j^\infty)(a_j)$, for all t and a_j and i and such that the following system of equations holds true for all $i = 1, 2$ and $i \neq j$:

$$v_{it}(a_j) = \max_{a_i} \{(1-\delta)\pi_i(a_i, a_j) + \delta w_{it+1}(a_i)\}, \text{ all } a_j \tag{12}$$

$$(w_{it}+v_{jt})(a_i) = \max_{a_j}\{(1-\delta)[\pi_i(\overline{a}_i, a_j)+\pi_j(\overline{a}_i, a_j)]+\delta(w_{it+1}+v_{jt+1})(a_j)\}, \text{ all } a_i$$

The last equations imply that $(w_{it}+v_{jt})(a_i)$ is the value of the Pareto problem 5 at a_i which we denote by $r(a_i)$. Therefore, the sequences $\{[v_{1t}(a_1), w_{2t}(a_1)]_{a_1\in A_1}; [v_{2t}(a_2), w_{1t}(a_2)]_{a_2\in A_2}\}_{t\geq 0}$ must be a solution to equations 12 and to

$$v_{it}(a_j) + w_{jt}(s_j) = r(s_2), \text{ all } a_j \text{ and } i. \tag{13}$$

We rewrite equations 12 and 13 in a more convenient form. Solving equations 13 for $w_{2t}(a_2)$ when $i = 1$ and for $v_{2t}(a_1)$ when $i = 2$ and them substitute those values into equations 12 when $i = 2$ obtaining:

$$r(a_1) - w_{1t}(a_1) = \max_{a_2}\{(1-\delta)\pi_2(a_1, a_2) + \delta[r(a_2) - v_{1t+1}(a_1)\}, \text{ all } a_1,$$

or equivalently

$$w_{it}(a_i) = \min_{a_j}\{r(a_i, a_j) + \delta v_{it+1}(a_j)\}, \text{ all } a_i. \tag{14}$$

for $r(a_1, a_2) = r(a_1) - (1-\delta)\pi_2(a_1, a_2) - \delta r(a_2)$.Therefore in order to prove that the correspondences V_2^∞, W_1^∞ and V_1^∞, W_2^∞ are single valued, it is sufficient to show that there exists a unique and time invariant sequence $\{v_{1t}, w_{1t}\}_{t\geq 0}$, with $(v_{1t}(a_1), v_{1t}(a_2)) \in V_1^\infty(a_2) \times W_1^\infty(a_1)$, for all t and (a_1, a_2) solving equations 12 and 14. This is shown in the next claim.

Claim 8 *: The solution to equations 12 and 14 is unique and time invariant.*

Proof First we show that $\{v_{1t}, w_{1t}\}_{t\geq 0}$ is unique. Then that uniqueness implies time invariance.

(Proof 1 - Contraction): Define the following two operators:

$$U_1 w(a_j) = \max_{a_i}\{(1-\delta)\pi_i(a_i, a_j) + \delta w(a_i)\} \tag{15}$$

and

$$U_2 v(a_i) = \min_{a_j}\{r(a_i, a_j) + \delta v(a_j)\} \tag{16}$$

wherer $w : A_i \longrightarrow \mathbb{R}$ and $v : A_j \longrightarrow \mathbb{R}$ (and hence $U_1 w : A_j \longrightarrow \mathbb{R}$ while $U_2 v : A_i \longrightarrow \mathbb{R}$). We shall show that both operators are contractions and hence have a unique fixed point - and that will prove the claim. The proof will follow Blackwell (1955) and is repeated here only because the minimization in Eq. 16 may appear non-standard to some readers who are familiar with Blackwell's proof. It is straightforward to see that

$$U_1(w+k)(a_j) = U_1 w(a_j) + \delta k, U_2(v+k)(a_i) = U_2 v(a_i) + \delta k \tag{17}$$

for any constant k. Equally clearly, if $w' \geqslant w$ then $U_1 w' \geqslant U_1 w$ while if $v' \leqslant v$ then $U_2 v' \leqslant U_2 v$. For any two functions w and w', denote the supnorm $\parallel w - w' \parallel = \sup_{a_i} \mid w(a_i) - w'(a_i) \mid$ (and likewise $\parallel v - v' \parallel = \sup_{a_j}$

$\mid v(a_j) - v'(a_j) \mid$. For any two pairs of functions (w, v) and (w', v') denote the supnorm as: $\parallel (v, w) - (v', w') \parallel = \max\{\parallel w - w' \parallel, \parallel v - v' \parallel\}$.

We will now show that

$$\parallel (U_2 v, U_1 w) - (U_2 v', U_1 w') \parallel \leqslant \delta \parallel (v, w) - (v', w') \parallel \tag{18}$$

Since it is always the case that $w' + \parallel w - w' \parallel \geqslant w$ (and similarly, $w + \parallel w - w' \parallel \geqslant w'$), from the monotonicity of the operator U_1 and Eq. 17 it follows that

$$\delta \parallel w - w' \parallel \geqslant \mid U_1 w(a_j) - U_1 w'(a_j) \mid \ \ \forall a_j$$

and hence, $\delta \parallel w - w' \parallel \geqslant \parallel U_1 w - U_1 w' \parallel$. Similarly, since it is always the case that $v' - \parallel v - v' \parallel \leqslant v$ (and $v - \parallel v - v' \parallel \leqslant v'$), the monotonicity of the operator U_2 implies that

$$\delta \parallel v - v' \parallel \geqslant \mid U_2 v(a_i) - U_2 v'(a_i) \mid \ \ \forall a_i$$

and hence, $\delta \parallel v - v' \parallel \geqslant \parallel U_2 v - U_2 v' \parallel$. The last two norm inequalities clearly prove Eq. 18. Hence, the operators are contractions and consequently have a unique fixed point. The claim is proved.

