A Welfare Theorem for Asynchronous Games with Transfers: Part II - Infinite Horizon*

Prajit K. Dutta   Paolo Siconolfi
Columbia University

October 12, 2017

Abstract

This paper - and its companion, Dutta-Siconolfi (2016a) - proves a First Welfare Theorem for Games. It shows that infinite horizon asynchronous dynamic games with voluntary one period ahead transfers have a unique equilibrium that coincides with the Utilitarian Pareto Optimum and hence can be computed from a (simpler) programming problem (rather than as a fixed point). The only way that multiplicity can arise is from strategies that have an infinite memory of transfers.

JEL: C3, C73.

Key words: transfers, asynchronous, Folk Theorem.

1 Introduction

This paper studies an infinite horizon dynamic game. It shows that if players *take turns* moving and can commit to a *one-period ahead transfer*, then there is a **unique** Subgame Perfect Equilibrium and it is **Pareto Optimal**. This stands in contrast to the most important and most cited result in dynamic games, the Folk Theorem which posits the exact opposite - multiple, indeed infinite number of, equilibria that are, generically, inoptimal.[1]

Consider the two italicized assumptions: *take turns* and *one-period ahead transfer.* Individually, the two assumptions are standard in the literature and have been used to model a variety of economic applications. Take the timing assumption of alternating moves. Although simultaneous moves is the more common assumption (and virtually all Folk Theorem results are proved under that assumption), many authors have argued that it should not be the *only*

*This research has benefited from the comments of Dilip Abreu, Paolo Barelli, Pierre-Andre Chiappori and Debraj Ray and seminar participants at Columbia, Cambridge, Paris I and Rochester. Errors are, sadly, our very own.

[1] By now, there are over a hundred papers that prove the result. The result has been proved for finite and infinite horizon, infinitely lived players, overlapping generations, complete and incomplete information, public and private monitoring...

benchmark for repeated interaction. Two seminal papers in which alternating moves are analyzed are Rubinstein (1982) and Maskin and Tirole (1988a, b). In the bargaining model of Rubinstein (1982), if a player turns down a proposed agreement, she commits to an alternative proposal for the next period. Similarly, Maskin and Tirole (1988a, b) assume that firms are committed to an action for a period due to exogenous technological reasons; installed capital that has little scrap value, or lags in producing and disseminating price lists or short-term contracts that bind a firm temporarily to an action. Alternating moves is implied if a competitor is able to react before the commitment expires.[2] [3]

Alternating moves lead to interesting equilibrium consequences in the above papers. Rubinstein (1982) showed that, unlike in simultaneous offer bargaining where any share is a static equilibrium, here there is a unique equilibrium (and it asymptotically replicates the Nash Bargaining Solution). Maskin and Tirole (1988a, b) showed that classic Industrial Organization predictions like the kinked demand curve or the Edgeworth cycle can emerge as equilibria when firms make short-term commitments.

Yet these results are still special in that there are many other equilibria as well. Alternating move games are examples of stochastic games (with the fixed actions in a period being the "state"). Hence, by Dutta (1995) there is a Folk Theorem in these games.[4]

What about *transfers*? That transfers are feasible is a standard assumption in, for example, (static) Mechanism Design and Auction Theory. Indeed, in those contexts, transfers have been shown to induce social optimality (via the Vickrey-Clarke-Groves mechanism). When interaction is repeated, it is also understood that transfers can be a more efficient way to distribute intertemporal incentives rather than (Abreu-Pearce-Stachetti constructions using) continuation values of the game. Yet, in general, when transfers are present and can be made contingent on players' own-actions, even a static game can have multiple equilbria. Worse, none of these equilibria need be efficient. This point has been forcefully made by Jackson and Wilkie (2005).

Hence, when the stage game (with transfers) is repeated, yet again a Folk Theorem emerges since multiple Nash equilibria in the stage game, some of them inefficient, is exactly the raw material from which the sauce of Folk Theorems is cooked.[5] So, having alternating moves by itself - or transfers by itself - still keep us trapped in the multiplicity and no prediction world of Folk Theorems.

What if the two assumptions are combined? This paper shows that has

[2] Other papers that study asynchronous moves (with action commitments) include Lagunoff and Matsui (1997), Kamada and Sugaya (2010), Takahashi (2005)and Dutta (2012) in coordination, Admati and Perry (1987) and Marx and Matthews (2000) in public goods and the vast literature in evolutionary games that study birth-death processes. See also Bhaskar (1998) and Bhaskar and Vega-Redondo (2008) and the Folk Theorem results in Yoon (2001), and Wen (1994 and 2012).

[3] The relational contracting literature - see Levin (2010) and the follow-up papers - offer a contrasting example where players take turns moving but at least one player, the Principal, cannot commit to an action.

[4] See also Yoon (2001), and Wen (1994 and 2012).

[5] A formal proof of this can be found in Goldücke and Krantz (2012).

dramatic consequences: it leads to a collapse of the Folk Theorem - and we have a unique, optimal equilibrium. This statement comes with one caveat: our result requires that players remember only a finite number of past transfers (though they may remember the entire past history of stage-game actions). Indeed we show in an example, that a Folk Theorem emerges if there is infinite memory of transfers as well.

How "real-world" are the two assumptions? There are many instances of non-simultaneity: collective bargaining[6], time-stamping/encryption[7], announcements of national policy[8] etc.. Transfers are also widespread. In the Paris Climate Accord, there is a plan - the "Green Climate Fund" - to provide US$100 billion a year in aid to developing countries to implement new procedures to minimize climate change (with higher amounts in subsequent years). Private companies routinely make transfers: they pay mineral royalties to extract resources[9], they pay licensing fees to use patented technologies[10], they pay copyright fees, trademark royalties, etc.

That players can commit to an action and a transfer schedule for one period is restrictive in some contexts. For instance, if other players' choices are unobserved. Or, as in relational contracting, if players can renege on payments. However, with an escrow account, a transfer commitment can be made. Similarly, if there is a legal system that enforces contracts, then companies can legally commit to future payments.

Note though that the commitment is short-term, i.e., is only for one period. It is therefore coterminous with the action commitment.[11] We are not allowing a player to commit to a life-time of transfers. That would turn our model into a one-period model and surely no one would be surprised at the efficiency result. The power of the result derives precisely from the fact that a minimal period commitment delivers a maximal period efficiency. Put differently, we propose a decentralized mechanism to implement the efficient solution, one in which players take turns being "Principals" and trade one-period contracts to generate dynamic efficiency.

Most of this paper analyzes a two-player alternating move game played over

---

[6] Collective bargaining accounts for trillions of dollars on an annual basis in the United States. At the Federal level alone, the President and Congress negotiate an annual spending authorization of around $4 trillion.

[7] *Timestamping* establishes priority based on when messages are sent to a common pool and is widespread in encryption, including in Bitcoins.

[8] For example, the Paris Accord of 2015 was arrived at in a "bottoms up" way by individual countries asynchronously submitting "nationally determined contributions" (INDCs), in contrast to past attempts - like Kyoto - where there has been an (unsuccessful) attempt to simultaneously determine each country's commitments.

[9] From mining companies, the US Government earns between $10bn and $15bn a year in mineral leases. See https://cdn.americanprogress.org/wp-content/uploads/2015/06/RevenueOilGas-brief.pdf.

[10] The size of these royalties can be quite significant; average royalty rates are about 15% of EBITDA. See, Kemmerer and Lu (2012).

[11] Our transfers are conditioned on just the other player's action in the next period. More generally one could condition transfers on the entire history. However, as the past history is sunk, this reduces to our formalization.

an infinite horizon - though in Section 6, and in detail in Dutta and Siconolfi (2016a), we show that the same result holds for any N player, stationary asynchronous game. There is a fixed stage game. In each period, one player moves (and the other player's move remains fixed). A move has two components; first, an action from the stage game. Second, a non-negative transfer schedule according to which she pays her opponent next period (the actual payment depending on the opponent's stage-game action). Period payoffs for the mover are the sum of stage-game payoffs and the transfer from her opponent while, for the non-mover, they are stage-game payoffs less transfer.

To contextualize this paper, consider the following typological table:

	Simultaneous	*Asynchronous*
No Transfers	*Canonical*[1] *FT*	*Stochastic Game*[2] *FT*
Transfers	*Also*[3] *FT*	**This Paper**

where references are: 1− Canonical Model (Fudenberg and Maskin (1986) plus 100 papers)[12]; 2− Stochastic Game (Dutta (1995) plus 10 papers)[13]; 3− Simultaneous Transfers (Goldücke and Krantze (2012), Dutta and Siconolfi (2016b)).

As discussed above, in the related literature *either* asynchronicity has been studied or (static) transfers but not both and that is our paper (in the bottom right). With only asynchronicity, top right, there is a Folk Theorem - by Dutta (1995) - because these are special cases of stochastic games. Also see Yoon (2001) who provides a direct proof of the Folk Theorem for Asynchronous Repeated Games. What of the bottom left, transfers alone? For instance, what if the (static) Prisoners Dilemma is played simultaneously but, ahead of play and simultaneously, players pick transfer schedules? There is a literature on Public Goods (Guttman (1978), Varian (1994)) that showed that one could indeed get the Pareto optimal outcome by this two-stage procedure. However, Jackson and Wilkie (2005) pointed out that those results are crucially predicated on the restriction that transfers can be made dependent only on the opponent's action but not on one's own. Once a richer set of transfer schedules is allowed, Jackson and Wilkie show that there are going to be multiple equilibria and none of them need be efficient. This paper considers the same set of transfer schedules as Jackson and Wilkie.[14]

[12] Folk Theorems show up in the canonical complete information model of infinitely and finitely repeated games (respectively, Fudenberg and Tirole (1986) and Benoit and Krishna (1985)) as also in variants of the canonical model - in the more general set-up of stochastic games, (Dutta (1995), Fudenberg and Yamamoto (2013) and Horner, Sugaya, Takahashi, Vielle (2011)), and with overlapping generations of players (Smith (1997)). They also show up in informational variants - with imperfect public monitoring (Fudenberg, Levine and Maskin (1994)), and even with private monitoring (Horner and Olszewski (2006)) and Sugaya (2016)). Sugaya (2016) is a fairly up-to-date reference on informational variants (and contains about fifty citations). Peski 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).

