# The Folk Theorem for Irreducible Stochastic Games with Imperfect Public Monitoring*

Drew Fudenberg and Yuichi Yamamoto
Department of Economics, Harvard University†

First version: November 15, 2009
This version: March 5, 2010

**Abstract**

In an irreducible stochastic game, no single player can prevent the stochastic process on states from being irreducible, so the other players can ensure that the current state has little effect on events in the distant future. This paper introduces stochastic games with imperfect public signals, and provides a sufficient condition for the folk theorem when the game is irreducible, thus generalizing the folk theorems of Dutta (1995) and Fudenberg, Levine, and Maskin (1994). To prove this theorem, the paper extends the concept of self-generation (Abreu, Pearce, and Stachetti (1990)) to "return generation," which explicitly tracks actions and incentives until the next time the state returns to its current value, and asks that players not wish to deviate given the way their continuation payoffs from the time of this return depend on the public signals that have been observed.

Keywords: stochastic game, public monitoring, perfect public equilibrium, folk theorem.

*We thank Johannes Hörner, Scott Kominers, and Satoru Takahashi for helpful conversations, and Yuhta Ishi and Lauren Merril for comments on a previous draft.

†Littauer Center, Harvard University, 1805 Cambridge Street, Cambridge, MA 02138

# 1 Introduction

Most social and economic interactions occur repeatedly, and the fact that agents can condition their current play on information about past outcomes makes it possible to support outcomes that are not equilibria of one-shot interactions. In particular, the folk theorems for discounted repeated games show, roughly speaking, that any feasible individually rational payoffs can be generated by an equilibrium if the players are sufficiently patient. These theorems hold both in the classic observed-actions case and also when players receive imperfect public signals of one another's actions.

Stochastic games (Shapley (1953)) generalize repeated games with observed actions by allowing each period's payoff functions to depend on a state variable whose evolution can be influenced by the players' actions; the state variable can capture intertemporal links such as technological innovations, persistent demand shocks, savings, and capital stocks. When there is an irreversible component to the evolution of the state, single deviations can have permanent effects on the sets of feasible and individually rational payoffs, and the structure of stochastic-game equilibria can be very different than that of repeated games. Conversely, if no single player can make the stochastic process irreversible- that is, when the other players have a strategy that makes the process irreducible- then the feasible, individually rational discounted average payoffs converge to a limit that is independent of the current state as the discount factor converges to 1. Dutta (1995) establishes a folk theorem for these irreducible stochastic games.

This paper introduces the class of stochastic games with imperfect public monitoring, where players observe the state and public signal that is related to the actions played, and shows that when the game is irreducible the folk theorem applies. Our proof is based on the extension of self-generation (Abreu, Pearce, and Stachetti (1990, hereafter APS)) to "return-generation," and on extensions of the full-rank conditions of Fudenberg, Levine, and Maskin (1994, hereafter FLM). The idea of return-generation is to explicitly track actions and incentives until the next time the state returns to its current value, and then use the recursive structure of the game to relate the equilibrium payoff in the current state to the equilibrium continuation payoff when the current state next occurs, which will be a function of the public signals that are observed. In our proof of the folk theorem, we

construct a return-generating set of "Markov review strategies": The idea is that players condition their actions only on the state, and not on the signals of each other's actions, until the first time the state returns to its initial position, and that the incentive to conform to the specified Markov strategies is provided by the continuation payoffs from the time the state returns, which typically will depend on the history of both states and public signals during this "review phase."

Hörner, Sugaya, Takahashi, and Vieille (2009) independently developed a different proof of the folk theorem for irreducible stochastic games. They first provide a linear programming characterization of the limit equilibrium payoffs by using an extension of self-generation that tracks actions and incentives for the next $T$ periods, regardless of the realization of the state, instead of using return-generation. They then use somewhat weaker full-rank conditions to conclude that the solution of the linear programming problems implies the folk theorem. Because we do not provide a characterization of payoffs when the folk theorem fails, our proof more clearly highlights the link between the folk theorem in stochastic games and the folk theorem of FLM. We also fill in the details needed to adapt FLM's approach to account for the fact that "convex monotonicity" fails, as we explain in Remark 6.

# 2 Stochastic Games and Perfect Public Equilibria

Let $\boldsymbol{I} = \{1, \cdots, I\}$ be the set of players. At the beginning of the game, Nature chooses the state of the world $\omega^1$ from a finite set $\Omega$. The state may change as time passes; let $\omega^t \in \Omega$ denote the state in period $t$.

In each period $t$, players observe the state $\omega^t \in \Omega$, and then move simultaneously, with player $i \in \boldsymbol{I}$ choosing an action $a_i$ from a finite set $A_i$.[1] Given an action profile $a = (a_i)_{i \in \boldsymbol{I}} \in A \equiv \times_{i \in \boldsymbol{I}} A_i$, players observe a public signal $\mathbf{y}^t$ from a finite set $Y$ and the state in period $t+1$ is determined. Let $\pi^{\omega}(y, \omega' | a)$ denote the probability that players observe a signal $y$ and and the state for the next period is $\omega'$ when today's state is $\omega$ and players play action profile $a$. (Note that the distributions of $y$ and $\omega'$ may be correlated.) Player $i$'s realized payoff is $u_i^{\omega}(a_i, y)$, so her expected

[1] For notational convenience, we assume that $A_i$ does not depend on $\omega$. But with no difficulty our results extend to the case where $A_i$ depends on $\omega$.

payoff conditional on $\omega$ and $a$ is $g_i^\omega(a) = \sum_{\omega' \in \Omega} \sum_{y \in Y} \pi^\omega(y, \omega' | a) u_i^\omega(a_i, y)$; $g^\omega(a)$ denotes the vector of expected payoffs associated with action profile $a$.

In the infinitely repeated game, players have a common discount factor $\delta \in (0,1)$. Let $(\omega^\tau, a_i^\tau, y^\tau)$ be the state, player $i$'s pure action, and observed signal in period $\tau$, and denote player $i$'s private history at the end of period $t \geq 1$ by $h_i^t = (\omega^\tau, a_i^\tau, y^\tau)_{\tau=1}^t$. Let $h_i^0 = \emptyset$, and for each $t \geq 1$, let $H_i^t$ be the set of all $h_i^t$. Likewise, a public history up to period $t \geq 1$ is denoted by $h^t = (\omega^\tau, y^\tau)_{\tau=1}^t$, and $H^t$ denotes the set of all $h^t$. A strategy for player $i$ is a mapping $s_i : \bigcup_{t=0}^\infty H_i^t \times \Omega \to \triangle A_i$: Here $s_i(h^t, \omega)$ denotes player $i$'s action for period $t+1$ if the history through period $t$ is $h^t$ and the state for period $t+1$ is $\omega$. Let $S_i$ be the set of all strategies for player $i$, and let $S = \times_{i \in \boldsymbol{I}} S_i$. Let $v_i^\omega(\delta, s)$ denote player $i$'s average payoff in the stochastic game when the initial state is $\omega$, the discount factor is $\delta$, and players play strategy profile $s$. Let $v^\omega(\delta, s) = (v_i^\omega(\delta, s))_{i \in \boldsymbol{I}}$.

**Remark 1.** The state has to encode any and all history-dependent variation in payoffs and signal structure, which puts some constraints on how small the state space can be. However, because $y$ and $\omega$ are both public information, a stochastic game with state space $\Omega$ and signal space $Y$ is equivalent to a stochastic game with larger state space $\Omega^* = Y \times \Omega$ and a null signal space. In this latter case payoffs and state transitions are the same for states $\omega^* = (y, \omega) \in \Omega^*$ and $\tilde{\omega}^* = (\tilde{y}, \tilde{\omega}) \in \Omega^*$ if $\omega = \tilde{\omega}$. However, when we restrict attention to irreducible games, we will see that the game may be irreducible on $\Omega$ but not on $\Omega^*$. For example when actions are observable ($Y \simeq A$) transitions on $\Omega^*$ are not irreducible but transitions on $\Omega$ may be. Roughly speaking we would like the state space to be as small as possible, as in this representation our assumptions are most likely to be satisfied.

This paper studies a special class of Nash equilibria called *perfect public equilibria* or *PPE*. The notion of PPE was introduced by FLM for repeated games with public monitoring. This paper extends their definition to stochastic games.

**Definition 1.** A strategy $s_i \in S_i$ is *public* if it depends only on public information, i.e., $s_i(h_i^t, \omega^{t+1}) = s_i(\tilde{h}_i^t, \tilde{\omega}^{t+1})$ for all $t \geq 1$, $h_i^t = (\omega^\tau, a_i^\tau, y^\tau)_{\tau=1}^t$, $\tilde{h}_i^t = (\tilde{\omega}^\tau, \tilde{a}_i^\tau, \tilde{y}^\tau)_{\tau=1}^t$, $\omega^{t+1}$, and $\tilde{\omega}^{t+1}$ satisfying $\omega^\tau = \tilde{\omega}^\tau$ for all $\tau \leq t+1$ and $y^\tau = \tilde{y}^\tau$ for all $\tau \leq t$. A strategy profile $s \in S$ is *public* if $s_i$ is public for all $i \in \boldsymbol{I}$. A public strategy is *Markov* if it depends only on the current state, i.e., $s_i(h_i^t, \omega^{t+1}) = s_i(\tilde{h}_i^t, \tilde{\omega}^{t+1})$ for all $t \geq 1$, $h_i^t$, $\tilde{h}_i^t$, $\omega^{t+1}$, and $\tilde{\omega}^{t+1}$ satisfying $\omega^{t+1} = \tilde{\omega}^{t+1}$.