(Proof 2) By contradiction let $\{v_{1t}^m, w_{1t}^m\}_{t \geq 0}$, $m = 1, 2$, be two distinct sequences solving equations 12 and 14. Therefore, the sequence $\{v_{1t}^{m^*} - v_{1t}^{m'}\}_{t \geq 0}$, for $m^* \neq m'$, solves

$$\begin{aligned} v_{1t}^{m^*}(a_2) - v_{1t}^{m'}(a_2) \quad &= \quad \max_{a_1}\{(1-\delta)\pi(a_1, a_2) + \delta w_{1t+1}^{m^*}(a_1)\} \\ &\quad - \max_{a_1}\{(1-\delta)\pi(a_1, a_2) + \delta w_{1t+1}^{m'}(a_1)\}\}, \text{ all } a_2 \end{aligned}$$

Since for any two functions f and g, it is $\max_x f(x) - \max_x g(x) \leq \max_x f(x) - g(x)$, it is:

$$v_{1t}^{m^*}(a_2) - v_{1t}^{m'}(a_2) \leq \delta\{w_{1t+1}^{m^*}(a_1^*) - w_{1t+1}^{m'}(a_1^*)\}, \text{ all } a_2, \tag{19}$$

where $a_1^* \in \arg\{\max_{a_1}\{w_{1t+1}^{m^*}(a_1^*) - w_{1t+1}^{m'}(a_1^*)\}$. Then, by equation 14, inequalities 19 can be rewritten as:

$$v_{1t}^{m^*}(a_2) - v_{1t}^{m'}(a_2) \leq \delta\{\min_{a_2'}\{r(a_1^*, a_2) + \delta v_{1t+2}^{m^*}(a_2')\} - \min_{a_2'}\{r(a_1^*, a_2') + \delta v_{1t+2}^{m'}(a_2')\}\}, \text{ all } a_2,$$

or, for $a_{2,t+2}^m \in \arg\min_{a_2'}\{r(a_1^*, a_2') + \delta v_{1t+2}^m(a_2')\}\}$, $m = m^*, m'$, as

$$v_{1t}^{m^*}(a_2) - v_{1t}^{m'}(a_2) \leq \delta\{r(a_1^*, a_{2,t+2}^{m^*}) + \delta v_{1t+2}^{m^*}(a_{2,t+2}^{m^*}) - r(a_1^*, a_{2,t+2}^{m'}) + \delta v_{1t+2}^{m'}(a_{2,t+m}^{m'})\}, \text{ all } a_2$$

As by definition, $r(a_1^*, a_{2,t+2}^{m^*}) + v_{t+2}^{m^*}(a_{2,t+2}^{m^*}) \leq r(a_1^*, a_{2,t+2}^{m'}) + v_{t+2}^{m^*}(a_{2,t+2}^{m'})$, the last inequalities imply the following:

$$v_{1t}^{m^*}(a_2) - v_{1t}^{m'}(a_2) \leq \delta^2\{v_{1t+2}^{m^*}(a_{2,t+2}^{m'}) - v_{1t+2}^{m'}(a_{2,t+2}^{m'})\}, \text{ all } a_2.$$

Then arguing recursively, for all a_2:

$$\begin{aligned} v_{1t}^{m^*}(a_2) - v_{1t}^{m'}(a_2) &\leq \delta^2\{v_{1t+2}^{m^*}(a_{2,t+2}^{m'}) - v_{1t+2}^{m'}(a_{2,t+2}^{m'})\} \\ &\leq \lim_{n\to\infty} \delta^{2n}\{v_{1t+2n}^{m^*}(a_{2,t+2n}^{m'}) - v_{1t+2n}^{m'}(a_{2,t+2n}^{m'})\} \text{ all } a_2. \end{aligned}$$

Since V_1^{∞} is a compact valued correspondence, the values $v_t^m(a_2) - v_t^m(a_2)$ are uniformly bounded. Then, $\delta < 1$ implies that $\lim_{n\to\infty} \delta^{2n}\{v_{1t+2n}^{m^*}(a_2) - v_{1t+2n}^{m'}(a_2)\} = 0$, all a_2. As m^* and m' are arbitrary, the latter implies $\{v_{1t}^1(a_2)\}_{t\geq 0} = \{v_{1t}^2(a_2)\}_{t\geq 0}$, for all a_2, establishing uniqueness of the solution. Moreover, if time invariance would fail there would be $t' > t^*$ such that $(v_{1t^*})(a_2) \neq (v_{1t'})(a_2)$ for some $a_2 \in A_2$. Then the two distinct sequences $\{w_{1t+t^*}, v_{1t+t^*}\}_{t\geq 0}$ and $\{w_{1t+t'}, v_{1t+t'}\}_{t\geq 0}$ would be solutions to equations 12 and 14 contradicting the first part of the argument. ■

By Lemma 8, SPE value sequences are time invariant, that is, they depend on the state (or action) a_i or a_j. If the solutions to the planner problem 5 are unique for all initial states $\bar{a}_i$ or $\bar{a}_j$, SPE action and transfer profiles are unique with SPE actions identical to the planners optimal actions. However, if for some $\bar{a}_i$ or $\bar{a}_j$, there are multiple solutions to the planner problem, there are multiple action and transfer SPE profiles (with actions identical to the planner's actions) all of them obviously payoff equivalent.

3.1.3 Equivalence Between Fixed Point and the SPE Value Set

Let V_j^*, W_i^* denote the SPE equilibrium value correspondence when the initial mover is player j (with associated SPE payoffs $v_j^*, w_i^* \in V_j^*, W_i^*$). Similarly, let V_i^*, W_j^* denote the SPE equilibrium value correspondence when the initial mover is player i. In compact notation, suppressing player subscripts, let us write that quartet as V^*, W^*.