[13] Dutta (1995), Fudenberg, Yamamoto (2013)), Horner, Sugaya, Takahashi, Vielle (2011).

[14] In our setting, a mover at t decides on a transfer schedule for $t+1$ (with the actual transfer depending on the opponent's action choice in that period). Typically, the schedule chosen

The companion piece, Dutta and Siconolfi (2016a) looks at the finite horizon case. The reader may wonder about the relation between the two papers. They are quite different in much the same way that in simultaneous move Repeated Games, the finite and infinite horizon models are different; they require different approaches and have different results.[15] In the simultaneous move model - modulo a dimensionality requirement - the infinite horizon model always has a Folk Theorem but the finite horizon model has one only with additional assumptions on multiplicity of stage game Nash Equilibria.[16] With alternating moves and transfers, it is the infinite horizon model that is more demanding. Dutta and Siconolfi (2016a) show that in the finite horizon model, with no restrictions on strategies, there is a unique SPE that is UPO. In this paper, dealing with the infinite horizon model - we show that the result is only true if the memory on transfer histories is bounded. Else, there is a Folk Theorem. The proof method is also very different - an operator based proof here and backwards induction in the companion paper.

The paper is organized as follows. In Section 2 we detail the two-player alternating move model and in Section 3 we state the main theorem. To better understand the (long) proof of the main theorem - and the APS-like operator that is crucial for the proof - that is in Section 5, we provide an Illustrative Example using the Prisoners Dilemma in Section 4. Section 6 contains an example that illustrates the folk theorem that arises with infinite transfer histories. In Section 7 we extend the result to N players and any asynchronous game.

2 Model

Let G denote a two player stage game. 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.[17] Let i denote the generic player and let j denote the "other" player, $j \neq i$.

The timing structure is **alternating moves**.[18] 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 can then be changed only at $t+2$. Payoffs are ongoing. If $\bar{a}_j$ denotes the (fixed) action of the other player, then the period t stage game payoffs are $\pi_i(a_{it}, \bar{a}_j)$, $\pi_j(a_{it}, \bar{a}_j)$. There is an *initial action state* $\bar{a}$ for the

will also depend on the mover's action choice at t. Hence, the transfer at $t+1$ will depend on both players' actions. Note also that, with simultaneous transfers, Dutta and Siconolfi (2016b) and Goodlucke and Krantz (2015) have shown that there is a Folk Theorem.

[15] Another way of saying this is that games are different from decision problems with regard to the effect of horizon. In Dynamic Programming, value functions in the finite horizon converge to the infinite horizon value function. However, in games, the SPE value correspondence is not continuous in the horizon.

[16] See Fudenberg and Maskin (1986) and Benoit and Krishna (1985).

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

[18] In Section 6 we discuss why the order of moves is unimportant.

game.[19] 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** next period to the other player, tied to his then action. For example, if in the previous period the then mover (player j) had played $\bar{a}_j$ and promised the current mover that she would be paid according to the schedule $\theta_j(a_i) \geq 0$ then the total payoffs of player i inclusive of transfers will be $\pi_i(a_{it}, \bar{a}_j) + \theta_j(a_i)$ while that of player j would be $\pi_j(a_{it}, \bar{a}_j) - \theta_j(a_i)$.

Lifetime payoffs are evaluated according to the discounted *average;* player $i's$ evaluation of the payoff will be given by

$$(1-\delta)\sum_{t=0}^{\infty}\delta^t\{[\pi_i(a_{it}, a_{jt-1}) + \theta_{jt-1}(a_{it}) \parallel i \; mover] + \\ +[\pi_i(a_{it-1}, a_{jt}) - \theta_{it-1}(a_{jt}) \parallel j \; mover]\}$$

2.1 Equilibrium

Let us start with an assumption on what is publicly observable/verifiable:

Informational Assumption - 1. *All past stage game actions are observable.* Hence, an *action history* h_t^a at time t is $h_t^a = (\bar{a}, a_{i0}, ..., a_{jt-1})$.

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 is $h_t^\theta = (\theta_{it-M}(\cdot), ...\theta_{jt-1}(\cdot))$.

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 $\theta_{jt-1}(\cdot)$, 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 $\geqslant 1$. (In the build-up to the main theorem we will prove a preliminary result when $M = 1$.) Let $h_t = (h_t^a, h_t^\theta)$.

For the $t-th$ period *action choice* a_{it}, the *payoff-relevant history* is only the preceding period's choices: $a_{jt-1}, \theta_{jt-1}(\cdot)$, the fixed action of the other player j and the transfer schedule that j has picked. For the $t-th$ period *transfer choice* $\theta_{it}(\cdot)$, 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. We will exploit this difference in payoff-relevance between a_{it} and $\theta_{jt-1}(\cdot)$ in the analysis that follows.

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 $\theta_{it}(\cdot)$ 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. (The arguments below will show that the restriction to pure strategies is without loss of generality.) A strategy vector - one strategy for every player - defines in the usual way a (possibly probabilistic)

[19] 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 $(a_{i0}, \overline{a})$.

history h_t. Denote the lifetime payoff of the mover at time t, v_i:

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

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

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

where h_{t+1} and h_{t+2} are the histories at periods t+1 and t+2 caused by player $i's$ period t actions, that is, $h_{t+1} = (h_t, a_{it}, \theta_{it}(\cdot))$ and $h_{t+2} = (h_{t+1}, a_{jt+1}, \theta_{jt+1}(\cdot))$. In that history, the follower player $j's$ best response action in period $t+1$, $a_j(h_{t+1})$, comes from the $t+1$ period analog of Eq. 1:[20]

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

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

$$w_j(h_{t+1}, a_j) = \max_{\theta_j(\cdot)}\{(1-\delta)[\pi_j(a_i(h_{t+2}), a_j) - \theta_j(a_i)] + \delta v_j(h_{t+3}))\}$$

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

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 $a_{jt-1}, \theta_{jt-1}(\cdot)$ and a function $\theta_{it}(\cdot)$ that maps from a_{it}. 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: Unique Optimal Equilibrium

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

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:*

$$\max_{\{a_{it}, a_{jt}\}_{t\geq 0}} (1-\delta)\sum_{t=0}^{\infty}\delta^t[\pi_i(a_{it}, a_{jt}) + \pi_j(a_{it}, a_{jt})] \tag{2}$$

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

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

It is well known that we can restrict attention to Markovian strategies in searching for a UPO solution, i.e., the mover's action at t depends only on the action chosen by the non-mover at $t-1$. Hereafter, we assume that the Markovian solution to the UPO problem is unique. The argument does not change if there exist multiple solutions.

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

Theorem 2 *There is a unique SPE. The sum of payoffs is equal to the UPO value, the SPE actions coincides with the UPO solution, and the transfers are uniquely determined.*

Since the UPO is a Markovian strategy, it follows from the theorem that SPE are, in fact, MPE. It is immediate to check from the argument that in environments with multiple UPO solutions uniqueness of SPE values is preserved, while SPE strategies may be multiple.

4 An Illustrative Example

Consider the following Prisoners Dilemma stage game:

1\2	$confess\ (C)$	$not\ confess\ (N)$
$confess\ (C)$	$0, 0$	c, b
$not\ confess\ (N)$	b, c	d, d

with $c > d > 0 > b$ (making $confess\ (C)$ a dominant strategy in the stage game) and $2d > b+c$ (making (N, N) the static Pareto optimum). The dynamic Pareto optimum - the UPO defined in the previous section - is a Markovian strategy; the action chosen at t depending only on the action chosen at $t-1$. Since the latter is either C or N, the UPO strategy is itself given by a pair - $\hat{f}(N)$, $\hat{f}(C)$ - where $\hat{f}(N)$ (respectively, $\hat{f}(N)$) is the action to be played by j at period t if i played N (respectively, C) in period $t-1$.

Computing the UPO Clearly, given that $2d$ is the highest total stage payoff, $\hat{f}(N)$ must be N. What if the previous period's action was C? Then, picking N yields a lifetime payoff of $(1-\delta)(b+c)+2\delta d$ while picking C - invoking unimprovability - yields 0. Clearly, we have the following two cases:

Case I - $b+c > 0$, then, regardless of δ, $\hat{f}(C) = N$. Consequently, regardless of the initial state, play converges to the static Pareto optimum (within two periods).

Case II - $b+c < 0$, then for "low" $\delta \leqslant \hat{\delta}$,[21] $\hat{f}(C) = C$, i.e., the UPO strategy $\hat{f}$ is to continue the previous period's action (whether C or N). However, for $\delta > \hat{\delta}$ the UPO strategy is $\hat{f}(C) = N$, i.e., always play N. So, for $\delta \leqslant \hat{\delta}$, play in the UPO need not converge to the static Pareto optimum if it initially starts at C.