Given a public strategy profile $s \in S$, let $s|_{h^t}$ denote its continuation strategy profile after public history $h^t \in H^t$, and let $s|_{(h^t,\omega^{t+1})}$ denote the continuation strategy profile given $(h^t,\omega^{t+1}) \in H^t \times \Omega$.

**Definition 2.** A public strategy profile $s$ is a *perfect public equilibrium* or *PPE* if for every $(h^t,\omega) \in H^t \times \Omega$ the profile $s|_{(h^t,\omega)}$ is a Nash equilibrium of the infinite-horizon stochastic game with initial state $\omega$. A PPE is *Markov-perfect* if it uses Markov strategies.

It is a best response for each player to ignore the public signal and play a Markov strategy if all of the other players do. Since a Markov-perfect equilibrium exists (Sobel (1971)) there is always a PPE where all players ignore the signal.

Given a discount factor $\delta \in (0,1)$, let $E^\omega(\delta)$ denote the set of PPE payoffs with initial state $\omega$, i.e., $E^\omega(\delta)$ is the set of all vectors $v = (v_i)_{i\in \boldsymbol{I}} \in \boldsymbol{R}^I$ such that there is a PPE $s \in S$ satisfying

$$(1-\delta)E\left[\sum_{t=1}\delta^{t-1}g_i^{\omega^t}(a^t)\middle| s,\omega^1=\omega\right] = v_i$$

for all $i \in \boldsymbol{I}$.

In repeated games, the set of all PPE payoffs at time $t$ coincides with that at time $t+1$; this recursive structure is what permits the application of dynamic programming ideas as in the self-generation results of APS. In a stochastic game, the set of PPE payoffs conditional on the state has a recursive structure, which will allow us to apply a modified version of self-generation.[2] Self-generation is based on the decomposition of the set of equilibrium payoffs into the sum of current period payoffs and continuation equilibrium payoffs from the next period on. To develop an analog of this decomposition for stochastic games, we instead decompose the payoffs into the sum of payoffs until the state returns to its current value and continuation payoffs from the time of this return.

To simplify the notation, let $\Pr(h^t,\omega'|s,\omega)$ be the probability under profile $s$ that the $t$-period public history is $h^t$ and $\omega^{t+1}=\omega'$ given that the initial state is $\omega$. (Note that $\Pr(h^t,\omega'|s,\omega)=0$ for all $h^t=(\omega^\tau,y^\tau)_{\tau=1}^t$ such that $\omega^1 \neq \omega$.)

[2]Just as in repeated games with imperfectly observed actions, the set of all Nash or sequential equilibria need not have a recursive structure; see Fudenberg, Levine, and Maskin (1994, hereafter FLM) or Exercise 5.10 of Fudenberg and Tirole (1991).

Also, for each $t \geq 1$ and $\omega$, let $\tilde{H}^{\omega,t}$ be the set of all $t$-period public histories $h^t = (\omega^\tau, y^\tau)_{\tau=1}^t$ such that the state does not reach $\omega$ after period two. (That is, $\omega^\tau \neq \omega$ for $1 < \tau \leq t$.) Let $\tilde{H}^{\omega,t} = H^0$ for $t = 0$. Let $S^p$ denote the set of all public strategy profiles.

**Definition 3.** For $W \subseteq \boldsymbol{R}^I$, a pair $(s,v) \in S^p \times \boldsymbol{R}^I$ of a public strategy profile and a payoff vector is *enforceable with respect to $\delta$ and $W$ from state $\omega$* if there is a function $w = (w_i)_{i \in \boldsymbol{I}} : \bigcup_{t=1}^\infty H^t \to W$ such that

$$\begin{aligned} v_i =& (1-\delta)\left[ g_i^\omega(s(h^0,\omega)) + \sum_{t=1}^\infty \sum_{h^t \in \tilde{H}^{\omega,t}} \sum_{\omega' \neq \omega} \delta^t \Pr(h^t, \omega' | s, \omega) g_i^{\omega'}(s(h^t,\omega')) \right] \\ &+ \sum_{t=1}^\infty \sum_{h^t \in \tilde{H}^{\omega,t}} \delta^t \Pr(h^t, \omega | s, \omega) w_i(h^t) \end{aligned}$$

for all $i \in \boldsymbol{I}$, and

$$\begin{aligned} &(1-\delta)\left[ g_i^{\omega'}(s|_{h^t}(h^0,\omega')) + \sum_{\tau=1}^\infty \sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}} \sum_{\omega'' \neq \omega} \delta^\tau \Pr(\tilde{h}^\tau, \omega'' | s|_{h^t}, \omega') g_i^{\omega''}(s|_{h^t}(\tilde{h}^\tau,\omega'')) \right] \\ &+ \sum_{\tau=1}^\infty \sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}} \delta^\tau \Pr(\tilde{h}^\tau, \omega | s|_{h^t}, \omega') w_i((h^t, \tilde{h}^\tau)) \\ &\geq (1-\delta)\left[ g_i^{\omega'}(s'|_{h^t}(h^0,\omega')) + \sum_{\tau=1}^\infty \sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}} \sum_{\omega'' \neq \omega} \delta^\tau \Pr(\tilde{h}^\tau, \omega'' | s'|_{h^t}, \omega') g_i^{\omega''}(s'|_{h^t}(\tilde{h}^\tau,\omega'')) \right] \\ &\quad + \sum_{\tau=1}^\infty \sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}} \delta^\tau \Pr(\tilde{h}^\tau, \omega | s'|_{h^t}, \omega') w_i((h^t, \tilde{h}^\tau)) \end{aligned}$$

for all $i \in \boldsymbol{I}$, for all $s' \in S$ such that $s'_{-i} = s_{-i}$, for all $(t,\omega')$ such that (i) $t = 0$ and $\omega' = \omega$ or (ii) $t \geq 1$ and $\omega' \neq \omega$, and for all $h^t \in \tilde{H}^{\omega,t}$ such that $\omega^1 = \omega$. We say $(s,v) \in S^p \times \boldsymbol{R}^I$ is *enforced by $w$ for $\delta$ at $\omega$* if the above conditions are satisfied.

Here the function $w$ specifies the continuation payoffs from the time the state returns to $\omega$. The equality condition says that using the specified strategy will yield the target payoff, provided that the continuation payoffs from the return time onwards are as specified by $w$. The inequality condition is the incentive compatibility at states that are reached before the return to $\omega$.[3] Note that in games with a single state, this definition of enforceability reduces to that of APS.

[3]To see this, suppose that the history up to period $t$ is $h^t$ , $\omega^{t+1} = \omega'$, and consider the contin-

For each $(\delta, W, s)$, let $B^\omega(\delta, W, s)$ denote the set of all payoff vectors $v \in \boldsymbol{R}^I$ such that $(s, v)$ is enforceable with respect to $\delta$ and $W$ from $\omega$. Let $B^\omega(\delta, W)$ be the union of $B^\omega(\delta, W, s)$ over all $s$.

The first theorem notes that the equilibrium payoff set $E^\omega(\delta)$ is a fixed point of the operator $B^\omega$; it is the analog of the "factorization" theorem in APS.

**Theorem 1.** *For any $\omega \in \Omega$ and $\delta \in (0,1)$, $E^\omega(\delta) = B^\omega(\delta, E^\omega(\delta))$.*

*Proof.* We first prove $E^\omega(\delta) \subseteq B^\omega(\delta, E^\omega(\delta))$. Let $v \in E^\omega(\delta)$. Then there is a PPE $s$ with payoff $v$ for initial state is $\omega$. Since $s$ is a PPE, nobody wants to deviate until the state returns to $\omega$. Also, the continuation play after returning is a PPE. Therefore $v \in B^\omega(\delta, E^\omega(\delta))$, so that $E^\omega(\delta) \subseteq B^\omega(\delta, E^\omega(\delta))$.

To prove the converse, we construct a PPE with payoff $v \in B^\omega(\delta, E^\omega(\delta))$. Let $s$ and $w$ be such that $(s, v)$ is enforced by $w$ and $w(h^t) \in E(\delta)$. Consider a strategy profile such that players follow $s$ until the state returns to $\omega$, and once the state returns to $\omega$ after $t$-period play then players play a PPE with payoff $w(h^t)$ thereafter. It is easy to check that this strategy profile is a PPE with payoff $v$. *Q.E.D.*

The next theorem asserts that if a bounded set $W$ is "return-generating" then $W$ is in the equilibrium payoff set. Note that it reduces to APS's self-generation theorem if there is only a single state.

**Definition 4.** A subset $W$ of $\boldsymbol{R}^I$ is *return-generating with respect to $\delta$ and initial state $\omega$* if $W \subseteq B^\omega(\delta, W)$.

---

uation game from such that the game ends when the state reaches $\omega$. $\Pr(\tilde{h}^\tau, \omega'' | s|_{h^t}, \omega')$ denotes the probability that the history in the first $\tau$ periods of this continuation game is $\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}$ (so that the state does not return to $\omega$ during these periods) and the state in period $\tau+1$ is $\omega'' \neq \omega$. Also, the stage game payoff after such a history $(\tilde{h}^\tau, \omega'')$ (i.e., the stage game payoff in the $(\tau+1)$st period of the continuation game) is $g_i^{\omega''}$. Thus the first term of left-hand side denotes the expectation of the discounted sum of the stage game payoffs in this continuation game. On the other hand, $\Pr(\tilde{h}^\tau, \omega | s|_{h^t}, \omega')$ denotes the probability that the history in the first $\tau$ periods of the continuation game is $\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}$ and the state returns to $\omega$ in period $\tau+1$, and $w_i((h^t, \tilde{h}^\tau))$ denotes the continuation payoff after such a history. Thus the second term of the left-hand side denotes the expectation of the continuation payoffs $w$. Overall, the left-hand side is player $i$'s payoff in the continuation game corresponding to $(h^t, \omega')$. Likewise, the right-hand side is player $i$'s payoff of the continuation game when he deviates to $s_i'$.