The argument that this SPE value correspondence is nothing but the fixed point V^{∞}, W^{∞} computed above will draw on the argument in Abreu-Pearce-Stachetti for discounted Repeated Games (see Abreu-Pearce-Stachetti (1986)) which in turn is a strategic version of the Bellman Principle for Dynamic programming. The **first step** will show that the SPE value correspondence V^*, W^* is "self-generated" (in APS terminology) in that $V^*, W^* \subset T(V^*, W^*)$. (This is the game-theoretic analog to the (Necessity Part of the) Optimality Principle in Dynamic Programming, the assertion that the value function must be bounded above by the Bellman operator applied to itself.) The **second step** will be an appeal to the monotonicity property of the operator to show that $V^*, W^* \subset V^n, W^n$ for all n and hence that $V^*, W^* \subset V^{\infty}, W^{\infty}$. The **third step** - and the final one - will follow from the fact that we have already shown - see Lemma 6 above - that V^{∞}, W^{∞} is a singleton; hence, it must be the case that $V^*, W^* = V^{\infty}, W^{\infty}$. That will evidently complete the proof.

The first step hence is:

Lemma 9 *For the SPE equilibrium value correspondence, it is the case that*

$$\begin{aligned} V_i^*, W_j^* &\subset T(V_j^*, W_i^*) \\ V_j^*, W_i^* &\subset T(V_i^*, W_j^*) \end{aligned}$$

Proof. Suppose that $v_j^*, w_i^* \in V_j^*, W_i^*(\overline{a}_i)$, i.e., when player i contemplates an action $\overline{a}_i$, these are SPE payoffs that arise based on equilibrium strategies - say, $\sigma_i^*(\overline{a}_i)$ and $\sigma_j^*(\overline{a}_i)$.[7] In particular, the pair of best response strategies imply a concurrent transfer schedule $\theta_i^*(.)$ (picked by i) and an action a_j^* (picked by j) and continuation payoffs $w_j(a_j \mid \overline{a}_i)$. Since this is a SPE, the continuation payoffs do not depend on the t period transfer scheme $\theta_i^*(.)$ (since that is not part of the public history at time $t+1$, the initial period for the continuation $w_j(a_j \mid \overline{a}_i)$).

Let us detail each player's incentives more carefully starting with player j:

(IC - Player j): The action a_j^* maximizes player $j's$ lifetime payoffs

$$a_j^* \in \arg\max\{(1-\delta)[\pi_j(\overline{a}_i, a_j) + \theta_i^*(a_j)] + \delta w_j(a_j \mid \overline{a}_i)\} \tag{22}$$

(IC - Player i): There must not exist an alternative strategy for player i, $\sigma_i^{'}(\overline{a}_i)$, with, say, an immediate transfer schedule $\theta_i^{'}(.)$, and a best response public strategy of player j, $\sigma_j^{'}(\overline{a}_i; \theta_i^{'}(.))$, with continuation payoffs $w_j(a_j \mid \overline{a}_i) \in W_j^*$, so that an action $a_j^{'}$ is incentive compatible for player j, i.e.,

a) $a_j^{'}$ solves Eq. 22 for $\theta_i^{'}(.)$ and $w_j(a_j \mid \overline{a}_i)$,and that

b) player i prefers it because it gives her higher lifetime payoffs, i.e.,

$$(1-\delta)[\pi_i(\overline{a}_i, a_j^{'}) - \theta_i^{'}(a_j^{'})] + \delta v_i(a_j^{'} \mid \overline{a}_i) > (1-\delta)[\pi_j(\overline{a}_i, a_j^*) - \theta_i^*(a_j^*)] + \delta v_i(a_j^* \mid \overline{a}_i) \tag{23}$$

Clearly, from the definition it follows that $v_j^*, w_i^* \in \Gamma(V_j^*, W_i^*)(\overline{a}_i)$. However, we have already seen that $\Gamma(V_j^*, W_i^*)(\overline{a}_i) = T(V_j^*, W_i^*)(\overline{a}_i)$. Hence, the lemma has been proved. A swap of player subscripts proves similarly that $v_j^*, w_i^* \in T(V_i^*, W_j^*)$. ■

From Lemma 4 above, it follows that we have $V^*, W^* \subset T(V^n, W^n) = V^{n+1}, W^{n+1}$ for all n and hence that $V^*, W^* \subset V^\infty, W^\infty$. Hence, Step 2 is proved. Finally, from Lemma 8 above we have the unique valuedness of the correspondence V^∞, W^∞. That then implies that one of two things must be true: a) the SPE value correspondence $V^*, W^* \to \mathbb{R}$ is empty-valued. Or, b) that it is non-empty-valued and is in fact identical to the function V^∞, W^∞. That the former is not possible, that there always exists a SPE, follows - for example - from Harris (1985) who showed that SPE always exist in perfect information games with compact histories and continuous payoffs.

All three steps have been proved and we have shown that the SPE value correspondence V^*, W^* agrees with the fixed point of the operator V^∞, W^∞.

[7] Note that if this is play commencing after a history, say h_t, then the strategies $\sigma_i^*(\overline{a}_i)$ and $\sigma_j^*(\overline{a}_i)$ might be indexed by the history and should be written more completely as $\sigma_i^*(\overline{a}_i; h_t)$ and $\sigma_j^*(\overline{a}_i; h_t)$. Since the *set* of SPE is independent of history, and we are investigating a generic SPE, it saves notation to drop h_t in what follows.

3.2 Completing the Proof for Finite Transfer Histories

In this section we show that the result extends to the case of any $M < \infty$.

Lemma 10 *The conclusion that there is a unique SPE which coincides with the Utilitarian maximum holds for all* $1 < M < \infty$.