[21] The cut-off $\widehat{\delta}$ is computed from solving $(1-\delta)(b+c) + 2\delta d = 0$.

For the rest of this section - purely for expositional ease - suppose we have *Case I.* We now illustrate the power of transfers by showing

1. that the UPO is *never* a SPE *without* transfers, but
2. *is* a SPE *with* transfers (and transfers are easily computed).

Without Transfers the UPO is not a SPE

A straightforward way to see that the UPO strategy, $\hat{f}(N) = \hat{f}(C) = N$, is not an SPE is to note that the initial state N is absorbing. Yet, a one-shot deviation, in which player i deviates to C, leads to "no punishment" since the other player j plays N next period and two periods later, play is back to N. Hence, the lifetime payoff stream to deviation is $c, c, d, d, ...$ and that is clearly better than that to not deviating $d, d, d, d, ...$ (independent of δ).

With Transfers the UPO is a SPE

Now suppose the UPO strategy, $\hat{f}(N) = \hat{f}(C) = N$, is embellished with transfers, θ_N, θ_C that each player offers the other, j, if she plays N in the next period. Hence, θ_N (respectively, θ_C) is the transfer promised if i played N (respectively, played C) in the current period. (The transfers typically depend on what i played himself in the current period, i.e., $\theta_N \neq \theta_C$.)

We check for two deviations, one on actions and the other on transfers: given transfers, θ_N, θ_C, will players play the UPO strategy $\hat{f}(\cdot)$? Second, will they offer the transfers? Since there is always going to be (large enough) transfers that will induce the UPO action, the first question can be re-thought as: what are the *minimum* transfers θ_N, θ_C such that they will induce the UPO strategy $\hat{f}(\cdot)$?

To simplify the analysis, note that, without loss, we can restrict attention to one-shot deviations, i.e., playing C instead of N in one period but therafter reverting to N - the same deviation we checked in the without transfers case. So, if the initial state is N, then a deviation to C leads to a stream of payoffs c, $c - \theta_C$, $d + \theta_N$, $d - \theta_N, ...$ Note that in the first period, the player gets no transfer since transfers are only made for playing N. Having played C, she then has to go back to the putative strategy and offer θ_C to the other player to get him to play N. By the third period, play is back to the equilibrium path. Not deviating yields $d + \theta_N$, $d - \theta_N$, $d + \theta_N$, $d - \theta_N, ...$

Since the only difference is in the first two periods, a simple computation yields that the minimum transfer must satisfy

$$(1 - \delta)\,\theta_N + \delta\theta_C = (1 + \delta)(c - d). \tag{3}$$

A simple special case to understand Eq. 3 is $\delta = 1$. Then, deviating gets i two periods of c rather than d as payoff, i.e., an extra payoff $2(c - d)$. As long as he has to "pay" the same amount in transfer θ_C - in order to get j to play N next period - deviating is unprofitable.

Starting at initial state C, a similar two period analysis yields the following second condition

$$(1 + \delta)\,\theta_C - \delta\theta_N \geq \delta(c - d) - b, \tag{4}$$

with equality if $(\theta_N, \theta_C) \gg 0$.

In the case $\delta = 1$ - where $\theta_C = 2(c-d)$, it follows that $\theta_N = 3(c-d)+b$. Whereas $c > d$ always - and hence $\theta_C > 0$ - it need not be the case that $\theta_N > 0$. If it is not, then all that is implied is that there is no transfer required once the absorbing state N is reached. The precise characterization of the transfers, based on Eqs. 3 and 4 is:

Case A $3(c-d)+b < 0$: In this case, for "high" $\delta \geqslant \delta'$,[22] $\theta_N = 0$ and $\theta_C = \frac{1}{1+\delta}[\delta(c-d) - b]$. For "low" $\delta < \delta'$,

$$\theta_C = -b(1-\delta) + 2\delta(c-d),$$

$$\theta_N = (c-d)(1+2\delta) + \delta b$$

Case B $3(c-d)+b > 0$: $\delta' = 1$, i.e., both transfers are positive and given immediately above.

Having computed the minimum transfers required to induce UPO play, we can now turn to the second incentive check: will the donor want to make the transfer? Again, since play is identical after two periods, we need only consider what happens in the first two periods. Second, and importantly, note that we have computed the minimum transfers required by *making the recipient indifferent between deviating and not deviating.* Hence, checking whether the donor is better off from offering transfers is *equivalent to checking whether the sum of payoffs is higher from not deviating.* Finally, since transfers cancel out when added across donor and recipient, the sum of payoffs inclusive of transfers is higher whenever the sum of stage game payoffs, exclusive of transfers, is higher. For instance, starting from initial state N, the donor will offer a transfer provided

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

which, of course, holds. Similarly, starting from initial state C, the donor will offer a transfer provided

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

which, again, holds. Actually, the reasoning is straightforward. At the minimum transfers, the donor becomes the residual claimant of the social surplus regardless of the recipient's subsequent action. Hence, she acts like the social planner and induces that action which maximizes the surplus, i.e., she induces the UPO action. This logic is true at every point of time - and after every history. In other words, in all subgames the UPO is induced.

So far, we have shown that the UPO is an SPE. But the argument in the previous paragraph should be suggestive as to why it is also the *only* SPE. After all, best response must imply that transfers, when given, are minimal thereby turning the donor into the social planner and "leading" to the planner picking the UPO. One has work to do though because, for arbitrary strategies, play will not always return to the UPO after two periods. Moreover, if there is history dependence, one would need to check all subgames. Finally, the required

[22] δ' is the solution to $(c-d)(1+2\delta) + \delta b = 0$.

transfers need not be be simple Markovian strategies. And, of course, the stage game may be considerably more complex and have none of the simplifying features of the Prisoners Dilemma. This proof requires a generalized version of the Abreu, Pearce and Stacchetti (1986), hereafter APS, construction in repeated games.

The general analysis now follows.

5 Proof of the Main Result

5.1 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-APS (BAPS) fixed point argument.

As with all Bellman-APS constructions, the idea will be to

- a) define an operator Γ (in Section 3.1.1.) that reduces the infinite horizon problem to a sequence of one period problems and then

- b) use the Unimprovability Principle to show an equivalence between the fixed point(s) of the operator Γ and the SPE of the (original) infinite horizon problem (Section 3.1.3.).

- c) What we will additionally show is that, with alternating moves, the fixed point is unique (rather than set-valued), i.e., produces a function rather than a correspondence. (Section 3.1.2.)

- d) We will also show that the Bellman-APS operator Γ is functionally equivalent to another operator T which actually coincides with the UPO Dynamic Programming operator. Hence, its fixed point coincides with the UPO solution. (Section 3.1.1.)

5.1.1 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 $v_i(\cdot)$ and $w_j(\cdot)$ 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

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

$$\begin{aligned}
\Gamma(V_i, W_j)(\bar{a}_i) &= \{w_i, v_j : \exists a_j^*, \theta_i^*(a_j), (v_i, w_j)(a_j) \in (V_i, W_j)(a_j) : \\
a_j^* &\in \arg\max_{a_j} \{(1-\delta)[\pi_j(\bar{a}_i, a_j) + \theta_i^*(a_j)] + \delta w_j(a_j)\} \\
v_j &= (1-\delta)[\pi_j(\bar{a}_i, a_j^*) + \theta_i^*(a_j^*)] + \delta w_j(a_j^*) \\
\forall a_j', \theta_i'(\cdot),\ a_j' &\in \arg\max_{a_j} \{(1-\delta)[\pi_j(\bar{a}_i, a_j) + \theta_i'(a_j)] + \delta w_j(a_j)\} \\
w_i &= (1-\delta)[\pi_i(\bar{a}_i, a_j^*) - \theta_i^*(a_j^*)] + \delta v_i(a_j^*) \\
&\geqslant (1-\delta)[\pi_i(\bar{a}_i, a_j') - \theta_i'(a_j')] + \delta v_i(a_j')\}.
\end{aligned} \tag{5}$$

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 $\theta_i(\cdot|\bar{a}_i)$ 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 inclusive of transfers (respectively, $\pi_i(\bar{a}_i, a_j) - \theta_i(a_j)$ and $\pi_j(\bar{a}_i, a_j) + \theta_i(a_j)$), 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.[23]

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

$$\begin{aligned}
T(V_i, W_j)(\bar{a}_i) &= \{w_i, v_j : \exists \hat{a}_j, (v_i, w_j)(a_j) \in (V_i, W_j)(a_j) : \\
\hat{a}_j &\in \arg\max_{a_j} \{(1-\delta)\pi_j(\bar{a}_i, a_j) + \delta w_j(a_j)\} \\
v_j &= (1-\delta)\pi_j(\bar{a}_i, \hat{a}_j) + \delta w_j(\hat{a}_j) \\
w_i &= \max_{\theta_i, a_j} \{(1-\delta)[\pi_i(\bar{a}_i, a_j) - \theta_i(a_j)] + \delta v_i(a_j)\} \\
\text{s.t. } v_j &= (1-\delta)[\pi_j(\bar{a}_i, a_j) + \theta_i(a_j)] + \delta w_j(a_j),\ a_j \in A_j\}.
\end{aligned} \tag{6}$$

As with the operator Γ, the simpler operator T also defines "continuation payoffs" starting at a period when player i is the mover. Note that in the definition of T, v_j is the mover's best response payoffs absent transfers in the current period. That is the *Binding Individual Rationality* (BIR) principle. Note that in T, the transfer schedule $\theta_i(\cdot)$ is implicitly defined given the continuation selections $v_i(\cdot), w_j(\cdot)$ whereas in Γ the transfer is seemingly determined alongside the continuations.