**Theorem 2.** *If a subset $W$ of $\boldsymbol{R}^I$ is bounded and return-generating with respect to $\delta$ and $\omega$ then $W \subseteq E^\omega(\delta)$.*

*Proof.* As in APS, we first construct a candidate strategy profile, and then note that it has the specified payoffs and satisfies the no-one-stage-deviations test. In APS the definition of self-generation is used to find new actions in every period. Here the fact that $W$ is return-generated for $\omega$ implies that for $v \in W$ there is an associated repeated game strategy $s$; the construction of the overall equilibrium specifies that players conform to $s$ until the state returns to $\omega$. *Q.E.D.*

**Remark 2.** It follows from Theorems 1 and 2 that the equilibrium payoff set $E^\omega(\delta)$ for a given $\delta$ is the largest fixed point of the operator $B^\omega$. Note that these theorems do not assume the game is irreducible. Note also that they are vacuously true at states $\omega$ that are transient regardless of play.

**Remark 3.** Hörner, Sugaya, Takahashi, and Vieille (2009) use "$T$-period generation," which looks $T$ periods ahead instead of looking to the time of first return. In irreducible games, this property is sufficient for characterizing the equilibrium payoff set in the limit as $\delta$ goes to one. However, for a fixed discount factor, the set of equilibrium payoffs need not be a fixed point of the "$T$-period generation" operator.

**Remark 4.** Our proof of the folk theorem will use "Markov return generation," meaning that the constructed strategies starting at $\omega$ will be Markov until the state returns to $\omega$. This restricted form of return generation is more tractable but does not have the factorization property.

# 3 The Folk Theorem in Irreducible Stochastic Games

## 3.1 Irreducible Stochastic Games

In general, stochastic games can be very different than infinitely repeated games, as shown by the fact that any static game can be viewed as the first period of a stochastic game with two states, one of which is absorbing. More generally, the irreversibility created by absorbing states can support various sorts of backwards induction arguments with no real analog in the infinitely repeated setting. For

the rest of this paper we will restrict attention to irreducible stochastic games; this restriction rules out absorbing states and more strongly implies that no single player can prevent some state from being reached.

**Definition 5.** A stochastic game is *irreducible despite player i* if for each pair of states $(\omega, \omega') \in \Omega \times \Omega$, there is a $T > 0$ and a sequence $(\omega^1, \cdots, \omega^T)$ of states such that $\omega^1 = \omega$, $\omega^T = \omega'$, and for each $t < T$, there is a profile $\alpha$ such that $\sum_{y \in Y} \pi^{\omega^t}(y, \omega^{t+1} | a_i, \alpha_{-i}) > 0$ for all $a_i \in A_i$. The game is *irreducible* if it is irreducible despite each player $i$.

One trivial case where our irreducibility condition is satisfied is when the state represents a persistent demand shock whose evolution is independent of the players' actions, as in Rotemberg and Saloner (1986) and Ellison (1994). Besanko, Doraszelski, Kryukov, and Satterthwaite (2008) study an irreducible stochastic game where the actions do influence the state transitions: In their model the state is the knowledge or know-how of each firm, and a firm's state increases stochastically when it makes a sale, and decreases stochastically due to "organizational forgetting."

Let $V^\omega(\delta)$ be the set of feasible payoffs when the initial state is $\omega$ and the discount factor is $\delta$; i.e., $V^\omega(\delta) = \{v^\omega(\delta, s), \text{ for all } s \in S\}$. Since the game is irreducible, the set of feasible payoffs is independent of the initial state in the limit as $\delta$ goes to one. (See Dutta (1995).) Let $V$ denote this limit set, that is, $V = \lim_{\delta \to 1} V^\omega(\delta)$.

The minimax payoff to player $i$ in a stochastic game with initial state $\omega$ is defined to be

$$\underline{v}_i^\omega(\delta) = \inf_{s_{-i} \in S_{-i}} \sup_{s_i \in S_i} v_i^\omega(\delta, s).$$

The limit of the minimax payoff as $\delta \to 1$ is independent of the initial state from irreducibility, again as in Dutta (1995). (See also Bewley and Kohlberg (1976) and Mertens and Neyman (1981).) Let $V^*$ denote the limit of the set of feasible and individually rational payoffs.

The following lemma records an immediate but important consequence of irreducibility: For each player $i$, there is a Markov strategy $s_{-i}$ of the opponents so that for any pair of states $(\omega, \omega')$ there is strictly positive probability that the state reaches $\omega$ from $\omega'$ within $|\Omega|$ periods. When players are sufficiently patient,

this will allow us to provide adequate current incentives with strategies that are Markov until the state returns to its current position.

Given a Markov strategy profile $s$, let $s(\omega)$ denote the action profile played by $s$ in state $\omega$.

**Lemma 1.** *If the game is irreducible, then there is a Markov strategy profile $s_{-i}$ such that for each pair of states $(\omega,\omega')$, there is an integer $T>0$ and a sequence $(\omega^1,\cdots,\omega^T)$ of states such that $\omega^1=\omega$, $\omega^T=\omega'$, and $\sum_{y\in Y}\pi^{\omega^t}(y,\omega^{t+1}|a_i,s_{-i}(\omega^t))>0$ for all $a_i\in A_i$ and $t<T$. Moreover for such $s_{-i}$ there is $p_i\in(0,1)$ such that for each pair of states $(\omega,\omega')$, the state reaches $\omega$ from $\omega'$ within $|\Omega|$ periods with at least probability $p_i$ if players play $s'$ such that $s'_{-i}=s_{-i}$.*

*Proof.* One such $s_{-i}$ is for players $-i$ to randomize with each player using a uniform distribution over his actions in each state. Consider any such $s_{-i}$ and fix a pair of states $(\omega,\omega')$, and let $T>0$ and $(\omega^1,\cdots,\omega^T)$ be such that $\omega^1=\omega$, $\omega^T=\omega'$, and $\sum_{y\in Y}\pi^{\omega^t}(y,\omega^{t+1}|a_i,s_{-i}(\omega^t))>0$ for all $a_i\in A_i$ and $t<T$. Without loss of generality we assume $T\leq|\Omega|$. The state reaches $\omega$ from $\omega'$ within $T$ periods with at least probability $p_i(\omega,\omega')$, where

$$p_i(\omega,\omega')=\prod_{t=1}^{T-1}\min_{a_i^t\in A_i}\sum_{y^t\in Y}\pi^{\omega^t}(y^t,\omega^{t+1}|a_i^t,s_{-i}(\omega^t)).$$

Letting $p_i$ be the minimum of $p_i(\omega,\omega')$ over all $(\omega,\omega')$, the lemma follows. *Q.E.D.*

## 3.2 Full Rank Conditions

As in repeated games with imperfectly observed actions, the key to whether the folk theorem holds is whether incentives can be provided along "hyperplanes" that correspond to trading utility between the players at various rates, and this is possible when the information revealed by the public outcomes permits deviations by one player to be statistically identified and distinguished from deviations by others. The following "full rank" conditions are sufficient for this.

For each $i$, $(\omega,\omega')$, and $\alpha$, let $\Pi_i^{(\omega,\omega')}(\alpha)$ be a matrix with rows $(\pi^{\omega}(\omega',y|a_i,\alpha_{-i}))_{y\in Y}$ for all $a_i\in A_i$. Let $\Pi_{(i,j)}^{(\omega,\omega')}(\alpha)$ be the matrix constructed by stacking matrices $\Pi_i^{(\omega,\omega')}(\alpha)$ and $\Pi_j^{(\omega,\omega')}(\alpha)$.

**Definition 6.** Profile $\alpha$ has *individual full rank for* $(i,\omega,\omega')$ if $\Pi_i^{(\omega,\omega')}(\alpha)$ has rank equal to $|A_i|$.

Note that this condition implies $\pi^{\omega}(\omega'|a_i,\alpha_{-i})>0$ for all $a_i$. It also implies that each action by player $i$ leads to a different distribution on $y$, conditional on a transition to $\omega'$ (though this need not be true if some other state occurs.) We will use this in Lemmas 5 and 6 to construct return payoffs that make the player indifferent between all of his actions.

**Definition 7.** For each $i,j$, $i\neq j$ and $(\omega,\omega')$, profile $\alpha$ has *pairwise full rank for* $(i,j)$ *and* $(\omega,\omega')$ if $\Pi_{(i,j)}^{(\omega,\omega')}(\alpha)$ has rank equal to $|A_i|+|A_j|-1$.

Pairwise full rank for $(\omega,\omega')$implies that conditional on a transition from $\omega$ to $\omega'$ deviations by player $i$ can be statistically distinguished from deviations by $j$. It is satisfied for generic distributions on signals provided that $|Y|$ is at least $|A_i|+|A_j|-1$, as in FLM, and that the transition from $\omega$ to $\omega'$ has positive probability regardless of the play of $i$ or $j$.

**Condition IFR.** For each $(\omega,\omega')\in\Omega\times\Omega$, there is an integer $T>0$ and a sequence $(\omega^1,\cdots,\omega^T)$ of states such that $\omega^1=\omega$, $\omega^T=\omega'$, and for each $t<T$, every pure action profile has individual full rank for $(\omega^t,\omega^{t+1})$ and every player $i$.

**Condition PFR.** For each $(\omega,\omega')\in\Omega\times\Omega$, there is an integer $T>0$ and a sequence $(\omega^1,\cdots,\omega^T)$ of states such that $\omega^1=\omega$, $\omega^T=\omega'$, and for each $(i,j)$ and $t<T$, there is a profile $\alpha^t$ that has pairwise full rank for $(\omega^t,\omega^{t+1})$ and for $(i,j)$.