Proof. Recall that a *history* h_t at time t is given by $h_t = (h_t^a, h_t^\theta)$ where the finite transfer history is $h_t^\theta = (\theta_{jt-M}(.), \theta_{it-M+1}(.), ...\theta_{jt-1}(.))$. Recall also the "Stackelberg leader" player $i's$ optimal transfer problem, i.e.,

$$w_i(h_t, a_{it}) = \max_{\theta_i(.)}\{(1-\delta)[\pi_i(a_{it}, a_j(\theta_i)) - \theta_i(a_j)] + \delta v_i(h_{t+2}(a_{it}, \theta_{it}))\} \quad (24)$$

where $h_{t+2}(a_{it}, \theta_{it})$ is the history at period $t+2$ caused by player $i's$ period t actions, $h_{t+2}(a_{it}, \theta_{it}) = (h_t^a, a_{it}, a_{jt+1}(\theta_{it}); \theta_{jt-M+2}(.), ...\theta_{it}(.), \theta_{jt+1}(.))$. In that history, the follower player $j's$ best response action in period $t+1$, $a_j(\theta_{it})$, comes from:[8]

$$v_j(h_{t+1}(a_{it}, \theta_{it})) = \max_{a_j}\{[\pi_j(a_{it}, a_j) + \theta_{it}(a_j)] + \delta w_j(h_{t+1}, a_j)\} \quad (25)$$

where, similarly, h_{t+1} is the history at period $t+1$ caused by player $i's$ period t actions, $h_{t+1}(a_{it}, \theta_{it}) = (h_t^a, a_{it}; \theta_{jt-M+1}(.), ...\theta_{it}(.))$. Hence, from Eq. 25 it is clear that $a_j(\theta_{it})$ does not contain any reference to $\theta_{jt-M}(.)$ since, by period $t+1$ when that action is chosen, $\theta_{jt-M}(.)$ is no longer in the finite transfer memory. In particular, $w_i(h_t, a_{it})$ is then independent of $\theta_{jt-M}(.)$. Since $w_i(h_t, a_{it})$ is independent of $\theta_{jt-M}(.)$ it then follows, from the definition of player $i's$ lifetime payoffs starting at time t,

$$v_i(h_t) = \max_{a_i}\{(1-\delta)[\pi_i(a_i, a_{jt-1}) + \theta_{jt-1}(a_i)] + \delta w_i(h_t, a_i)\} \quad (26)$$

that $v_i(h_t)$ is also independent of $\theta_{jt-M}(.)$. By the same logic, the payoffs for player j starting in the subsequent period $t+1$, $v_j(h_{t+1}(a_{it}, \theta_{it}))$, must be independent of $\theta_{it-M+1}(.)$. Then, from Eq. 25, it follows that player $j's$ best response action in period $t+1$, $a_j(\theta_{it})$, which is the solution of Eq. 25, must itself be independent of $\theta_{it-M+1}(.)$. But, invoking Eqs. 24 and 26, that implies that $w_i(h_t, a_{it})$ and $v_i(h_t)$ must both be independent of $\theta_{it-M+1}(.)$.

At this point the argument in the previous paragraph can be repeated to show that if $w_i(h_t, a_{it})$ and $v_i(h_t)$ are independent of $\theta_{it-M+1}(.)$, then $v_j(h_{t+1}(a_{it}, \theta_{it}))$ and $a_j(\theta_{it})$ must be independent of $\theta_{it-M+2}(.)$ thereby implying that $w_i(h_t, a_{it})$ and $v_i(h_t)$ are in fact independent of $\theta_{it-M+2}(.)$. And so on. Leading finally to the conclusion that no part of the transfer history enters w_i, i.e., it is a function of a_{it} alone and that only $\theta_{jt-1}(.)$ enters v_i - and only through the payoff relevant effect in the first term of Eq. 26. Put another way, the general case of any M has been reduced to $M = 1$. The proof of the previous sub-section then applies. ■

[8] Note that $h_{t+1}(a_{it}, \theta_{it}) = (h_t^a, a_{it};\ \theta_{jt-M+1}(.), ...\ \theta_{it}(.))$.

4 A General Result for the Finite Horizon

The theorem that we proved in the previous section applies to a finite horizon game as well.[9] Indeed, in many ways, the argument is simpler because one can rely on the stick of backwards induction. Also, the result does not require a restriction to finite transfer histories.

We now prove the relevant result. For purely notational reasons, suppose that the last player to move is player j. For any a_i, define the one-period utilitarian maximum:

$$\widehat{\sigma}_1(a_i) = \arg\max_{a_j}[\pi_i(a_i, a_j) + \pi_j(a_i, a_j)]$$

and let the associated one-period utilitarian maximum value be denoted

$$U_1(a_i) = \pi_i(a_i, \widehat{\sigma}_1(a_i)) + \pi_j(a_i, \widehat{\sigma}_1(a_i))$$

Inductively, whenever player j is the mover, define the $T+1$ period utilitarian maximum

$$\widehat{\sigma}_{T+1}(a_i) = \arg\max_{a_j}\{[\pi_i(a_i, a_j) + \pi_j(a_i, a_j)] + U_T(a_i)\}$$

and let the associated $T+1$-period utilitarian maximum value be denoted

$$U_{T+1}(a_i) = \pi_i(a_i, \widehat{\sigma}_{T+1}(a_i)) + \pi_j(a_i, \widehat{\sigma}_{T+1}(a_i)) + U_T(\widehat{\sigma}_{T+1}(a_i))$$

with analogous optimal action, $\widehat{\sigma}_{T+1}(a_j)$ and utilitarian value, $U_{T+1}(a_j)$ defined for periods in which player i is the mover.

Define now individual incentives in the following manner. For any a_i define the one-period unilateral maximum:

$$\sigma_1^*(a_i) = \arg\max_{a_j} \pi_j(a_i, a_j)$$

and let the associated one-period unilateral maximum be denoted

$$v_1^*(a_i) = \pi_j(a_i, \sigma_1^*(a_i))$$

For every action a_i define the transfer required to induce one-period play of a_j as

$$\theta_1(a_j \mid a_i) = v_1^*(a_i) - \pi_j(a_i, a_j) \tag{27}$$

Clearly, $\theta_1(a_j \mid a_i) \geqslant 0$ for all a_j. Define the non-mover's, player $i's$, payoff in the last period as

$$W_1(a_i) = \max_{a_j}[\pi_i(a_i, a_j) - \theta_1(a_j \mid a_i)]$$

[9] The result applies both to the discounted as well as the undiscounted model. For notational simplicity, we restrict ourselves to the undiscounted case.