We first show that two operators are equivalent:

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

[23] 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. However, since only one-period memory of transfers is allowed, for the moment, $\theta_i^*(a_j|\bar{a}_i)$, the transfer made currently, is not present in the termination payoffs that originate two periods from the present.

Proof We first show that $T(V_i, W_j)(\bar{a}_i) \subseteq \Gamma(V_i, W_j)(\bar{a}_i)$. Given $w_i, v_j \in T(V_i, W_j)(\bar{a}_i)$, we have continuations $v_i(\cdot), w_j(\cdot)$ and an associated "no transfer" most preferred action $\hat{a}_j$ that satisfy Eq. 6. From the last line of that equation, we can derive the transfer schedule $\theta_i^*(\cdot)$ as:

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

Note that it follows that $\theta_i^*(\hat{a}_j|\bar{a}_i) = 0$. It further follows from the last line of Eq. 6 that all actions a_j yield the same "one-shot" payoff $(1-\delta)[\pi_j(\bar{a}_i, a_j) + \theta_i^*(a_j)] + \delta w_j(a_j|\bar{a}_i)$ to player j. Define a_j^* as

$$a_j^* \in \arg\max_{a_j} \{(1-\delta)[\pi_i(\bar{a}_i, a_j) - \theta_i^*(a_j)] + \delta v_i(a_j|\bar{a}_i)\} \tag{7}$$

Consider an alternative transfer schedule $\theta_i'(\cdot)$ and a best action choice a_j', i.e., suppose that $a_j' \in \arg\max_{a_j}\{(1-\delta)[\pi_j(\bar{a}_i, a_j) + \theta_i'(a_j)] + \delta w_j(a_j|\bar{a}_i)\}$.

In particular, then

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

the second inequality following from the fact that $\theta_i^*(\hat{a}_j) = 0$ (as established above) and the 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. 7 above we know that

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

and the last inequality follows from $\theta_i'(a_j') \geqslant \theta_i^*(a_j')$. Hence, all the requirements of Eq. 5 have been satisfied, i.e., we have established that $w_i, v_j \in T(V_i, W_j)(\bar{a}_i)$ imply that $w_i, v_j \in \Gamma(V_i, W_j)(\bar{a}_i)$.

For the opposite inclusion, suppose that $w_i, v_j \in \Gamma(V_i, W_j)(\bar{a}_i)$. Given $w_i, v_j \in \Gamma(V_i, W_j)(\bar{a}_i)$, we know that we have continuations $v_i(\cdot), w_j(\cdot)$ and we can define an associated "no transfer" most preferred action $\hat{a}_j$ that satisfies Eq. 6. Let us first show that $\theta_i^*(\hat{a}_j) = 0$. Suppose instead that $\theta_i^*(\hat{a}_j) = \zeta > 0$. We are also given a transfer inclusive most preferred action a_j^*. It follows that $\theta_i^*(a_j^*) > \zeta$. Consider an alternative transfer schedule $\theta_i'(\cdot)$ that reduces the transfers at $\hat{a}_j$ and a_j^* by ζ and reduces transfers at other actions such that the lifetime payoffs everywhere is the same as at $\hat{a}_j$, i.e.,

$$(1-\delta)\theta_i'(a_j) = (1-\delta)[\pi_j(\bar{a}_i, \hat{a}_j) - \pi_j(\bar{a}_i, a_j)] + \delta[w_j(\hat{a}_j) - w_j(a_j)] \tag{8}$$

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 $\theta_i^*(\hat{a}_j) = 0$. The same argument also establishes that the transfer at a_j^* must be such that player j is just indifferent between a_j^* and $\hat{a}_j$. Hence the first two lines of Eq. 6 have been established. That a_j^* must be derived from a maximization problem follows, for example, by considering a transfer scheme that satisfies Eq. 8. From Eq. 5 we know that

$$w_i = (1-\delta)[\pi_i(\bar{a}_i, a_j^*) - \theta_i^*(a_j^*)] + \delta v_i(a_j^*) \geqslant (1-\delta)[\pi_i(\bar{a}_i, a_j') - \theta_i'(a_j')] + \delta v_i(a_j')$$

and that is a re-statement of the penultimate equation in Eq. 6. 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. 6:

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

Substituting that into the definition of w_i, the second-last line in Eq. 6,

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

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

$$\begin{aligned} T(V_i, W_j)(\bar{a}_i) &= \{w_i, v_j : \exists \hat{a}_j, (v_i, w_j)(a_j) \in (V_i, W_j)(a_j) : \\ \hat{a}_j &\in \arg\max_{a_j}\{(1-\delta)\pi_j(\bar{a}_i, a_j) + \delta w_j(a_j)\} \\ v_j &= (1-\delta)\pi_j(\bar{a}_i, \hat{a}_j) + \delta w_j(\hat{a}_j) \\ w_i &= \max_{a_j}\{(1-\delta)[\pi_i(\bar{a}_i, a_j) + \pi_j(\bar{a}_i, a_j)] + \delta[v_i(a_j) + w_j(a_j)]\} - v_j\}. \end{aligned} \tag{9}$$

The last equation illustrates the *One is the Many* (OIM) Principle: to maximize her own payoffs, she has to maximize the sum of payoffs given that the non-mover gets a constant (IR payoff) across his actions.

Similarly, define the operator $T(V_j, W_i)(\bar{a}_j)$ 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. ■

5.1.2 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

$$\begin{aligned} V_i^{n+1}, W_j^{n+1} &= T(V_j^n, W_i^n) \\ V_j^{n+1}, W_i^{n+1} &= T(V_i^n, W_j^n) \end{aligned} \tag{10}$$

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 V_j^∞, W_i^∞ and V_i^∞, W_j^∞ as $n \to \infty$.*

Proof From the fact that the operator preserves compact-valuedness and non-emptiness it follows that each correspondence has that property along the sequence.[24] 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 singletons, for all (a_i, a_j).

Lemma 7 *The correspondences V_j^∞, W_i^∞ and V_i^∞, W_j^∞ are single valued.*

Proof By definition of V_i^∞, W_j^∞, $i \neq j$, $i = 1, 2$, and of the operator T, we have $T(V_i^\infty, W_j^\infty) = V_j^\infty, W_i^\infty$. Therefore, $(v_j, w_i; v_i, w_j) \in \times_{a_j \in A_j} V_j^\infty, W_i^\infty(a_j) \times \{\times_{a_i \in A_j} V_i^\infty, W_j^\infty(a_i)\}$ implies that

$$v_i(\bar{a}_j) = \max_{a_i}\{(1-\delta)\pi_i(a_i, \bar{a}_j) + \delta w_i'(a_i)\}, \text{ all } \bar{a}_j \tag{11}$$

$$(w_i + v_j)(\bar{a}_i) = \max_{a_j}\{(1-\delta)[\pi_i(\bar{a}_i, a_j) + \pi_j(\bar{a}_i, a_j)] + \delta(w_i' + v_j')(a_j)\}, \text{ all } \bar{a}_i \tag{12}$$

for some $(v_j', w_i'; v_i', w_j') \in \times_{a_j \in A_j} V_j^\infty, W_i^\infty(a_j) \times \{\times_{a_i \in A_j} V_i^\infty, W_j^\infty(a_i)\}$. By Eq. 12, $(w_i + v_j)(\bar{a}_i)$ (and therefore $(w_i' + v_j')(a_j)$) is the value of the Pareto problem 2 at $\bar{a}_i$ which we denote by $\mu(\bar{a}_i)$. Therefore:

$$v_j(\bar{a}_i) = \mu(\bar{a}_i) - w_i(\bar{a}_i), \text{ and } w_j'(a_j) = \lambda(a_j) - v_i'(a_j)$$

where, similarly, λ is the value of the Pareto problem 2 at a_j.[25] Substituting the latter into equations 11 for player j, we get:

$$\mu(\bar{a}_i) - w_i(\bar{a}_i) = \max_{a_j}\{(1-\delta)\pi_j(\bar{a}_i, a_j) + \delta[\lambda(a_j) - v_i'(a_j)\}, \text{ all } \bar{a}_i,$$

or equivalently

$$w_i(\bar{a}_i) = \min_{a_j}\{\rho(\bar{a}_i, a_j) + \delta v_i'(a_j)\}, \text{ all } a_i. \tag{13}$$

for $\rho(\bar{a}_i, a_j) = \mu(\bar{a}_i) - (1-\delta)\pi_j(\bar{a}_i, a_j) - \delta\lambda(a_j)$.

[24] Note that V_i^0, W_j^0 are compact because they are bounded: above by the 45° line going through $\max \pi_i + \pi_j$ and below by the quadrant defined by $\min_a \pi_i$ and $\min_a \pi_j$.

[25] μ and λ are both values for the UPO problem but in one case with a fixed action of player i as the initial state and, in the other, a fixed action of j. In a symmetric set-up, $\rho = \lambda$ but that need not to be true with asymmetry.