Note that (IFR) and (PFR) are weaker than assuming that full rank conditions are satisfied for all pairs of states $(\omega,\omega')$, as these conditions require full rank only for certain $(\omega,\omega')$. Note also that (IFR) applies to all pure action profiles, while (PFR) only asks that for each pair of players there is at least one profile with the stronger pairwise full rank property for that pair. These conditions are generically satisfied if the game is irreducible and $|Y|\geq|A_i|+|A_j|-1$ for all $(i,j)$ with $i\neq j$.[4] In particular they are satisfied in irreducible games with observed actions, the case studied by Dutta (1995), provided that the observations of the actions are

[4]These conditions could be relaxed along the lines of Kandori and Matsushima (1998).

represented by the signal $y$ as opposed to being embedded in the state. If the actions are represented as part of the state then both full rank and irreducibility fail.

## 3.3 The Folk Theorem

**Definition 8.** A subset $W$ of $\boldsymbol{R}^I$ is *smooth* if it is closed and convex; it has a non-empty interior; and there is a unique unit normal for each point on the boundary of $W$.[5]

The following theorem is the main result of this paper; it shows that (IFR) and (PFR) are sufficient for the folk theorem.

**Theorem 3.** *Suppose (IFR) and (PFR) hold. Then, for any smooth subset $W$ of the interior of $V^*$, there is $\overline{\delta} \in (0,1)$ such that $W \subseteq E^\omega(\delta)$ for all $\omega$ and $\delta \in (\overline{\delta},1)$. Therefore if $V^*$ has dimension $I$, then $\lim_{\delta\to 1} E^\omega(\delta) = V^*$.*

**Remark 5.** Because the state transitions are irreducible, it might seem natural to consider strategies that track separate histories for each state, so that play in state $\omega_1$ depends only on observations from previous periods where the state was $\omega_1$. This approach works if the state transitions are independent of the actions, and yields a folk theorem whenever the folk theorem holds in each state considered in isolation. However, stratification by the state does not work when actions have an effect on the state transitions, as even in a control problem optimization requires that the player take into account the way his current action influences tomorrow's state and thus tomorrow's payoffs.[6] For this reason our proof does not stratify the histories by state but instead works with the game as a whole, keeping track of the effect of actions on state transitions.

[5] A sufficient condition for each point on the boundary of $W$ to have a unique unit normal is that the boundary is a $C^2$-submanifold of $\boldsymbol{R}^I$.

[6] Suppose for example there are two actions $a'$ and $a''$ and three states, $\omega_1$, $\omega_2$, and $\omega_3$; $(\omega_1,a')$ is followed by $\omega_2$, and $(\omega_1,a'')$ is followed by $\omega_3$; both $\omega_2$ and $\omega_3$ are followed by $\omega_1$ regardless of the action played. In state $\omega_1$, $a'$ gives payoff 1 and $a''$ gives payoff 0; in state $\omega_2$ payoffs are 0 regardless of the action played, while in state $\omega_3$ payoffs are identically equal to 3. State-by-state optimization leads the player to choose $a'$ in state $\omega_1$ but this is not optimal if $\delta > \frac{1}{3}$. Then even when $\delta$ is close to 1, equilibrium play for the "state $\omega_1$ game" need not be consistent with overall equilibrium, as players will give roughly equal weight to their current period payoff and the payoff in the next period.

**Remark 6.** Much of the proof of this theorem is similar to FLM: We first develop a local condition -one that applies pointwise to each $v \in W$- for a set to be return-generating for large discount factors. We then show that the local condition is satisfied under the full-rank assumptions, basically because they allow $v$ on the boundary of $W$ to be generated by continuation payoffs that lie on the hyperplanes tangent to $W$ at $v$ and whose variation is of order $(1-\delta)$. One key difference with FLM is that we use return-generation instead of self-generation to show that various payoff vectors can be supported by equilibrium strategies. The particular strategies that we show are return-generated have the form of stochastic review strategies, in the sense that the strategies are Markov (depend only on the state and not on the sequence of signals) until the state returns to $\omega$.[7] The second key difference is that the local condition, which we call "uniform decomposability," is more complicated then the "local self-decomposability" used in FLM's Lemma 4.2. The reason for this is as follows. In a repeated game, if payoff vector $v$ is generated using a function $w$ for some $\delta$, then $v$ is generated by $w'$ for $\delta' \in (\delta, 1)$ where the function $w'$ is a convex combination of $w$ and the constant function $v$. In particular, if $v$ and all continuation payoffs specified by $w$ are chosen from a convex set $W$, then the continuation payoffs specified by $w'$ are in $W$ as well. Using this monotonicity result, FLM show that for $W$ to be self-generating for sufficiently large $\delta$, it is sufficient that $W$ is "locally decomposable." In stochastic games this monotonicity result fails because for any fixed $\delta$, the payoff to a given strategy depends on the initial state.[8]

[7] More specifically, in order to generate a target payoff, we let players follow a Markov strategy profile until the state returns to $\omega$, and we choose the continuation payoffs when the state returns to $\omega$ in such a way that no player wants to deviate. This construction requires that a player's deviation be statistically distinguished from a distribution of a public signal $y$ given that the state tomorrow is $\omega$. This is the reason why we need full-rank conditions for a distribution of $y$ given a state transition.

[8] For example, suppose there is one player and two states, $\omega_1$ and $\omega_2$. The state transitions follow a cycle regardless of the actions played: $\omega_1$ is followed by $\omega_2$ and $\omega_2$ is followed by $\omega_1$. The player has a single action (so the incentive constraint is vacuous) and his payoff is 1 in state $\omega_1$ and 0 in state $\omega_2$. For $\delta' = \frac{1}{2}$, the average payoff during the first two periods is $\frac{1-\delta}{1-\delta^2}(1+0) = \frac{2}{3}$, so that the payoff $v = \frac{2}{3}$ is enforced by the constant continuation payoff $w' = \frac{2}{3}$. On the other hand, for $\delta'' > \frac{1}{2}$, the average payoff during the first two periods is less than $\frac{2}{3}$, and hence for $v$ to be enforced, the constant continuation payoff $w''$ must be greater than $\frac{2}{3}$, which is not a convex combination of $v$ and $w'$. Thus local decomposability is not sufficient for self-generation, and

We first introduce the concept of *uniform decomposability*. We say that $\lambda \in \boldsymbol{R}^I$ is *regular* if it has at least two non-zero components, and is *singular* if it has exactly one non-zero component. Let $\Lambda$ be the set of all $\lambda$ with $|\lambda| = 1$. Given any $v \in \boldsymbol{R}^I$, $\lambda \in \Lambda$, $\varepsilon > 0$, $K > 0$, and $\delta \in (0,1)$, let $G_{v,\lambda,\varepsilon,K,\delta}$ be the set of all $v'$ such that $\lambda \cdot v \geq \lambda \cdot v' + (1-\delta)\varepsilon$ and such that $v'$ is within $(1-\delta)K$ of $v$. (See Figure 1, where this set is labeled "$G$.")

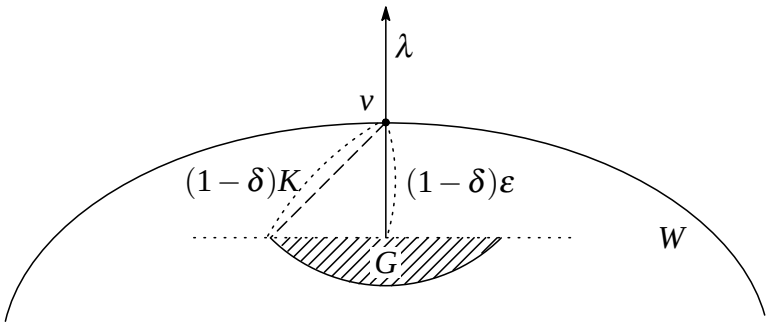


Figure 1: Continuation payoff $w(h^t)$ is chosen from $G$.

**Definition 9.** A subset set $W$ of $\boldsymbol{R}$ is *uniformly decomposable for initial state* $\omega$ if there are $\varepsilon > 0$, $K > 0$, and $\overline{\delta} \in (0,1)$ such that for all $v \in W$, $\delta \in (\overline{\delta}, 1)$, and $\lambda \in \Lambda$, there are $s \in S$ and $w : \bigcup_{t=1}^{\infty} H^t \to \boldsymbol{R}^I$ such that $(s,v)$ is enforced by $w$ for $\delta$ at $\omega$, and such that $w(h^t) \in G_{v,\lambda,\varepsilon,K,\delta}$ for all $t$ and $h^t$.

In words, uniform decomposability requires that any $v \in W$ is enforced by a function $w$ that chooses continuation payoffs from the set $G_{v,\lambda,\varepsilon,K,\delta}$. Note that the set $G_{v,\lambda,\varepsilon,K,\delta}$ is bounded uniformly in $\lambda$; indeed, for each $\lambda$, the set $G_{v,\lambda,\varepsilon,K,\delta}$ is in the ball with center $v$ and radius $(1-\delta)K$.

The following lemma shows that if a smooth set $W$ is uniformly decomposable, then it is return-generating for sufficiently large $\delta$.

**Lemma 2.** *Suppose that a smooth and bounded subset $W$ of $\boldsymbol{R}^I$ is uniformly decomposable for initial state $\omega$. Then there is $\overline{\delta} \in (0,1)$ such that $W$ is return-generating for $\delta \in (\overline{\delta}, 1)$ and for $\omega$, and hence $W \subseteq E^{\omega}(\delta)$ for $\delta \in (\overline{\delta}, 1)$ and for $\omega$.*

FLM's proof does not directly apply.