Substituting for $\theta_1(a')$ from Eq. 27 it follows that the maximization above is equivalent to

$$W_1(a_i) = \max_{a_j}[\pi_i(a_i, a_j) + \pi_j(a_i, a_j)] - v_1^*(a_i) = U_1(a_i) - v_1^*(a_i)$$

so that we have the following claim:

Claim 11 *When there is one period left in the game, regardless of the action a_i played in the immediately preceding period by (the non-mover) player i, the unique equilibrium continuation is for i to offer a transfer $\theta_1(\widehat{\sigma}_1(a_i) \mid a_i)$ to induce the play of the utilitarian maximum $\widehat{\sigma}_1(a_i)$ in the last period. The associated payoffs for the mover, player j, and the non-mover are then, respectively, the unilateral maximum value $v_1^*(a_i)$ and the residual from the utilitarian maximum value $U_1(a_i) - v_1^*(a_i)$.*

Similarly when there are 2 periods left, for any a_j define the two-period unilateral maximum:

$$\sigma_2^*(a_j) = \arg\max_{a_i}[\pi_i(a_i, a_j) + W_1(a_i)]$$

and let the associated two-period unilateral maximum be denoted

$$v_2^*(a_j) = \pi_i(\sigma_2^*(a_j), a_j) + W_1(\sigma_2^*(a_j))$$

For every action a_i define the transfer required to induce penultimate-period play of a_i as

$$\theta_2(a_i) = v_2^*(a_j) - [\pi_i(a_i, a_j) + W_1(a_i)] \tag{28}$$

Clearly, $\theta_2(a_i) \geqslant 0$ for all a_i. Define the non-mover's, player $j's$, continuation payoff starting in the penultimate period as

$$W_2(a_j) = \max_{a_i}[\pi_j(a_i, a_j) - \theta_2(a_i) + v_1^*(a_i)]$$

Substituting for $\theta_1(a')$ from Eq. 28 it follows that the maximization above is equivalent to

$$W_2(a_j) = \max_{a_i}[\pi_i(a_i, a_j) + \pi_j(a_i, a_j) + W_1(a_i) + v_1^*(a_i)] - v_2^*(a_j) = U_2(a_j) - v_2^*(a_j)$$

so that we have the following claim:

Claim 12 *When there are two periods left in the game, regardless of the action a_j played in the third to last period by current non-mover j, the unique equilibrium continuation in the penultimate period is for to offer a transfer $\theta_2(\widehat{\sigma}_2(a_j) \mid a_j)$ to induce the play by i of the utilitarian maximum $\widehat{\sigma}_2(a_j)$ in the second to last period. The associated payoffs in the last two periods for the mover and non-mover are then, respectively, the unilateral maximum value $v_2^*(a_j)$ and the residual from the utilitarian maximum value $U_2(a_j) - v_2^*(a_j)$.*

Inductively define, for periods when player i is the mover, optimal action, $\widehat{\sigma}_{T+1}(a_j)$ and utilitarian value, $U_{T+1}(a_j)$. Similarly, for periods when player j is the mover, define an optimal action, $\widehat{\sigma}_{T+1}(a_i)$ and utilitarian value, $U_{T+1}(a_i)$. The same proof as above applies to give us the main result which is:

Theorem 13 *Consider a T period game. The utilitarian solution $\{\widehat{\sigma}_t(a) : t = 1, ...T\}$ is the unique SPE action profile in the game. That is supported by transfers that induce the mover to pick that action by way of a compensating transfer that makes her as well off as she would be from her most preferred transfer.*

5 An Example with Infinite Transfer History

If infinite transfer histories are allowed, whilst equilibria are still Pareto optimal, uniqueness of the transfer scheme need no longer hold. The example below establishes that.

Consider the following stage game

1\2	A	B
*	(0,0)	(1,0)

An action is $a \in \{A, B\}$. We use the following convention: a) at T even, $T = 0, 2, \ldots$, player 2 selects an action a_T given the history $h_T = (\theta_0, a_1, \ldots, a_{T-1}, \theta_T)_{t \leq T}$; at T odd player 2 selects a transfer θ_{T+1} given the history $h_T = (\theta_0, a_1, \ldots, \theta_{T-1}, a_T)$. The set of histories is $\mathbb{H}$.

We limit attention to strategies where when player 2 is not a mover, he sets $\theta = 0$, and we never record this entry in the histories. It should be clear from the argument that this is without loss of generality. Similarly we never record the single action of player 1. When player 2 is a mover at h_T, the last entry of h_T is θ_T. When player 2 is not a mover at h_T, the last entry is a_T, the last action taken by 2. When player 1 is either not a mover or a mover at h_T, the last entry is a_T, as 1 does not choose actions.

For given strategies of the other players, the lifetime payoffs for the players are:

$$\begin{aligned} W_1(h_T) &= \max_{\theta}(1-\delta)(\pi_1(a_{T+1}(h_T,\theta)) - \theta(a_{T+1}(h_T,\theta)) + \delta V_1(h_T, \theta, a_{T+1}(h_T,\theta)) \\ V_1(h_T) &= (1-\delta)\pi_1(a_T) + \delta W_1(h_T) \end{aligned}$$

and

$$\begin{aligned} W_2(h_T) &= \delta V_2(h_T) \\ V_2(h_T) &= \max_{a}(1-\delta)(\theta_T(a)) + \delta W_2(h_T, a) \end{aligned}$$

Our equilibrium strategies depend on two subsets of $\mathbb{H}$ and two parameters, α and θ', that we define below.

Let $\mathbb{H}_+$ and $\mathbb{H}_-$, $\mathbb{H}_- \subset \mathbb{H}_+$, be two subsets of the set of histories defined as follows:

Definition of $\mathbb{H}_+$: $h_T \in \mathbb{H}_+$ if there exists t such that a) either $\theta_t(B) - \theta_t(A) < \alpha$ or $a_t \neq B$; and b) $\theta_{t'}(B) - \theta_{t'}(A) < \theta'$, all $t \leq t' \leq T$.