Claim 8 *The solution to equations 11 and 13 is unique.*

Proof Define the following two operators:

$$U_1 w(a_j) = \max_{a_i}\{(1-\delta)\pi_i(a_i,a_j) + \delta w(a_i)\} \tag{14}$$

and

$$U_2 v(a_i) = \min_{a_j}\{\rho(a_i,a_j) + \delta v(a_j)\} \tag{15}$$

where $w : A_i \longrightarrow \mathbb{R}$ 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 (1965) and is repeated here only because the minimization in Eq. 15 may appear non-standard to some readers who are familiar with Blackwell's proof. It is straightforward to see that

$$U_1(w+k)(a_j) = U_1 w(a_j) + \delta k, U_2(v+k)(a_i) = U_2 v(a_i) + \delta k \tag{16}$$

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

We will now show that

$$\| (U_2 v, U_1 w) - (U_2 v', U_1 w') \| \leqslant \delta \| (v,w) - (v',w') \| \tag{17}$$

Since $w' + \| w - w' \| \geqslant w$ (and similarly, $w + \| w - w' \| \geqslant w'$), from the monotonicity of the operator U_1 and Eq. 16 it follows that

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

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

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

and hence, $\delta \| v - v' \| \geqslant \| U_2 v - U_2 v' \|$. The last two norm inequalities clearly prove Eq. 17. Hence, the operators are contractions and consequently have a unique fixed point. The claim is proved. ■

By Lemma 8, SPE value sequences are time invariant, that is, they depend on the state (or action) a_i or a_j.

5.1.3 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 and Stacchetti (1990) for discounted Repeated Games 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 $v_j^*, w_i^* \in (V_j^*, W_i^*)(\bar{a}_i)$, i.e., when player i contemplates an action $\bar{a}_i$, these are SPE payoffs that arise based on equilibrium strategies - say, $\sigma_i^*(\bar{a}_i)$ and $\sigma_j^*(\bar{a}_i)$.[26] In particular, the pair of best response strategies imply a concurrent transfer schedule $\theta_i^*(\cdot)$ (picked by i) and an action a_j^* (picked by j) and continuation payoffs $w_j(a_j)$. Since this is a SPE, continuation payoffs do not depend on the transfer scheme $\theta_i^*(\cdot)$ picked at $t-1$ (since that is not part of the public history at $t+1$, the initial period for the continuation $w_j(a_j)$).

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

$$a_j^* \in \arg\max\{(1-\delta)[\pi_j(\bar{a}_i, a_j) + \theta_i^*(a_j)] + \delta w_j(a)\} \tag{18}$$

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

a) a_j' solves Eq. 18 for $\theta_i'(\cdot)$ and $w_j(a_j)$,and that

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

$$(1-\delta)[\pi_i(\bar{a}_i, a_j') - \theta_i'(a_j')] + \delta v_i(a_j') > (1-\delta)[\pi_j(\bar{a}_i, a_j^*) - \theta_i^*(a_j^*)] + \delta v_i(a_j^*) \tag{19}$$

[26] Note that if this is play commencing after a history, say h_t, then the strategies $\sigma_i^*(\bar{a}_i)$ and $\sigma_j^*(\bar{a}_i)$ might be indexed by the history and should be written more completely as $\sigma_i^*(\bar{a}_i; h_t)$ and $\sigma_j^*(\bar{a}_i; h_t)$. 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.

Clearly, from the definition it follows that $v_j^*, w_i^* \in \Gamma(V_j^*, W_i^*)(\bar{a}_i)$. However, we have already seen that $\Gamma(V_j^*, W_i^*)(\bar{a}_i) = T(V_j^*, W_i^*)(\bar{a}_i)$. 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^∞.

5.2 Equilibrium transfers

We show that equilibrium values V_i, W_i are uniformly bounded in discount rates and that this property implies the uniform boundness in δ of the equilibrium transfers.

Proposition 10 *Equilibrium transfers are uniformly bounded in discount factors δ.*

Proof: Through out the argument, we fix $\delta \in (0, 1)$ and show that the values of the variables of interest are bounded above and below by values independent of δ thereby implying uniform boundness. We start by showing a preliminary result stated under the form of a Claim.

Claim 11 *The equilibrium values V_i, W_i are uniformly bounded below and above in δ.*

Proof Both the mover and the non-mover can choose to not make positive transfers and to play any fix action in each period. Therefore, V_i, W_i are uniformly (on δ) bounded below by the minimum payoff of the stage game, $\pi_i^m = \min_{a_i, a_j} \pi_i(a_j, a_i)$. By Theorem 2, $V_i(a_j) + W_j(a_j) = U(a_j)$, all $a_j \in A_j$ and $j, i,\ i \neq j$. Denote by Π the aggregate payoff of the stage game, that is, $\Pi = \pi_i + \pi_j$. The UPO value $U(a_j)$ is bounded above by Π^M, the maximum aggregate payoff of the stage game. Hence the values V_i, W_i are bounded above by $\Pi^M - \pi_j^m$. ■

Let σ denote the UPO strategy, a map from A_j to A_i, $i \neq j$, associating the action a_j with the next period action σa_j. By the definition of the operator T, Eq. (5), on path transfers satisfy

$$\theta_j(\sigma a_j) = \frac{V_i(a_j) - \delta W_i(\sigma a_j)}{1 - \delta} - \pi_i(a_i, \sigma_i a_j),$$

while off equilibrium transfers are indeterminate and they satisfy the same condition as a weak inequality

$$\theta_j(a_i) \leq \frac{V_i(a_j) - \delta W_i(a_i)}{1-\delta} - \pi_i(a_i, a_j),\ a_i \neq \sigma a_j.$$

Therefore, as $\theta \geq 0$, we just need to show that $\frac{V_i(a_j)-\delta W_i(a_i)}{1-\delta}$ is uniformly bounded above for all δ.

By the definition of the operator T, the equilibrium values V_i satisfy

$$V_i(a_j) = \max_{a_i}\{(1-\delta)\pi_i(a_i, a_j) + \delta W_i(a_i)\}$$

and we call $\phi_i(a_j)$ (one of) the optimizer. Then by definition of V_i, the following inequalities hold true:

$$\frac{V_i(a_j) - \delta W_i(a_i)}{1-\delta} \geq \pi_i(a_i, a_j), \text{ all } (a_i, a_j). \tag{I}$$

Moreover, as $V_j(a_i) = U(a_i) - W_i(a_i)$, we get the following system of equations in the variables V_i, W_i

$$\frac{V_i(a_j) - \delta W_i(\phi_i(a_j))}{1-\delta} = \pi_i(\phi_i(a_j), a_j) \tag{V}$$

$$\frac{W_i(a_i) - \delta V_i(\phi_j(a_i))}{1-\delta} = -\pi_j(a_i, \phi_j(a_i)) + \frac{U(a_i) - \delta U(\phi_j(a_i))}{1-\delta} \tag{W}$$

Therefore, by equations (V) and (W), it is:

$$\frac{V_i(a_j) - \delta W_i(a_i)}{1-\delta} = \delta\frac{W_i(\phi_i(a_j)) - \delta V_i(\phi_j(a_i))}{1-\delta} +$$

$$+\pi_i(\phi_i(a_j), a_j) + \delta\pi_j(a_i, \phi_j(a_i)) - \delta\frac{U(a_i) - \delta U(\phi_j a_i)}{1-\delta}$$

$$= \frac{\delta W_i(\phi_i(a_j)) - V_i(\phi_j(a_i))}{1-\delta} + (1+\delta)\delta V_i(\phi_j(a_i))$$

$$+\pi_i(\phi_i(a_j), a_j) + \delta\pi_j(a_i, \phi_j(a_i)) - \delta\frac{U(a_i) - \delta U(\phi_j(a_i))}{1-\delta}$$

Exploiting inequality (I), we get:

$$\frac{V_i(a_j) - \delta W_i(a_i)}{1-\delta} \leq -\delta\pi_j(\phi_i(a_j), \phi_j(a_i)) + (1+\delta)V_i(\phi_j(a_i))$$

$$+\pi_i(\phi_i(a_j), a_j) + \delta\pi_j(a_i, \phi_j(a_i)) - \delta\frac{U(a_i) - \delta U(\phi_j(a_i))}{1-\delta}$$

By definition of Utilitarian Pareto optimum, it is

$$U(a_i) = (1-\delta)\Pi(\sigma a_i, a_i) + \delta U(\sigma a_i) \geq (1-\delta)\Pi(\phi_j a_i, a_i) + \delta U(\phi_j a_i; \delta).$$

The latter together with Claim 11 implies the thesis. ■

5.3 Completing the Proof for Finite Transfer Histories

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

Lemma 12 *The conclusion that there is a unique SPE which coincides with the UPO solution 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 $h_t^\theta = (\theta_{jt-M}(\cdot), ..., \theta_{jt-1}(\cdot))$. Recall also the "Stackelberg leader" player $i's$ optimal transfer problem, i.e.,

$$w_i(h_t, a_{it}) = \max_{\theta_i(\cdot)}\{(1-\delta)[\pi_i(a_{it}, a_j(h_{t+1})) - \theta_i(a_j)] + \delta v_i(h_{t+2})\} \tag{20}$$

where h_{t+1} and h_{t+2} are the histories at $t+1$ and $t+2$ caused by $i's$ period t actions, $h_{t+1} = (h_t, a_{it}, \theta_{it}(\cdot))$ and $h_{t+2} = (h_{t+1}, a_{jt+1}, \theta_{jt+1}(\cdot))$. In the $t+1$ history, the follower player $j's$ best response action in period $t+1$, $a_j(h_{t+1})$, comes from:[27]

$$v_j(h_{t+1}) = \max_{a_j}\{[\pi_j(a_{it}, a_j) + \theta_{it}(a_j)] + \delta w_j(h_{t+1}, a_j)\}. \tag{21}$$

Hence, from Eq. 21 it is clear that $a_j(h_{t+1})$ does not contain any reference to $\theta_{jt-M}(\cdot)$ since, by period $t+1$ when that action is chosen, $\theta_{jt-M}(\cdot)$ is no longer in the finite transfer memory. In particular, $w_i(h_t, a_{it})$ is independent of $\theta_{jt-M}(\cdot)$. It then follows, from the definition of player $i's$ lifetime payoffs starting at time t,

$$v_i(h_t) = \max_{a_i}\{(1-\delta)[\pi_i(a_i, a_{jt-1}) + \theta_{jt-1}(a_i)] + \delta w_i(h_t, a_i)\} \tag{22}$$

that $v_i(h_t)$ is also independent of $\theta_{jt-M}(\cdot)$. By the same logic, the payoffs for player j starting in the subsequent period $t+1$, $v_j(h_{t+1})$, must be independent of $\theta_{it-M+1}(\cdot)$. Then, from Eq. 21, it follows that player $j's$ best response action in period $t+1$, $a_j(h_{t+1})$, which is the solution of Eq. 21, must itself be independent of $\theta_{it-M+1}(\cdot)$. But, invoking Eqs. 20 and 22, that implies that $w_i(h_t, a_{it})$ and $v_i(h_t)$ must both be independent of $\theta_{it-M+1}(\cdot)$.