*Proof.* First of all, we prove the following claim: For any smooth set $W$, and for any $\varepsilon > 0$, $K$, and $v \in W$, there are $\delta_v \in (0,1)$, $\lambda_v \in \Lambda$, and an open set $U_v$ containing $v$ such that for any $v' \in U_v \cap W$ and $\delta \in (\delta_v, 1)$, the set $G_{v',\lambda_v,\varepsilon,K,\delta}$ is in $W$.

If $v$ is in the interior of $W$, then the statement is obvious. So consider $v$ on the boundary of $W$. Let $\lambda_v$ be a unit normal at $v$. Since $W$ is smooth, its boundary is locally flat, so that (as in Fudenberg and Levine (1994)) there is $\delta_v \in (0,1)$ such that $G_{v,\lambda_v,\varepsilon,K,\delta_v}$ is in the interior of $W$. Then there is an open set $U_v$ containing $v$ such that for any $v' \in U_v \cap W$, the set $G_{v',\lambda_v,\varepsilon,K,\delta_v}$ is in the interior of $W$. Note that for any $\delta \in (\delta_v, 1)$ and $v' \in U_v \cap W$, any point $v'' \in G_{v',\lambda_v,\varepsilon,K,\delta}$ is a strict convex combination of $v'$ and some point $v''' \in G_{v',\lambda_v,\varepsilon,K,\delta_v}$. This shows that such a $v''$ is in the interior of $W$, since $W$ is convex, $v'$ is an element of $W$, and $v'''$ is an interior point of $W$. (See Figure 2.) Therefore, for any $\delta \in (\delta_v, 1)$ and $v' \in U_v \cap W$, the set $G_{v',\lambda_v,\varepsilon,K,\delta}$ is in the interior of $W$. This completes the proof of the above claim.

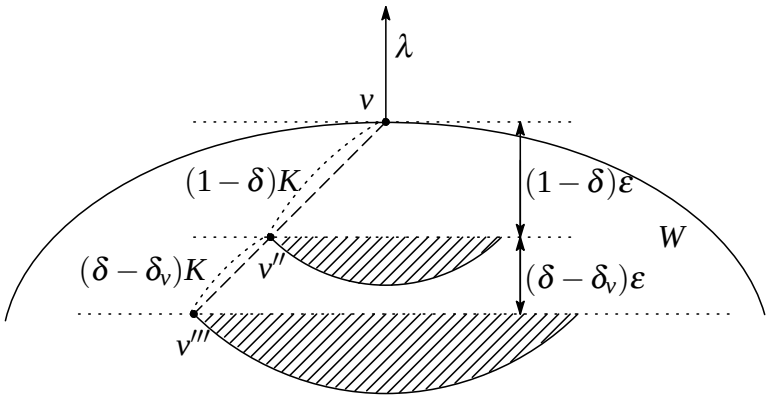


Figure 2: $v''$ is a convex combination of $v$ and $v'''$.

Let $U_v$ and $\delta_v$ be as in the claim for each $v \in W$. Note that $\{U_v\}_{v \in W}$ is an open cover of $W$. Since $W$ is compact, $\{U_v\}_{v \in W}$ has a finite subcover. Let $\overline{\delta}$ be the maximum of $\delta_v$'s on this subcover. Then the above claim implies that for each point $v \in W$, there is $\lambda \in \Lambda$ such that for any $\delta \in (\overline{\delta}, 1)$, the set $G_{v,\lambda,\varepsilon,K,\delta}$ is in $W$. Since $W$ is uniformly decomposable, it follows that $W$ is return-generating for $\delta \in (\overline{\delta}, 1)$. *Q.E.D.*

In what follows, we show that any smooth subset $W$ of the interior of $V^*$ is uniformly decomposable. Let $S_{-i}^{\mathrm{IFR}}(i)$ be the set of Markov strategies $s_{-i}$ such that

for any player $i$'s pure Markov strategy $s_i$ and for any $(\omega, \omega') \in \Omega \times \Omega$, there is an integer $T > 0$ and a sequence $(\omega^1, \cdots, \omega^T)$ of states such that $\omega^1 = \omega$, $\omega^T = \omega'$, and for each $t < T$, profile $s(\omega^t)$ has individual full rank for $(\omega^t, \omega^{t+1})$ and for all $j \neq i$. Let $S^{\text{IFR}}(i, \delta)$ be the set of Markov strategy profile $s$ such that player $i$'s play is pure and optimal in every history for $\delta$ and $s_{-i} \in S^{\text{IFR}}_{-i}(i)$.[9] Note that $S^{\text{IFR}}(i, \delta)$ is non-empty if (IFR) holds.

Likewise, let $S^{\text{PFR}}$ be the set of Markov strategy profiles $s$ such that for any $(\omega, \omega') \in \Omega \times \Omega$, there is an integer $T > 0$ and a sequence $(\omega^1, \cdots, \omega^T)$ of states such that $\omega^1 = \omega$, $\omega^T = \omega'$, and for each $t < T$, profile $s(\omega^t)$ has pairwise full rank for $(\omega^t, \omega^{t+1})$ and for all pairs of players.

The next two lemmas are straightforward extensions of FLM: Lemma 3 shows that $S^{\text{PFR}}$ is non-empty if (PFR) holds. Lemma 4 asserts that player $i$'s minimax payoff can be approximated by $s \in S^{\text{IFR}}(\delta, i)$, if (IFR) holds. The proofs are presented in Appendix.

**Lemma 3.** *Suppose that (PFR) holds. Then the set $S^{PFR}$ is dense in the set of all Markov strategy profiles.*

**Lemma 4.** *Suppose that (IFR) holds. Then for any $\omega \in \Omega$, $\delta \in (0,1)$, and $\varepsilon > 0$, there is $s \in S^{IFR}(\delta, i)$ such that $|v_i^{\omega}(\delta, s) - \underline{v}_i^{\omega}(\delta)| < \varepsilon$.*

The next lemma is an extension of Theorem 5.1 of FLM. Roughly speaking, it shows that any $s \in S^{\text{PFR}}$ is "enforceable with respect to all regular hyperplanes," and that the continuation payoffs used to enforce it can be chosen to converge to a constant at rate $(1 - \delta)$. Note that the latter conclusion uses the irreducibility assumption. As Lemma 1 shows, irreducibility assures that there is strictly positive probability that the state returns to the initial state within $|\Omega|$ periods. We use this and individual full rank to show that a player's short-run incentive to deviate can be offset by an appropriate specification of the continuation payoffs on return to the initial state; the ability to do this with payoffs on the specified hyperplane comes from the pairwise full rank assumption.

In the proof of the lemma, we explicitly construct the continuation payoffs $w$ as follows: We first show that for a given state $\omega'$, player $i$ will be indifferent

[9] Note the dependence on $\delta$ here, which does not occur in the analogous definition in FLM. This is because our analog of a static best response is a response that is optimal until the state returns to $\omega$, and this sort of intermediate-horizon optimality depends on $\delta$.

over all actions at $\omega'$ if he receives a bonus continuation payoff $z_i^{\omega'}$ when the state evolves following a specific chain from $\omega'$ to $\omega$, with no bonus paid if the state follows a different path. Then we define $w_i$ as the collection of the bonuses $z_i$ that should be paid at the corresponding histories. See Appendix for a detailed proof.

**Lemma 5.** *For each initial state $\omega$, Markov strategy profile $s \in S^{PFR}$, and $\overline{\delta} \in (0,1)$, there is $\kappa > 0$ such that for each regular direction $\lambda \in \Lambda$, there is $K > 0$ such that for each $\delta \in (\overline{\delta},1)$ and for each $v \in \boldsymbol{R}^I$ such that $\lambda \cdot v^\omega(\delta,s) \geq \lambda \cdot v$, there is $w$ such that*

*(i)* *$(s,v)$ is enforced by $w$ for $\delta$ at $\omega$,*

*(ii)* *$\lambda \cdot w(h^t)$ is independent of $t$ and $h^t$ and $\lambda \cdot w(h^t) \leq \lambda \cdot v - (1-\delta)(\lambda \cdot v^\omega(\delta,s) - \lambda \cdot v)$, and*

*(iii)* *$|v - w(h^t)| < (1-\delta)(\kappa |v^\omega(\delta,s) - v| + K)$ for each $t$ and $h^t$.*

The next lemma is an extension of Lemma 5.2 of FLM; it shows that any $s \in S^{\text{IFR}}$ is "enforceable with respect to all coordinate hyperplanes orthogonal to the $i$th coordinate axis." Here we choose player $i$'s strategy $s_i$ depending on $\delta$, since player $i$'s best reply against $s_{-i}$ might vary with the discount factor. Again, the proof can be found in Appendix.

**Lemma 6.** *For each initial state $\omega$, Markov strategy $s_{-i} \in S_{-i}^{IFR}(i)$, and $\overline{\delta} \in (0,1)$, there is $\kappa > 0$ such that for each direction $\lambda$ such that $\lambda_i \neq 0$ and $\lambda_j = 0$ for all $j \neq i$, there is $K > 0$ such that for each $\delta \in (\overline{\delta},1)$ and for each $v \in \boldsymbol{R}^I$ such that $\lambda_i v_i \leq \lambda_i \max_{s_i'} v_i^\omega(\delta, s_i', s_{-i})$, there is player $i$'s pure Markov strategy $s_i$ and $w$ such that*

*(i)* *$(s,v)$ is enforced by $w$ for $\delta$ at $\omega$,*

*(ii)* *$\lambda_i w_i(h^t)$ is independent of $t$ and $h^t$ and $\lambda_i w_i(h^t) \leq \lambda_i v_i - (1-\delta)\lambda_i(\max_{s_i'} v_i^\omega(\delta, s_i', s_{-i}) - v_i)$, and*

*(iii)* *$|v - w(h^t)| < (1-\delta)(\kappa |v^\omega(\delta,s) - v| + K)$ for each $t$ and $h^t$.*

The previous lemmas each apply for a given $\delta$; the next lemma states conclusions that hold for all $\delta$ sufficiently close to 1. The proof is given in Appendix.