Definition of $\mathbb{H}_-$: $h_T \in \mathbb{H}_-$ if $h_T \in \mathbb{H}_+$ and there exists $t < T$ such that $h_T = (h_t, h_{T-t})$ with $h_t \in \mathbb{H}_+$ and $a_{t+1} = B$.

Let α be a positive number to be specified below and let

$$\theta' = \frac{1+\delta^2}{\delta^2}\alpha. \tag{29}$$

Later, we define α so that the following are equilibrium strategies:

Equilibrium strategies:
Player 2: $a(h_T) = B$, if $h_T \in \mathbb{H}\backslash\mathbb{H}_+$ or $h_T \in \mathbb{H}_-$ and $\theta_T(A) \leq \theta_T(B)$. While $a(h_T) = A$, otherwise (that is, $h_T \in \mathbb{H}_+\backslash\mathbb{H}_-$, or $h_T \in \mathbb{H}_-$ and $\theta_T(A) > \theta_T(B)$).

Player 1: $\theta(h_T) = (\theta(A), \theta(B)) = (0, \alpha)$, if $h_T \in \mathbb{H}\backslash\mathbb{H}_+$; $\theta(h_T) = (0, \theta')$, $h_T \in \mathbb{H}_+\backslash\mathbb{H}_-$, $\theta(h_T) = (0,0)$, $h_T \in \mathbb{H}_-$.

5.1 Equilibrium payoffs

5.1.1 Player 1

Histories in $\mathbb{H}\backslash\mathbb{H}_+$.

Let $h_T \in \mathbb{H}\backslash\mathbb{H}_+$ be such that $a_T = B$. By definition of the equilibrium strategies, $\theta(h_T) = \theta^* = (0, \alpha)$, while, as $(h_T, \theta^*) \in H\backslash H+$, $a(h_T, \theta^*) = B$. Hence $W_1(h_T) = W_1(\bar{h}_{\bar{T}})$, and $V_1(h_T, (0,\alpha), B) = V_1(\bar{h}_{\bar{T}}, (0,\alpha), B)$, for all histories h_T and $\bar{h}_{\bar{T}} \in \mathbb{H}\backslash\mathbb{H}_+$ with $a_T = \bar{a}_{\bar{T}} = B$. We denote by $W_1(+)$ and $V_1(B)$ such values.

Then:

$$\begin{aligned} W_1(+) &= (1-\delta)(1-\alpha) + \delta V_1(B), \\ V_1(B) &= (1-\delta) + \delta W_1(+) \end{aligned}$$

and therefore

$$W_1(+) = \frac{1-\alpha+\delta}{1+\delta} \tag{30}$$

$$V_1(B) = \frac{1 + \delta(1-\alpha)}{1+\delta} \tag{31}$$

Histories in $\mathbb{H}_+ \backslash \mathbb{H}_-$

Let $h_T \in \mathbb{H}_+ \backslash \mathbb{H}_-$. Then $\theta(h_T) = \hat{\theta} = (0, \theta')$ while, as $h_T, \hat{\theta} \in \mathbb{H} \backslash \mathbb{H}_+$, it is $a_T(h_T, \hat{\theta}) = B$ and $(h_T, \hat{\theta}) \in \mathbb{H} \backslash \mathbb{H}_+$. Thus, the payoff $W_1(h_T)$ is invariant for all histories $h_T \in \mathbb{H}_+ \backslash \mathbb{H}_-$ and it is denoted by $W_1(-)$. Then:

$$W_1(-) = (1-\delta)(1-\theta') + \delta V_1(B)$$

and then, by equation 31,

$$W_1(-) = (1-\delta)(1-\theta') + \delta \frac{1+\delta(1-\alpha)}{1+\delta} \tag{32}$$

Histories in $\mathbb{H}_-$

Let $h_T \in \mathbb{H}_-$, then $\theta(h_T) = \bar{\theta} = (0,0)$ and $a(h_T, \bar{\theta}) = B$. If $h_T \in \mathbb{H}_-$ and player 1 adopts the equilibrium strategies, all the followers of $h_T \in \mathbb{H}_-$ so generated stay in the set $\mathbb{H}_-$. We denote by $W_1(0)$ and $V_1(0)$ the payoff generated by the equilibrium strategies for histories $h_T \in \mathbb{H}_-$. It is

$$\begin{aligned} W_1(0) &= (1-\delta) + \delta V_1(0) \\ V_1(0) &= (1-\delta) + \delta W_1(0) \end{aligned}$$

thus

$$W_1(0) = V_1(0) = 1 \tag{33}$$

5.1.2 Player 2

Histories in $\mathbb{H} \backslash \mathbb{H}_+$.

There are three different types of these histories indexed by θ_T : a) $\theta_T = (0, \alpha)$, b) $\theta_T = (0, \theta')$, c) $\theta_T = (\theta_T(A), \theta_T(B))$, with either $\theta_T(B) - \theta_T(B) \geq \alpha$ or with $\theta_T(B) - \theta_T(B) \geq \theta'$.

a) By definition of equilibrium strategies, $a(h_T) = B$, for all $h_T \in \mathbb{H} \backslash \mathbb{H}_+$. Let $V_2(\alpha)$ denote the value of 2 which is invariant for all $h_T \in \mathbb{H} \backslash \mathbb{H}_+$ with $\theta_T = (0, \alpha)$. The value for player 2 at $(h_T, B) \in \mathbb{H} \backslash \mathbb{H}_+$ is invariant for all histories $h_{T+1} \in \mathbb{H} \backslash \mathbb{H}_+$ with $a_{T+1} = B$ and is denoted by $W_2(B)$. Then

$$\begin{aligned} V_2(\alpha) &= (1-\delta)\alpha + \delta W_2(B) \\ W_2(B) &= \delta V_2(\alpha) \end{aligned}$$