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 $\theta_{it-M+1}(\cdot)$, then $v_j(h_{t+1})$ and $a_j(h_{t+1})$ must be independent of $\theta_{it-M+2}(\cdot)$ thereby implying that $w_i(h_t, a_{it})$ and $v_i(h_t)$ are in fact independent of $\theta_{it-M+2}(\cdot)$. And so on. Leading 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 $\theta_{jt-1}(\cdot)$ enters v_i - and only through the payoff relevant effect in the first term of Eq. 22. Put another way, the general M case has been reduced to $M = 1$. The proof of the previous sub-section then applies. ■

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

6 An Example with Infinite Transfer History

If infinite transfer histories are allowed, it is possible to generate a Folk Theorem - multiple equilibria some of which are Pareto suboptimal. The example below illustrates this point.

Consider the following stage game

$1\backslash 2$	A	B
a	$(0,0)$	$(1,0)$

Note that in this (simple) example, Player 1 has no action and only offers a transfer (to induce Player 2 to play her preferred action B). Consequently, Player 2 can be thought of, without loss, as offering no transfer (and instead just as a provider of an action). We use the following convention: a) at T even, $T = 0, 2, \ldots$, Player 2 selects an action from $\{A, B\}$ given history $h_T = (\theta_{-1}, a_0, \theta_1 \ldots a_t, \theta_{t+1}, ..a_{T-1}, \theta_T)_{t \leq T}$; at T odd, Player 1 selects a transfer θ_{T+1} after a similar history. Let $h_{T/t}$ denote a history of play between periods $t+1$ and T.

Consider the following strategies defined by three transfers (α, β, γ) with $1 > \alpha > \beta > \gamma > 0$.

Norm: Player 1 starts by offering the transfer schedule $\theta^n = (0, \beta)$ - a positive transfer β paid iff the subsequent period's action is B - and continues to do so if the history is $h_T = (\theta^n, B, \ldots, \theta^n, B)$. Player 2 plays B provided that same history has been observed.

Punishment 1: Upon the first deviation from the Norm, say at date t, if it is by Player 1, Player 2 plays A. Thereafter, 1 plays $\theta^1 = (0, \alpha)$ and continues to do so provided the remaining history is $h_{T/t} = (\theta^1, B, \ldots, \theta^1, B)$. 2 plays B from period $t+3$ onwards, provided that same remaining periods history $h_{T/t}$ has been observed. In the event that there is a deviation by Player 1 (in period $t+2, t+4$...) *Punishment* 1 is re-started. In the event that there is a deviation by 2 (in period $t+1$, $t+3$...) *Punishment* 2 is started.

Punishment 2: Upon the first deviation from the Norm, say at date t, if it is by Player 2, Player 1 plays $\theta^2 = (0, \gamma)$. Thereafter, 1 plays θ^2 and continues to do so provided the remaining history is $h_{T/t} = (\theta^2, B, \ldots, \theta^2, B)$. Player 2 plays B from period $t+2$ onwards, provided that same remaining periods history $h_{T/t}$ has been observed. In the event that there is a deviation by 1 (in period $t+1$, $t+3$...) *Punishment* 1 is started. In the event that there is a deviation by 2 (in period $t+2$, $t+4$...) *Punishment* 2 is re-started.

It should be clear that all histories have associated player actions, i.e., we have well-defined strategies. We now check incentives. Since we are going to be interested only in high δ scenarios, it will suffice to consider undiscounted streams and show that there are strict incentives to not deviate at $\delta = 1$. Furthermore, we shall check incentives in "backward induction" fashion by first checking the incentives in the *Punishment* phases and we shall check for one-shot deviations only.

Checking for Best Responses

Punishment 1 Subgame: If player 1 deviates by picking $(\theta_A, \theta_B) \neq \theta^1$, she gets the stream of payoffs $0-\theta_A, 0, 1-\alpha, 1, 1-\alpha, 1, ...$ whereas by not deviating she gets the stream $1-\alpha, 1, 1-\alpha, 1, ...$ The latter is preferred if $2-\alpha > -\theta_A$. Since $\theta_A \geqslant 0$ and $\alpha \leqslant 1$, evidently not deviating is preferable.

If player 2 deviates at t,[28] then from $t+3$ onwards his payoff stream is γ, 0, γ, 0, ... while by not deviating the payoff stream is α, 0, α, 0, ... The latter is "infinitely bigger" than the former, as $\delta \uparrow 1$ (since $\alpha > \gamma$) and hence swamps any payoff comparison at $t+1$ and makes deviating the worse option. Hence, in the *Punishment* 1 subgame, the strategies are indeed equilibria.

Punishment 2 Subgame: If player 1 deviates at t by picking $(\theta_A, \theta_B) \neq \theta^2$, she gets the stream of payoffs $0-\theta_A, 0, 1-\alpha, 1, 1-\alpha, 1, ...$ whereas by not deviating she gets the stream $1-\gamma, 1, 1-\gamma, 1, ...$ The latter is "infinitely bigger" than the former, as $\delta \uparrow 1$ (since $\alpha > \gamma$) and hence swamps any payoff comparison at $t+1$ and makes deviating the worse option.

If player 2 deviates, then his payoff stream is $0, 0, \gamma, 0, \gamma, 0, ...$ while by not deviating it is $\gamma, 0, \gamma, 0, ...$ Since $\gamma > 0$, clearly, deviating is the worse option. Hence, in the *Punishment* 2 subgame, the strategies are indeed equilibria.

Norm Subgame: If player 1 deviates by picking $(\theta_A, \theta_B) \neq \theta^n$, she gets the stream of payoffs $0-\theta_A, 0, 1-\alpha, 1, 1-\alpha, 1, ...$ whereas by not deviating she gets the stream $1-\beta, 1, 1-\beta, 1, ...$ The latter is "infinitely bigger" than the former, as $\delta \uparrow 1$ (since $\alpha > \beta$) and hence makes deviating the worse option.

If player 2 deviates, then his payoff stream is $0, 0, \ \gamma, 0, \gamma, 0, ...$ while by not deviating the payoff stream is $\beta, 0, \beta, 0, ...$ Since $\beta > \gamma$, clearly, deviating is the worse option. Hence, in the *Norm* subgame, the strategies are indeed equilibria. More generally, we have shown that the given strategies are SPE.

Note that the lifetime payoffs are $(\frac{1-\beta+\delta}{1+\delta}, \frac{\beta}{1+\delta})$. Since β is arbitrarily chosen from $(0,1)$, we have shown that there are an uncountable number of Pareto optimal SPE in this game. It should be clear though that the construction can be repeated using an alternative norm and alternative punishment regimes in which, say, on the norm path, player 2 plays A with probability $1-\rho$ and plays B with probability ρ and gets a transfer β upon playing B. The associated lifetime payoffs are $\rho(\frac{1-\beta+\delta}{1+\delta}, \frac{\beta}{1+\delta})$. We have hence shown in this example:

Proposition 13 *Any payoff vector* $\rho(\frac{1-\beta+\delta}{1+\delta}, \frac{\beta}{1+\delta})$, $\rho \in (0,1), \beta \in (0,1)$ *is a SPE payoff vector in this game.*

Note that this construction works when players' memories about transfers are infinite, but it fails otherwise. Indeed suppose that $T < \infty$ is the length of the transfers' memory and take as given a history $\bar{h}$ putting the play in Punishment 1. Consider the strategies where player 1 selects for T periods θ^2 (rather than θ^1), while player 2 selects (consistently with the equilibrium strategy outlined above) A. After T periods these strategies generate the history

[28] Note that a deviation by Player 2 means either a) picking B instead of A, if it is the first period after a deviation from the *Norm*, or b) picking A instead of B, if it is the first period after a deviation by Player 1 from one of *Punishment* 1 or *Punishment* 2 regimes.

$h' = (\bar{h}, \theta^2, A, ..., A, \theta^2)$. As T is the length of transfers' memory, players cannot rule out the possibility that the history h' is generated by player 2 deviating by selecting A rather than B and player 1 consistently with *Punishment* 2 punishing her by choosing θ^2.

7 Extensions

7.1 N Players

To keep notation to a minimum, we analyze the three player game in some detail. The argument will be seen to extend straightforwardly to any number $N \geq 3$ of players (and further details may be found in Dutta and Siconolfi (2016a)).