**Lemma 7.** *Suppose that (IFR) and (PFR) hold. Then for any initial state $\omega \in \Omega$ and for any smooth subset $W$ of the interior of $V^*$, there are $\varepsilon > 0$ and $\overline{\delta} \in (0,1)$ such that the following properties hold:*

*(a) There is a finite set of Markov strategy profiles $\{s^1, \cdots, s^N\}$ such that $s^n \in S^{PFR}$ for all $n$, and for every $\delta \in (\overline{\delta}, 1)$ and for every regular $\lambda \in \Lambda$, $\max_{n \in \{1,\cdots,N\}} \lambda \cdot v^n > \max_{v \in W} \lambda \cdot v + \varepsilon$, where $v^n = v^\omega(\delta, s^n)$.*

*(b) For each $i \in \boldsymbol{I}$ and for each $\lambda$ such that $|\lambda_i| = 1$ and $\lambda_j = 0$ for all $j \neq i$, there is $s_{-i} \in S_{-i}^{IFR}(i)$ such that for every $\delta \in (\overline{\delta}, 1)$, there is player $i$'s Markov strategy $s_i$ such that $s = (s_i, s_{-i}) \in S^{IFR}(\delta, i)$ and $\lambda \cdot v > \max_{v' \in W} \lambda \cdot v' + \varepsilon$, where $v = v^\omega(\delta, s)$.*

Now we are in a position to prove uniform decomposability of a smooth subset $W$ of the interior of $V^*$. A main point in the proof is that the variation in continuation payoffs needed to enforce a given limit payoff $v$ on the hyperplane corresponding to $\lambda$ is bounded uniformly in $\lambda$.

**Lemma 8.** *Suppose that (IFR) and (PFR) hold. Then any smooth subset $W$ of the interior of $V^*$ is uniformly decomposable for any initial state $\omega \in \Omega$.*

*Proof.* Fix $\omega$ and $W$. Fix $\overline{\delta} \in (0,1)$ and $\tilde{\varepsilon}$ so that Lemma 7 holds. Applying Lemmas 5 and 6 to the strategy profiles specified in Lemma 7, it follows that there is $\kappa > 0$ such that for each $\lambda \in \Lambda$, there is $\tilde{K}_\lambda > 0$ such that for each $\delta \in (\overline{\delta}, 1)$ and $v \in W$, there is a Markov strategy profile $s_{v,\lambda,\delta}$ and a function $\tilde{w}_{v,\lambda,\delta}$ such that

(i) $(s_{v,\lambda,\delta}, v)$ is enforced by $\tilde{w}_{v,\lambda,\delta}$ for $\delta$ at $\omega$,

(ii) $\lambda \cdot \tilde{w}_{v,\lambda,\delta}(h^t) \leq \lambda \cdot v - (1-\delta)\tilde{\varepsilon}$ for each $t$ and $h^t$, and

(iii) $|v - w(h^t)| < (1-\delta)(\kappa |v^\omega(\delta, s_{v,\lambda,\delta}) - v| + \tilde{K}_\lambda)$ for each $t$ and $h^t$.

Set $\varepsilon = \frac{\tilde{\varepsilon}}{2}$, and for each $\lambda \in \Lambda$, let $K_\lambda > \kappa_2 |v^\omega(\delta, s_{v,\lambda,\delta}) - v| + \tilde{K}_\lambda$ for all $v \in W$ and $\delta \in (\overline{\delta}, 1)$. Then it follows from (ii) and (iii) that $w_{v,\lambda,\delta}(h^t) \in G_{v,\lambda,2\varepsilon,K_\lambda,\delta}$ for all $t$ and $h^t$.

Note that for each $v \in W$ and $\lambda \in \Lambda$, there is an open set $U_{v,\lambda,\delta} \subseteq \boldsymbol{R}^I$ containing $\lambda$ such that $\lambda' \cdot v \geq \lambda' \cdot w_{v,\lambda,\delta}(h^t) + (1-\delta)\varepsilon$ for all $t$, $h^t$, and $\lambda' \in \Lambda \cap U_{v,\lambda,\delta}$. In

particular, since $W$ is bounded, there is $U_{\lambda,\delta} \subseteq \boldsymbol{R}^I$ containing $\lambda$ such that $\lambda' \cdot v \geq \lambda' \cdot w_{v,\lambda,\delta}(h^t) + (1-\delta)\varepsilon$ for all $t$, $h^t$, $v \in W$ and $\lambda' \in \Lambda \cap U_{\lambda,\delta}$.

The set $\Lambda$ is compact, so $\{U_{\lambda,\delta}\}_{\lambda\in\Lambda}$ has a finite subcover $\{U_{\lambda,\delta}\}_{\lambda\in\Lambda^*}$. For each $v$ and $\lambda$, let $s^*_{v,\lambda,\delta} = s_{v,\lambda',\delta}$ and $w^*_{v,\lambda,\delta}(h^t) = w_{v,\lambda',\delta}(h^t)$, where $\lambda' \in \Lambda^*$ is such that $\lambda \in U_{\lambda',\delta}$. Then letting $K = \max_{\lambda\in\Lambda^*} K_\lambda$, the specified $(s^*_{v,\lambda,\delta}, w^*_{v,\lambda,\delta})$ satisfies all the desired conditions for a given $\lambda \in \Lambda$; that is, $(s^*_{v,\lambda,\delta}, v)$ is enforced by $w^*_{v,\lambda,\delta}$ and $w^*_{v,\lambda,\delta}$ chooses the continuation payoffs from the set $G_{v,\lambda,\varepsilon,K,\delta}$. Note that now $K$ is independent of $\lambda$, and hence the proof is completed. *Q.E.D.*

# 4 Conclusion

Our results leave open many interesting avenues of research. One of them is whether the folk theorem for stochastic games holds when strategies are restricted to have finite memory, as it does in games with a single state (Hörner and Olszewski (2009)). A second is to suppose that the discount factor tends to one because time periods grow short, so that the state transition probabilities will vary with the discount factor; we explore this in ongoing research with Johannes Horner. Yet another extension is to allow for private information about the time-varying state, as in Athey and Bagwell (2008) and Escobar and Toikka (2009).

# Appendix

## A.1 Proof of Lemma 3

**Lemma 3.** *Suppose that (PFR) holds. Then the set $S^{PFR}$ is dense in the set of all Markov strategy profiles.*

*Proof.* For each $\omega \in \Omega$, let $\Omega(\omega)$ denote the set of all $\omega' \in \Omega$ such that for each pair $(i,j)$ of players, there is a profile $\alpha$ that has pairwise full rank for $(i,j)$ and $(\omega,\omega')$. Let $A^{\mathrm{PFR}}(\omega,\omega')$ be the set of all action profiles $\alpha$ that has pairwise full rank for $(\omega,\omega')$ and for all pairs of players. As Lemma 6.2 of FLM shows, the set $A^{\mathrm{PFR}}(\omega,\omega')$ is open and dense in the set of all action profiles for each $\omega \in \Omega$ and $\omega' \in \Omega(\omega)$. Let $A^{\mathrm{PFR}}(\omega) = \bigcap_{\omega'\in\Omega(\omega)} A^{\mathrm{PFR}}(\omega,\omega')$. Then $A^{\mathrm{PFR}}(\omega)$ is open and dense in the set of all action profiles. Note that under (PFR), for

any pair $(\omega,\omega')$, there is an integer $T$ and a sequence $(\omega^1,\cdots,\omega^T)$ such that $\omega^1=\omega$, $\omega^T=\omega'$, and $\omega^{t+1}\in\Omega(\omega^t)$ for all $t<T$. Thus we have $s\in S^{\mathrm{PFR}}$ for any Markov strategy profile $s$ such that $s(\omega)\in A^{\mathrm{PFR}}(\omega)$ for all $\omega$. This shows that $S^{\mathrm{PFR}}$ contains $\times_{\omega\in\Omega}A^{\mathrm{PFR}}(\omega)$. Therefore $S^{\mathrm{PFR}}$ is dense in the set of all Markov strategy profiles. *Q.E.D.*

## A.2 Proof of Lemma 4

**Lemma 4.** *Suppose that (IFR) holds. Then for any $\omega\in\Omega$, $\delta\in(0,1)$, and $\varepsilon>0$, there is $s\in S^{IFR}(\delta,i)$ such that $|v_i^\omega(\delta,s)-\underline{v}_i^\omega(\delta)|<\varepsilon$.*

*Proof.* For each $\omega\in\Omega$, let $\Omega(\omega)$ denote the set of all $\omega'\in\Omega$ such that any pure action profile has individual full rank for $(\omega,\omega')$ and for all players. For each $a_i\in A_i$, let $C(\omega,\omega',a_i)$ denote the set of all $\alpha_{-i}$ such that the profile $(a_i,\alpha_{-i})$ has individual full rank for all $j\neq i$ and for $(\omega,\omega')$. As Lemma 6.3 of FLM shows, the set $C(\omega,\omega',a_i)$ is open and dense in the set of all $\alpha_{-i}$, for each $\omega\in\Omega$, $\omega'\in\Omega(\omega)$, and $a_i\in A_i$. Let $C(\omega)=\bigcap_{\omega'\in\Omega(\omega)}\bigcap_{a_i}C(\omega,\omega',a_i)$. Let $C=\times_{\omega\in\Omega}C(\omega)$. Then $C$ is dense in the set of all Markov strategy profiles. Since (IFR) holds, for any Markov strategy profile $s$ in $C$ and for any $(\omega,\omega')$, there is a sequence $(\omega^1,\cdots,\omega^T)$ of states such that for each $t$, $s(\omega^t)$ has individual full rank for $(\omega^t,\omega^{t+1})$ and for all $j\neq i$.