that is

$$V_2(\alpha) = \frac{\alpha}{1+\delta} \; ; \tag{34}$$

$$W_2(B) = \delta \frac{\alpha}{1+\delta} \tag{35}$$

b) By definition of equilibrium strategies, $a(h_T) = B$, for all $h_T \in \mathbb{H} \backslash \mathbb{H}_+$. Let $V_2(\theta')$ denote the value of 2 which is invariant for all $h_T \in \mathbb{H} \backslash \mathbb{H}_+$ with $\theta_T = (0, \theta')$. Then

$$V_2(\theta') = (1-\delta)\theta' + \delta W_2(B) = (1-\delta)\theta' + \delta^2 \frac{\alpha}{1+\delta} \tag{36}$$

c) By definition of the equilibrium strategies, $a(h_T) = B$, while, as $(h_T, B) \in \mathbb{H} \backslash \mathbb{H}_+$, it is $\theta(h_T, B) = (0, \alpha)$. Then by a) and b), it is $W_2(h_T, B, (0, \theta') = W_2(B)$. Then we get

$$V_2(h_T) = (1-\delta)\theta_T(B) + \delta W_2(B) = (1-\delta)\theta_T(B) + \delta^2 \frac{\alpha}{1+\delta} \tag{37}$$

Histories in $\mathbb{H}_+ \backslash \mathbb{H}_-$.

For such histories $a(h_T) = A$ and, since $(h_T, A) \in \mathbb{H}_+ \backslash \mathbb{H}_-$, $\theta(h_T, A) = (0, \theta')$ thereby implying that $(h_T, A, (0, \theta')) \in \mathbb{H} \backslash \mathbb{H}_+$. Thus, by b), $W_2(h_{T+1})$ is invariant for all $h_{T+1} \in \mathbb{H}_+ \backslash \mathbb{H}_-$ with $a_{T+1} = A$ and it is denoted $W_2(A)$. It is

$$W_2(A) = \delta V_2(\theta') = \delta(1-\delta)\theta' + \delta^3 \frac{\alpha}{1+\delta} \tag{38}$$

and thus by 36

$$\begin{aligned} V_2(h_T) &= (1-\delta)\theta_T(A) + \delta W_2(A) = (1-\delta)\theta_T(B) + \delta^2 V_2(\theta') \\ &= (1-\delta)\theta_T(B) + \delta^2(1-\delta)\theta' + \delta^4 \frac{\alpha}{1+\delta} \end{aligned} \tag{39}$$

Histories in $\mathbb{H}_-$.

We distinguish three such histories parametrized by θ_T : i) $\theta_T = (0,0)$, ii) $\theta_T = (\theta_T(A), \theta_T(B))$, with $\theta_T(A) > \theta_T(B)$, and iii) $\theta_T = (\theta_T(A), \theta_T(B))$, with $\theta_T(A) \leq \theta_T(B)$.

i) Then, $a(h_T) = B$. For all histories following (h_T, B) generated by equilibrium strategies, player 1 transfer is $\bar{\theta} = (0,0)$, while player 2 action is B. We call $V_2(0)$ and $W_2(0)$, the payoff so generated. Thus,

$$V_2(0) = \delta W_2(0) \text{ and } W_2(0) = \delta V_2(0)$$

and hence

$$V_2(0) = W_2(0) = 0 \tag{40}$$

ii) Then $a(h_T) = A$ and hence

$$V_2(h_T) = (1-\delta)\theta_T(A) + \delta W_2(0) = (1-\delta)\theta_T(A).$$

iii) Then $a(h_T) = B$ and hence

$$V_2(h_T) = (1-\delta)\theta_T(B) + \delta W_2(0) = (1-\delta)\theta_T(B).$$

5.1.3 Player 2 does not have incentives to deviate

We analyze one-shot deviations. We have to analyze four possible such deviations: a) player 2 picks A at $h_T \in \mathbb{H} \backslash \mathbb{H}_+$ with $\theta_T(B) - \theta_T(A) \geq \alpha$, b) picks A at $h_T \in \mathbb{H} \backslash \mathbb{H}_+$ with $\theta_T(B) - \theta_T(B) \geq \theta'$, c) picks A (B) at $h_T \in \mathbb{H}_-$ with $\theta(A) > \theta(B)$ ($\theta(A) \leq \theta(B)$), d) picks B at $h_T \in \mathbb{H}_+ \backslash \mathbb{H}_-$.

a) Player 2 payoffs generated by such deviation are:

$$(1-\delta)\theta_T(A) + \delta W_2(h_T, A)$$

By definition of equilibrium strategies $\theta(h_T, A) = (0, \theta')$ and thus

$$W_2(h_T, A) = \delta V_2(\theta').$$

Thus such a deviation is not profitable if

$$\begin{aligned} V_2(h_T) &= (1-\delta)\theta_T(B) + \delta W_2(B) \geq (1-\delta)\theta_T(A) + \delta^2 V_2(\theta') \\ &= (1-\delta)\theta_T(A) + \delta^2[(1-\delta)\theta' + \delta W_2(B)] \end{aligned}$$

As $\theta_T(B) - \theta_T(A) \geq \alpha$, the last inequality is satisfied if

$$(1-\delta)\alpha + \delta W_2(B) \geq +\delta^2[(1-\delta)\theta' + \delta W_2(B)]$$

This is the case as by equation 35 and equation 29, the left and right hand side of the inequality above are identical.

b) Player 2 payoffs generated by such deviation are:

$$(1-\delta)\theta_T(A) + \delta W_2(h_T, A).$$

As $(h_T, A) \in \mathbb{H}_+ \backslash \mathbb{H}_-$, $\theta(h_T, A) = (0, \theta')$. Thus, by definition of equilibrium strategies

$$W_2(h_T, A) = \delta V_2(\theta').$$

If player 2 follows equilibrium strategies at h_T, payoffs are $V_2(h_T)$ as in equation 37. Thus, by taking into account that $\theta_T(B) - \theta_T(B) \geq \theta'$, such a deviation is not profitable if

$$(1-\delta)\theta' \geq \delta[\delta V_2(\theta') - W_2(B)] = \delta[(1-\delta)\delta\theta' - (1-\delta^2)W_2(B)$$

which is obviously verified as both $W_2(B) > 0$ and $\theta' > 0$.