Let G now denote a three player stage game (in strategic form) with strategy sets A_i, $A \equiv A_1 \times A_2 \times A_3$ and payoff functions $\pi_i : A \to \mathbb{R}$, $i = 1, 2, 3$. Let i, j, k denote the generic players. Asynchronocity will continue to be central: action and transfer choices will be in sequence. Notationally, we will say that the sequence of action choices is: i followed by j followed by k - and then repeated.

Transfer commitments will continue to be one-period ahead commitments. So, if, say, Player i has a move in the current period t, then Players j and k make transfer commitments $\theta_{ji}(\cdot)$ and $\theta_{ki}(\cdot)$ in (sequence) in the previous period, $t-1$. Without loss, we will assume that the sequence of transfer commitments follow the sequence of action choices; since the choice a_j (was made in period $t-2$) and preceded a_k (chosen in $t-1$), hence the choice $\theta_{ji}(\cdot)$ precedes $\theta_{ki}(\cdot)$.[29] Put another way, in period $t-1$, prior to $i's$ move in period t, three choices are made in the following sequence: a_k, followed by $\theta_{ji}(\cdot)$ followed by $\theta_{ki}(\cdot)$:

time	t_-	$t_=$	t_+	$t+1$
choices	a_k	$\theta_{ji}(\cdot)$	$\theta_{ki}(\cdot)$	a_i

where period t has been divided into three subperiods with t_- preceding $t_=$ that precedes t_+.

If, at time t, it is player $i's$ turn to move, she can choose an action a_{it} from A_i - and this action then remains fixed till $t+3$. Denoting by $\bar{a}_{-it} = (a_{jt-2}, a_{kt-1})$ the (fixed) actions of the other players, the period t stage game payoffs are $\pi_i(a_t)$, $\pi_j(a_t)$, $\pi_k(a_t)$, where $a_t = a_{it}, \bar{a}_{-it}$. Hence, the payoffs at t of player i inclusive of transfers are

$$\pi_i(a_t) + \theta_{ji,t-1}(a_{it}) + \theta_{ki,t-1}(a_{it})$$

while those of players j and k are

$$\pi_k(a_t) - \theta_{ki,t-1}(a_{it}), \pi_j(a_t) - \theta_{ji,t-1}(a_{it})$$

If i is the initial period mover, then there is an *initial action state* $\bar{a}_j, \bar{a}_k$ for the game, which can be arbitrary. As in the two-player model, the *initial*

[29] The result is the same if the reverse ordering is used - choice $\theta_{ki}(\cdot)$ precedes $\theta_{ji}(\cdot)$.

transfer state $\overline{\theta}_{ki}, \overline{\theta}_{ji}$ is chosen in sequence by the non-movers in the "pre-period" -1.

7.1.1 Strategies and Equilibrium

As in the two player game, action histories h_t^a include all past actions, while transfer histories h_t^θ include M past transfer commitments.

The action mover, player i, makes an action choice a_{it} conditional on h_t. Thereafter, two transfer commitments - to next period's mover j - are made in the sequence $\theta_{kj,t}$ and $\theta_{ij,t}$ conditional, respectively, on (h_t, a_{it}) and $(h_t, a_{it}, \theta_{kj,t})$. Given strategies $\sigma = (\sigma_i, \sigma_j, \sigma_k)$, and for given history h_t, we denote by h_{t+n}^σ the history of length $t+n$ generated by σ following h_t. Denote the lifetime payoff of the mover at time t, v_i:

$$v_i(h_t) = \max_{a_i}\{(1-\delta)[\pi_i(a_i, a_{-it}) + (\theta_{ji,t-1} + \theta_{ki,t-1})(a_i)] + \delta r_i(h_t, a_i)\}$$

where r_{it+1} is

$$r_{it+1}(h_t, a_i) = \max_{\theta_{ij}}\{(1-\delta)[\pi_i(a_j(h_{t+1}^\sigma), a_{-j,t+1}) - \theta_{ij}(a_j(h_{t+1}^\sigma))] + \delta w_{it+2}(h_{t+2}^\sigma)\}$$

and w_{it+2} is

$$w_{it+2}(h_{t+2}^\sigma) = \max_{\theta_{ik}}\{(1-\delta)[\pi_i(a_k(h_{t+2}^\sigma), a_{-k,t+2}) - \theta_{ik}(a_k(h_{t+2}^\sigma))] + \delta v_{it+3}(h_{t+3}^\sigma)\}.$$

The **Utilitarian Pareto Optimum** problem is the maximization of the lifetime sum of players payoffs

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

subject to an initial action state $(\bar{a}_j, \bar{a}_k)$ and that player i can change her action in periods $0, 3, ...$ while player j can do so in periods $1, 4, ...$ and, finally, player k moves in periods $2, 5, ...$

It is well known that the solution is Markovian, i.e., the mover's action at t depends on the actions chosen by the non-movers at $t-1$ and $t-2$.

The **main theorem** is:

Theorem 14 *The SPE of the game coincides with the Utilitarian Pareto Optimum solution in terms of actions. SPE values, inclusive of transfers, are uniquely determined.*

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

For the proof, we mimic the argument for the two player game. We define, in the context of three players, the operators Γ and T when $M = 1$. Their equivalence follows from the analysis in Dutta and Siconolfi (2016a). Once that equivalence is established, it is easily verified that the remaining argument for the three players game is an obvious adaptation of the proof for the two players case and is therefore omitted.

7.1.2 The Operators Γ and T and their Equivalence

The principle of the construction is worth a brief explanation. In the two player game, recall, the player picking the transfer (in the immediately preceding period) acts like a Stackleberg leader. She picks the transfer to maximize her payoff subject to the subsequent period's mover picking a best response action for every transfer chosen. This is reflected in the construction of the operator Γ. Moreover, the leader holds the subsequent mover to his best payoff absent transfers, i.e., holds him to a constant irrespective of what action he chooses. Thus, the leader has an incentive to maximize the sum of two players' payoffs since her net is maximized if and only if the sum is maximized. This is the logic behind the construction of the equivalent operator T.

In the three player game, there are two players picking transfers in the previous period, in sequence. The first of them acts like a Stackleberg leader picking his transfer subject to *two* constraints: i) that the other transfer chooser - the middle player - will select a best response transfer, and ii) that the mover next period - the follower - will select a best response action to both transfers. The middle player takes as given the leader's transfer commitment. All this is reflected in the construction of the operator Γ. In particular the leader holds the middle player and the follower to their best sum of payoffs absent transfers, i.e., she holds them to a constant irrespective of what transfer and action they subsequently choose. Consequently, the leader's choice of transfer maximizes the (three player) sum of payoffs since her net is maximized if and only if the sum is maximized. The second player picking a transfer, typically but not always, acts as the residual transferee getting the desired utilitarian action taken by paying the balance of the aggregate transfer. This is the logic behind the construction of the equivalent operator T.

Unfortunately with more than two players, the definition of operators Γ and T requires some additional notation. As before, let V_i, W_j and R_k be correspondences of potential SPE continuation payoffs with domain $A_j \times A_k$ (and range $\mathbb{R}$). Pick $\bar{a}_{-k} = (\bar{a}_i, \bar{a}_j) \in A_i \times A_j$ and continuation payoffs $(w_i, v_j, r_k)(a_k) \in (W_i, V_j, R_k)(\bar{a}_j, a_k)$, $a_k \in A_k$. For given $\bar{\theta}_{ik}, \bar{\theta}_{jk}$, let $Z_k(\bar{a}_{-k}, \bar{\theta}_{ik}, \bar{\theta}_{jk}, r_k) \subset A_j$ be the set of solutions to:

$$\max_{a_k}\{(1-\delta)[\pi_k(a_k, \bar{a}_{-k}) + (\bar{\theta}_{ik} + \bar{\theta}_{jk})(a_k)] + \delta r_k(\bar{a}_j, a_k)\};$$

while for given $\bar{\theta}_{ik}$, let $Z_j(\bar{a}_{-k}, \bar{\theta}_{ik}, w_j, r_k) \subset A_j \times \mathbb{R}_+^{A_k}$ be the set of solutions to:

$$\begin{aligned} &\max_{a_k, \theta_{jk}(\cdot)}\{(1-\delta)[\pi_j(a_k, \bar{a}_{-k}) - \theta_{jk}(a_k)] + \delta w_j(\bar{a}_j, a_k)\}, \\ \text{s. t. } a_k \quad &\in \quad Z_k(\bar{a}_{-k}, \bar{\theta}_{ik}, \theta_{jk}, r_k); \end{aligned}$$

and finally let $Z_i(\bar{a}_{-k}, w_j, v_i, r_k) \subset A_j \times \mathbb{R}_+^{A_k} \times \mathbb{R}_+^{A_k}$ be the set of solutions to

$$\begin{aligned} &\max_{a_\kappa, \theta_{jk}, \theta_{ik}}\{(1-\delta)[\pi_i(a_k, \bar{a}_{-k}) - \theta_{ik}(a_k)] + \delta v_i(\bar{a}_j, a_k)\}, \\ \text{s. t. } (a_k, \theta_{jk}) \quad &\in \quad Z_j(\bar{a}_{-k}, \theta_{ik}, v_j, r_k), \end{aligned}$$

Notice that by construction if $(a_k^*, \theta_{jk}^*, \theta_{ik}^*) \in Z_i(\bar{a}_{-k}, v_i, w_j, r_k)$ then $(a_k^*, \theta_{jk}^*) \in Z_j(\bar{a}_{-k}, \theta_{ik}^*, w_j, r_k)$ and $a_k^* \in Z_k(\bar{a}_{-k}, \theta_{ik}^*, \theta_{jk}^*, r_k)$.