Let $\underline{s}$ be a minimax profile against player $i$ for $\delta$ and for initial state $\omega$. Without loss of generality, $\underline{s}$ is Markov, so write it as $(\underline{s}(\omega))_{\omega\in\Omega}$. Since $C$ is dense, there is a sequence $\{(s^k_{-i}(\omega))_{\omega\in\Omega}\}_{k=1}^\infty$ converging to $(\underline{s}_{-i}(\omega))_{\omega\in\Omega}$ with $(s^k_{-i}(\omega))_{\omega\in\Omega}\in C$ for all $k$. Let $s_i^k$ be a best response to $(s^k_{-i}(\omega))_{\omega\in\Omega}$. Without loss of generality $s_i^k$ is pure and Markov, so write it by $(s_i^k(\omega))_{\omega\in\Omega}$. By definition of $C$, for any pair $(\omega,\omega')$, there is an integer $T$ and a sequence $(\omega^1,\cdots,\omega^T)$ such that $\omega^1=\omega$, $\omega^T=\omega'$, and for each $t<T$, the profile $(s_i^k(\omega^t),s^k_{-i}(\omega^t))$ has individual full rank for all $j\neq i$ and for $(\omega^t,\omega^{t+1})$. Choose $(s_i(\omega))_{\omega\in\Omega}$ and a subsequence of $\{(s_i^k(\omega),s^k_{-i}(\omega))_{\omega\in\Omega}\}_{k=1}^\infty$ such that $(s_i^k(\omega))_{\omega\in\Omega}=(s_i(\omega))_{\omega\in\Omega}$. Since $(s^k_{-i}(\omega))_{\omega\in\Omega}$ converges to $(\underline{s}_{-i}(\omega))_{\omega\in\Omega}$, $(s_i(\omega))_{\omega\in\Omega}$ is a best reply to $(\underline{s}_{-i}(\omega))_{\omega\in\Omega}$ and gives $\underline{v}_i^\omega(\delta)$ to player $i$. Therefore for any $\varepsilon>0$, there is $k$ such that the Markov strategy profile corresponding to $(s^k(\omega))_{\omega\in\Omega}$ satisfies all the desired conditions. *Q.E.D.*

## A.3 Proof of Lemma 5

**Lemma 5.** *For each initial state $\omega$, Markov strategy profile $s \in S^{PFR}$, and $\overline{\delta} \in (0,1)$, there is $\kappa > 0$ such that for each regular direction $\lambda \in \Lambda$, there is $K > 0$ such that for each $\delta \in (\overline{\delta},1)$ and for each $v \in \boldsymbol{R}^I$ such that $\lambda \cdot v^{\omega}(\delta,s) \geq \lambda \cdot v$, there is $w$ such that*

*(i)* *$(s,v)$ is enforced by $w$ for $\delta$ at $\omega$,*

*(ii)* *$\lambda \cdot w(h^t)$ is independent of $t$ and $h^t$ and $\lambda \cdot w(h^t) \leq \lambda \cdot v - (1-\delta)(\lambda \cdot v^{\omega}(\delta,s) - \lambda \cdot v)$, and*

*(iii)* *$|v - w(h^t)| < (1-\delta)(\kappa |v^{\omega}(\delta,s) - v| + K)$ for each $t$ and $h^t$.*

*Proof.* Fix $\omega$, $s$, $\lambda$, $\delta$, and $v$. Since $s \in S^{\mathrm{PFR}}$, for each $\omega'$ there is an integer $T(\omega') > 0$ and a sequence $(\omega^t(\omega'))_{t=1}^{T(\omega')}$ of states such that $\omega^1(\omega') = \omega'$, $\omega^T(\omega') = \omega$, and for each $t < T(\omega')$, profile $s(\omega^{t-1}(\omega'))$ has pairwise full rank for each $(i,j)$ for $(\omega^{t-1}(\omega'), \omega^t(\omega'))$. To simplify the notation, we denote this sequence by $\vec{\omega}(\omega')$, that is, $\vec{\omega}(\omega') = (\omega^t(\omega'))_{t=1}^{T(\omega')}$ Note that $\vec{\omega}(\omega')$ is a path from $\omega'$ to $\omega$ that occurs with positive probability even if player $i$ unilaterally deviates.

Given each $\omega'$, consider the "return game" such that the game starts at initial state $\omega'$ (instead of $\omega$) and ends when the state reaches $\omega$, and player $i$ obtains a payment $\tilde{v}_i$ at the end of the game, where

$$\tilde{v}_i = \frac{v_i - (1-\delta)\left[g_i^{\omega}(s(\omega)) + \sum_{\tau=1}^{\infty} \sum_{\tilde{h}^{\tau} \in \tilde{H}^{\omega,\tau}} \sum_{\omega'' \neq \omega} \delta^{\tau} \Pr(\tilde{h}^{\tau}, \omega'' | s, \omega) g_i^{\omega''}(s(\omega''))\right]}{\sum_{\tau=1}^{\infty} \sum_{h^{\tau} \in \tilde{H}^{\omega,\tau}} \delta^{\tau} \Pr(h^{\tau}, \omega | s, \omega)}. \tag{1}$$

Let $v_i^{\omega'}$ denote player $i$'s payoff in this return game when players follow the Markov strategy profile $s$, that is,

$$\begin{aligned} v_i^{\omega'} =& (1-\delta)\left[g_i^{\omega'}(s(\omega')) + \sum_{\tau=1}^{\infty} \sum_{\tilde{h}^{\tau} \in \tilde{H}^{\omega,\tau}} \sum_{\omega'' \neq \omega} \delta^{\tau} \Pr(\tilde{h}^{\tau}, \omega'' | s, \omega') g_i^{\omega''}(s(\omega''))\right] \\ &+ \sum_{\tau=1}^{\infty} \sum_{h^{\tau} \in \tilde{H}^{\omega,\tau}} \delta^{\tau} \Pr(h^{\tau}, \omega | s, \omega') \tilde{v}_i. \end{aligned}$$

Also, let $v_i^{\omega'}(a_i^1)$ denote the analogous value when player $i$ chooses $a_i^1 \in A_i$ in the initial period of the return game and then follows $s$. Note that, from (1), we have

$v_i^{\omega} = v_i$. That is, the payment $\tilde{v}_i$ is chosen in such a way that player $i$'s payoff in the return game from initial state $\omega$ with constant payment $\tilde{v}_i$ is equal to $v_i$.

Choose $\{z_i^{\omega'}(y^1)\}_{(i,y^1)\in \boldsymbol{I}\times Y}$ in such a way that player $i$ is indifferent over all actions in the first period of the return game from initial state $\omega'$ with constant payment $\tilde{v}_i$ supposing that all players follow $s$ in all periods $t > 1$ and that player $i$ receives a bonus $z_i^{\omega'}(y^1)$ in period two when the evolution of states follows $(\omega^1(\omega'), \omega^2(\omega'))$, while he receives no bonus when the state reaches other state $\omega'' \neq \omega^2(\omega')$ in period two. That is, let $\{z_i^{\omega'}(y^1)\}_{(i,y^1)\in \boldsymbol{I}\times Y}$ be such that

$$v_i^{\omega'} = v_i^{\omega'}(a_i^1) + \delta \sum_{y_1 \in Y} z_i^{\omega'}(y^1)\pi^{\omega^1(\omega')}(y^1, \omega^2(\omega')|a_i^1, s_{-i}(\omega^1(\omega'))) \tag{2}$$

for all $i$ and $a_i^1 \in A_i$. The existence of such $\{z_i^{\omega'}(y^1)\}$ is guaranteed, because the profile $s(\omega^1(\omega'))$ has pairwise full rank for $(\omega^1(\omega'), \omega^2(\omega'))$ so that the state moves from $\omega^1(\omega')$ to $\omega^2(\omega')$ with positive probability regardless of player $i$'s action, and the coefficient matrix of system (2) has full rank. In particular, as in Lemmas 5.3 and 5.4 of FLM, $\{z_i^{\omega'}(y^1)\}$ can be set to lie on a hyperplane orthogonal to $\lambda$, that is,

$$\sum_{i\in \boldsymbol{I}} \lambda_i z_i^{\omega'}(y^1) = 0 \tag{3}$$

for all $y^1 \in Y$.[10]

For $\omega'$ such that $T(\omega') > 2$, we recursively construct $\{z_i^{\omega'}(y^1, \cdots, y^t)\}_{(i,y^1,\cdots,y^t)}$ for each $t \in \{2, \cdots, T(\omega') - 1\}$ as follows. Intuitively, $z_i^{\omega'}(y^1, \cdots, y^t)$ is a bonus paid to player $i$ in period $t+1$ of the return game from $\omega'$ with constant payment $\tilde{v}_i$, when the evolution of states is precisely $(\omega^\tau(\omega'))_{\tau=1}^{t+1}$. Let $t \in \{2, \cdots, T(\omega') - 1\}$ and $\{z_i^{\omega'}(y^1, \cdots, y^{t-1})\}$ be given. We choose $\{z_i^{\omega'}(y^1, \cdots, y^t)\}$ in such a way that, in the return game from state $\omega'$, giving the bonus $z_i^{\omega'}(y^1, \cdots, y^{t-1})$ in period $t$ is equivalent to giving the bonus $z_i^{\omega'}(y^1, \cdots, y^t)$ in period $t+1$, regardless of player $i$'s play in period $t$. That is,

$$z_i^{\omega'}(y^1, \cdots, y^{t-1}) = \delta \sum_{y^t \in Y} z_i^{\omega'}(y^1, \cdots, y^t)\pi^{\omega^t(\omega')}(y^t, \omega^{t+1}(\omega')|a_i^t, s_{-i}(\omega^t(\omega'))) \tag{4}$$

[10]Here each row of the coefficient matrix $\Pi_{(i,j)}^{(\omega^1(\omega'),\omega^2(\omega'))}(s(\omega^1(\omega'))$ may not be a probability distribution and some rows can have larger norm than others. But since the rows of the matrix are linearly independent the resulting linear system still has a solution.

for all $i$ and $a_i^t \in A_i$. Again, the existence of the $\{z_i^{\omega'}(y^1,\cdots,y^t)\}$is guaranteed because the profile $s(\omega^t(\omega'))$ has pairwise full rank for $(\omega^t(\omega'),\omega^{t+1}(\omega'))$. Also, $\{z_i^{\omega'}(y^1,\cdots,y^t)\}$ can be set to lie on a hyperplane orthogonal to $\lambda$. That is,

$$\sum_{i\in \boldsymbol{I}} \lambda_i z_i^{\omega'}(y^1,\cdots,y^t) = 0 \tag{5}$$

for all $(y^1,\cdots,y^t)$.