c) Since $h_T \in \mathbb{H}_-$, any story following h_T generated by player 1 playing equilibrium strategies are in $\mathbb{H}_-$. Thus, if $\theta_T(A) \leq \theta_T(B)$, by equation 40, the payoffs generated by such deviation are:

$$(1-\delta)\theta(A)$$

which is less or equal than $(1-\delta)\theta_T(B)$, the payoff generated by the equilibrium strategy. Identical argument applies if $\theta_T(B) \geq \theta_T(A)$.

d) Since $h_T \in \mathbb{H}_+ \backslash \mathbb{H}_-$, it is be such that $\theta_T(B) - \theta_T(A) < \theta'$. By deviating and selecting B, player 2 generates a history $(h_T, B) \in \mathbb{H}_-$. Thus payoffs are

$$(1-\delta)\theta_T(B)$$

Had player 2 followed the equilibrium strategy payoffs would have been $V_2(h_T)$ as in equation 39. Thus such a deviation is not profitable if:

$$(1-\delta)\theta_T(A) + \delta W_2(A) \geq (1-\delta)\theta_T(B)$$

which, since $\theta_T(B) - \theta_T(A) < \theta'$, it is certainly verified if, by equation 38

$$\delta^2(1-\delta)\theta' + \delta^4 \frac{\alpha}{1+\delta} \geq (1-\delta)\theta'$$

The latter, by equation 29, can be rewritten, after straightforward computations as

$$\delta^2(1+\delta^2) \geq 1 \tag{41}$$

which is certainly verified for δ close enough to 1.

5.1.4 1 does not have incentives to deviate

We have to check four possible deviations: a) player 1 offers $\hat{\theta}$ with $\hat{\theta}(B) - \hat{\theta}(A) < \alpha$, at $h_T \in \mathbb{H} \backslash \mathbb{H}_+$; b) player 1 offers $\theta \neq (0,0)$ at $h_T \in \mathbb{H}_-$; c) player 1 offers $\tilde{\theta}$, with $\tilde{\theta}(B) - \tilde{\theta}(A) < \theta'$ at $h_T \in \mathbb{H}_+ \backslash \mathbb{H}_-$.

a) Since $h_T \in \mathbb{H} \backslash \mathbb{H}_+$, it is $(h_T, \hat{\theta}) \in \mathbb{H}_+ \backslash \mathbb{H}_-$. Therefore, $a(h_T, \hat{\theta}) = A$ and then $\theta(h_T, \hat{\theta}, A) = \theta'$ implying that $W_1(h_T, \hat{\theta}, A) = W_1(-)$. Thus payoffs generated by such deviation are:

$$\begin{aligned}
-(1-\delta)\hat{\theta}(A) + \delta V_1(h_T, \hat{\theta}, A) &= -(1-\delta)\hat{\theta}(A) + \delta^2 W_1(h_T, \hat{\theta}, A) \\
&= -(1-\delta)\hat{\theta}(A) + \delta^2 W_1(-) \\
&= -(1-\delta)\hat{\theta}(A) + \delta^2[(1-\delta)(1-\theta') + \delta V_1(B)]
\end{aligned}$$

while equilibrium payoffs are

$$W_1(h_T) = (1-\delta)(1-\alpha) + \delta V_1(B)$$

As $\theta_T(A) \geq 0$ and, by equations 29 and 31, $\alpha < \theta'$ and $V_1(B) > 0$, such deviations is unprofitable.

b) If $h_T \in \mathbb{H}_-$ all the followers of h_T generated by player 1 following the equilibrium strategy are in $\mathbb{H}_-$. Furthermore, by equation 33, equilibrium payoffs for player 1 are the highest possible payoffs. Therefore, there cannot be a profitable deviation at $h_T \in \mathbb{H}_-$.

c) Since $h_T \in \mathbb{H}_+ \backslash \mathbb{H}_-$, if player 1 offers $\tilde{\theta}$, is $(h_T, \tilde{\theta}) \in \mathbb{H}_+ \backslash \mathbb{H}_-$ and hence $a(h_T, \tilde{\theta}) = A$, while $\theta(h_T, \tilde{\theta}, A) = (0, \theta')$. Thus payoffs generated by such deviation are:

$$\begin{aligned} -(1-\delta)\tilde{\theta}(A) + \delta V_1(h_T, \theta, A) &= -(1-\delta)\tilde{\theta}(A) + \delta^2 W_1(h_T, \theta, A) = \\ & -(1-\delta)\tilde{\theta}(A) + \delta^2 W_1(-) \ , \end{aligned}$$

while equilibrium payoffs at $h_T \in \mathbb{H}_+ \backslash \mathbb{H}_-$ are $W_1(-)$.

Thus, as $\tilde{\theta}(A) \geq 0$, the deviation is not profitable if $W_1(-) > 0$, that is, by equation 32, if

$$(1-\delta)(1-\theta') + \delta^2 \frac{1+\delta(1-\alpha)}{1+\delta} > 0$$

The latter defines an upper bound on α, $\bar{\alpha}(\delta)$, which by the definition of θ'(equation 29) can be written after trivial computations as:

$$\bar{\alpha}(\delta) = \frac{\delta^2(1+\delta^3)}{1+\delta^5-\delta^4}$$

which is strictly positive for $\delta > 0$.

Thus we have shown:

Fact: *Let δ satisfy inequality 41. Then, the set of SPNE contains a continuum of equilibria indexed by $\alpha \in [0, \bar{\alpha}(\delta)]$. Equilibrium payoffs are one to one in α.*

6 Extensions

6.1 N Players

Result generalizes - To be Written

6.2 Transfers with No Commitment

To be Written

6.3 Infinite Transfer Histories

To be Written

6.4 Mechanism Design and Folk Theorems in Repeated Games

To be Written

7 Conclusion

To be Written

8 References

To be Written