Define the operator Γ :

$$\begin{aligned}
\Gamma(V_i, W_j, R_k)(\bar{a}_{-k}) &= \{w_i, r_j, v_k : \exists(v_i, w_j, r_k)(a_k) \in (V_i, W_j, R_k)(\bar{a}_j, a_k) \\
&\quad \text{such that} \\
w_i &= (1-\delta)[\pi_i(a_k^*, \bar{a}_{-k}) - \theta_{ik}^*(a_k^*)] + \delta v_i(\bar{a}_j, a_k^*) \\
r_j &= (1-\delta)[\pi_i(a_k^*, \bar{a}_{-k}) - \theta_{jk}^*(a_k^*)] + \delta w_j(\bar{a}_j, a_k^*) \\
v_k &= \{(1-\delta)[\pi_k(a_k^*, \bar{a}_{-k}) + (\theta_{ik}^* + \theta_{jk}^*)(a_k^*)] + \delta r_k(a_k^*), \\
\text{for } (a_k^*, \theta_{ik}^*, \theta_{jk}^*) &\in Z_i(\bar{a}_{-k}, v_i, w_j, r_k)\}
\end{aligned}$$

Define on the same domain of Γ, the operator T :

$$\begin{aligned}
T(V_i, W_j, R_k)(\bar{a}_{-k}) &= \{w_i, r_j, v_k : \exists(v_i, w_j, r_k)(a_k) \in (V_i, W_j, R_k)(\bar{a}_j, a_k) \\
&\quad \text{such that}
\end{aligned}$$

$$\begin{aligned}
v_k &= v_k^* + \max\{0, r_j^* - [(1-\delta)\pi_j(a_k^*, \bar{a}_{-k}) + \delta w_j(\bar{a}_j, a_k^*)]\}, \\
r_j + v_k &= r_j^* + v_k^*,
\end{aligned}$$

$$\begin{aligned}
w_i + r_j + v_k &= \max_{a_k}\{(1-\delta)(\pi_i + \pi_j + \pi_k)(a_k, \bar{a}_{-k}) + \delta(v_i + w_j + r_k)(\bar{a}_j, a_k)\}, \\
\text{for } v_k^* &= \max_{a_k}\{(1-\delta)\pi_k(a_k, \bar{a}_{-k}) + \delta r_k(a_k)\},
\end{aligned}$$

$$\begin{aligned}
r_j^* + v_k^* &= \max_{a_k}\{(1-\delta)(\pi_j + \pi_k)(a_k, \bar{a}_{-k}) + \delta(w_j + r_k)(\bar{a}_j, a_k)\}, \\
a_k^* &\in \arg\max_{a_k}\{(1-\delta)(\pi_i + \pi_j + \pi_k)(a_k, \bar{a}_{-k}) + \delta(v_i + w_j + r_k)(\bar{a}_j, a_k)\}.
\end{aligned}$$

The equivalence of the two operators is an immediate consequence of Lemma 16 and 19 in Dutta and Siconolfi (2016a) and it is stated next.

Lemma 15 *Given correspondences* (V_i, W_j, R_k), (V_k, W_i, R_j) *and* (V_j, W_k, R_i), *it is* $\Gamma(V_i, W_j, R_k)(\bar{a}_{-k}) = T(V_i, W_j, R_k)(\bar{a}_{-k})$, *for all* $\bar{a}_{-k}$. *Similarly, it is* $\Gamma(V_k, W_i, R_j)(\bar{a}_{-j}) = T(V_k, W_i, R_j)(\bar{a}_{-j})$, *for all* $\bar{a}_{-j}$ *and* $\Gamma(V_j, W_k, R_i)(\bar{a}_{-i}) = T(V_j, W_k, R_i)(\bar{a}_{-i})$, *for all* $\bar{a}_{-i}$.

The operator T of the three players game enjoys the same properties of the operator T of the two players game. In particular, it is easily verified that the obvious adaptations of Lemmas 4, 6, 7, 9 and 12 hold true. Thus, Theorem 14 holds true.

7.2 Asynchronous Game

We conjecture that the results proved in this paper work for any asynchronous game, not just the specific case of alternating moves that we have analyzed here. The companion paper Dutta and Siconolfi (2016.a) provides a fairly general definition of asynchronous games which is independent of whether the time horizon is finite or infinite. Dutta and Siconolfi (2016.a) argues that payoffs and actions of asynchronous games coincide with payoffs and actions of the associated UPO whenever the time horizon is finite. When the time horizon is infinite and the asynchronicity has a recursive structure it seems obvious that the analysis goes through. Without a recursive structure, the same result may well hold. However, the argument is by far nontrivial and it is left to further research.

References

[1] D. Abreu, D. Pearce, and E. Stacchetti, Towards a Theory of Discounted Repeated Games with Imperfect Monitoring, 1990, Econometrica 58(5): 1041-1063

[2] A.R. Admati and M. Perry, Strategic Delay in Bargaining, 1987, Review of Economic Studies 54(3): 345-364

[3] JP. Benoit and V. Krishna, Finitely Repeated Games, 1985, Econometrica 53(4): 905-922

[4] D. Blackwell, Discounted dynamic programming, 1965, The Annals of Mathematical Statistics 36(1): 226-235

[5] P. Dutta, A Folk Theorem for Stochastic Games, 1995, Journal of Economic Theory 66(1): 1-32

[6] P. Dutta, Coordination Need Not Be a Problem, 2012, Games and Economic Behavior 76(2): 519-534

[7] P. Dutta, Compromise is Key to Repeated Bargaining, 2016, mimeo

[8] P. Dutta and P. Siconolfi, Finite Horizon Asynchronous Games with Transfers: A Welfare Theorem, 2017(a), mimeo

[9] P. Dutta and P. Siconolfi, Prisoners Dilemma with Simultaneous Transfers, 2016(b), mimeo

[10] D. Fudenberg and E. Maskin, The Folk Theorem in Repeated Games with Discounting or with Incomplete Information, 1986, Econometrica 54(3): 533-554

[11] D. Fudenberg, D. Levine and E. Maskin, The Folk Theorem with Imperfect Public Information, 1994, Econometrica 62(5): 997-1039

[12] D. Fudenberg and Y. Yamamoto, The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring, 2011, Journal of Economic Theory 146(4): 1664-1683

[13] S. Goldücke and S. Krantz, Infinitely Repeated Games with Public Monitoring and Monetary Transfers, 2012, Journal of Economic Theory 147(3): 1191-1221

[14] J. Guttman, "Understanding Collective Action: Matching Behavior," 1978, American Economic Review 68 (2): 251–255

[15] C. Harris, Existence and Characterization of Perfect Equilibrium in Games of Perfect Information, 1985, Econometrica 53(3): 613-628

[16] J. Horner and W. Olszewski, The Folk Theorem for Games with Private Almost-Perfect Monitoring, 2006, Econometrica 74(6): 1499-1545

[17] J. Horner, T. Sugaya, S. Takahashi, and N. Vieille, Recursive Methods in Discounted Stochastic Games: An Algorithm for $\delta \rightarrow 1$ and a Folk Theorem, 2011, Econometrica 79(4): 1277–1318

[18] M. Jackson and S.Wilkie, Endogenous Games and Mechanisms: Side Payments among Players, 2005, Review of Economic Studies 72(2): 543-566

[19] Y. Kamada, and T. Sugaya, Asynchronous Revision Games with Deadline: Unique Equilibrium in Coordination Games, 2010, mimeo

[20] J. Kemmerer, and J. Lu, "Profitability and royalty rates across industries: Some preliminary evidence." 2012, Research Report, Global Valuation Institute, KPMG International

[21] R. Lagunoff, and A. Matsui, Asynchronous Choice in Repeated Coordination Games, 1997, Econometrica 65(6): 1467-1478

[22] M. Pęski and T. Wiseman, A Folk Theorem for Stochastic Games with Infrequent State Changes, 2015, Theoretical Economics 10(1): 131-173

[23] E. Maskin and J. Tirole, A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed, 1988, Econometrica 56(3): 549-569

[24] E. Maskin and J. Tirole, A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles, 1988, Econometrica 56(3): 571-599

[25] L. Marx and S. Matthews, Dynamic Voluntary Contribution to a Public Project, 2000, Review of Economic Studies 67(2): 327-358

[26] E. Maskin, and J. Tirole, A Theory of Dynamic Oligopoly, I: Overview and Quantity Competition with Large Fixed Costs, 1988a., Econometrica 56 (3): 549-569

[27] E. Maskin, and J. Tirole, A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles, 1988b, Econometrica 56 (3): 571-599

[28] A. Rubinstein, Perfect Equilibrium in a Bargaining Model, 1982, Econometrica 50 (1): 97-109

[29] L. Smith, Folk Theorems in Overlapping Generations Games, 1992, Games and Economic Behavior 4(3): 426–449

[30] T. Sugaya, Folk Theorem in Repeated Games with Private Monitoring, 2016, mimeo

[31] S. Takahashi, Infinite Horizon Common Interest Games with Perfect Information, 2005, Games and Economic Behavior 53 (2): 231-247

[32] H. Varian, Sequential Provision of Public Goods, 1994, Journal of Public Economics 53 (2): 165–186.

[33] K. Yoon, Folk Theoreme for Asynchronously Repeated Games, 2001, Econometrica 69(1): 191-200