By construction, in the return game from state $\omega'$, giving the bonus $z_i^{\omega'}(y^1)$ in period two when the evolution of states follows $(\omega^1(\omega'),\omega^2(\omega'))$ is equivalent to giving the bonus $z_i^{\omega'}(y^1,\cdots,y^{T(\omega')-1})$ in period $T(\omega')$ when the evolution of states follows the sequence $\vec{\omega}(\omega')$. Therefore, player $i$ is indifferent over all actions in the first period of the return game from state $\omega'$ with constant payment $\tilde{v}_i$, if players do not deviate from $s$ in period $t>1$ and if player $i$ receives $z_i^{\omega'}(y^1,\cdots,y^{T(\omega')-1})$ in period $T(\omega')$ when the evolution of states follows the sequence $\vec{\omega}(\omega')$, with no bonus if the evolution of states does not follow this sequence. Indeed, from (2) and (4), we obtain

$$\begin{aligned} v_i^{\omega'} =& v_i^{\omega'}(a_i^1) \\ &+ \delta^{T(\omega')} \sum_{(y_1,\cdots,y^{T(\omega')-1})} z_i^{\omega'}(y^1,\cdots,y^{T(\omega')-1}) \prod_{t=1}^{T(\omega')-1} \pi^{\omega^t(\omega')}(y^t,\omega^{t+1}(\omega')|a_i^t, s_{-i}(\omega^t(\omega'))) \end{aligned}$$

for all $i$ and $(a_i^1,\cdots,a_i^{T(\omega')-1}) \in (A_i)^{T(\omega')-1}$, which implies indifference in the first period. Note also that the above equality holds for all $(a_i^1,\cdots,a_i^{T(\omega')-1}) \in (A_i)^{T(\omega')-1}$, so that player $i$'s action in period $t>1$ of the return game does not affect the expectation of $z_i^{\omega'}(y^1,\cdots,y^{T(\omega')-1})$. Moreover, from (3) and (5), the $\{z_i^{\omega'}(y^1,\cdots,y^{T(\omega')-1})\}$ lie on the hyperplane orthogonal to $\lambda$, i.e.,

$$\sum_{i\in \boldsymbol{I}} \lambda_i z_i^{\omega'}(y^1,\cdots,y^{T(\omega')-1}) = 0 \tag{6}$$

for all $(y^1,\cdots,y^{T(\omega')-1})$.

Now we are in a position to construct $w$. We construct $w$ in such a way that at the end of the return game, player $i$ receives a constant payment $\tilde{v}_i$ plus a bonus contingent on the past public history. Specifically, given any $t$ and $\omega'$, if the state in period $t$ is $\omega'$ and if the evolution of states follows the sequence $\vec{\omega}(\omega')$ thereafter, then player $i$ receives a bonus $z_i^{\omega'}(y^t,\cdots,y^{t+T(\omega')-2})$ in period $t+T(\omega')-1$ (i.e.,

when the state reaches the end of the chain $\omega^{T(\omega')}(\omega') = \omega$ so that the game ends). By the definition of $z_i^{\omega'}$, this return payoff makes player $i$ indifferent over all actions in any period $t$ with any state $\omega'$. Note that the sequences for two (or more) different states might overlap; for example, there might be $\omega'$ and $\omega'' \neq \omega'$ such that for some $t$, $(\omega^\tau(\omega'))_{\tau=t}^{T(\omega')} = \vec{\omega}(\omega'')$. In this case we construct $w$ in such a way that player $i$ receives both $z_i^{\omega'}$ and $z_i^{\omega''}$ simultaneously if the recent evolution of states is $\vec{\omega}(\omega')$.

The formal specification of $w$ is as follows. For each history $h^t = (\omega^\tau, y^\tau)_{\tau=1}^t \in \tilde{H}^{\omega,t}$, let $\Omega(h^t)$ be the set of all $\omega' \in \Omega$ such that the sequence of states for the last $T(\omega') - 1$ periods is identical to $(\omega^1(\omega'), \cdots, \omega^{T(\omega')-1}(\omega'))$, i.e., $\Omega(h^t)$ is the set of all $\omega' \in \Omega$ such that $(\omega^{t-T(\omega')+2}, \cdots, \omega^t) = (\omega^1(\omega'), \cdots, \omega^{T(\omega')-1}(\omega'))$. Note that if the history up to period $t$ is $h^t$ and the state returns to $\omega$ in period $t+1$, then the recent evolution of states is exactly the sequence $\vec{\omega}(\omega')$ for each $\omega' \in \Omega(h^t)$. Thus we pay $z_i^{\omega'}(y^{t-T(\omega')+2}, \cdots, y^t)$ to player $i$ after such a history. Specifically, we set

$$w_i(h^t) = \tilde{v}_i + \sum_{\omega' \in \Omega(h^t)} z_i^{\omega'}(y^{t-T(\omega')+2}, \cdots, y^t).$$

That is, the continuation payoff $w_i(h^t)$ when the public history up to period $t$ is $h^t$ and the state returns to $\omega$ in period $t+1$ is a sum of the constant payment $v_i^\omega(\delta, s)$ and the bonuses $z_i^{\omega'}(y^{t-T(\omega')+2}, \cdots, y^t)$ for all $\omega'$ such that the recent evolution of states is exactly $\vec{\omega}(\omega')$.

This $w$ enforces $(s, v)$ because it makes each player exactly indifferent between all actions, and it yields payoff $v$, by construction. (Recall that $\tilde{v}_i$ is chosen so that $v_i^\omega = v_i$.) Therefore clause (i) follows, and it remains to show clauses (ii) and (iii).

To simplify the notation, let

$$x^{\omega'} = \sum_{\tau=1}^{\infty} \sum_{h^\tau \in \tilde{H}^{\omega,\tau}} \delta^\tau \Pr(h^\tau, \omega | s, \omega')$$

and

$$x_i^{\omega'} = g_i^{\omega'}(s(\omega')) + \sum_{\tau=1}^{\infty} \sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}} \sum_{\omega'' \neq \omega} \delta^\tau \Pr(\tilde{h}^\tau, \omega'' | s, \omega') g_i^{\omega''}(s(\omega'')).$$

Note that player $i$'s payoff in the original stochastic game from state $\omega$ is $v_i^\omega(\delta, s)$, if players play $s$. Also, since $s$ is Markov, player $i$'s payoff in the continuation game such that the current state is $\omega$ is $v_i^\omega(\delta, s)$ as well. Therefore player $i$'s

payoff when players play $s$ in the return game from state $\omega$ with constant payment $v_i^\omega(\delta,s)$ is equal to $v_i^\omega(\delta,s)$; that is,

$$v_i^\omega(\delta,s) = (1-\delta)x_i^\omega + x^\omega v_i^\omega(\delta,s).$$

Arranging,

$$v_i^\omega(\delta,s) = \frac{(1-\delta)x_i^\omega}{1-x^\omega}. \tag{7}$$

Recall that $\tilde{v}_i = \frac{v_i-(1-\delta)x_i^\omega}{x^\omega}$. Plugging (7) to this, we have

$$\tilde{v}_i = v_i - \frac{1-x^\omega}{x^\omega}(v_i^\omega(\delta,s) - v_i) \tag{8}$$

for all $i \in \boldsymbol{I}$ so that

$$\lambda\cdot\tilde{v} = \lambda\cdot v - \frac{1-x^\omega}{x^\omega}(\lambda\cdot v^\omega(\delta,s) - \lambda\cdot v).$$

Then it follows from (6) that

$$\lambda\cdot w(h^t) = \lambda\cdot v - \frac{1-x^\omega}{x^\omega}(\lambda\cdot v^\omega(\delta,s) - \lambda\cdot v)$$

for all $t$ and $h^t$. Since $0 < x^\omega < \delta$, we have $\frac{1-x^\omega}{x^\omega} > 1-\delta$ for any $\delta$. Then for any $\delta$ and $v$, $\lambda\cdot w(h^t) \le \lambda\cdot v - (1-\delta)(\lambda\cdot v^\omega(\delta,s) - \lambda\cdot v)$. This proves clause (ii).

Next we prove (iii). Letting $p_i$ be as in Lemma 1, it follows that if players $-i$ play the strategy $s$, then given any period $t$, the state reaches $\omega$ within $|\Omega|$ periods with at least probability $p_i$. Therefore,

$$\begin{aligned}
1-x^{\omega'} &= 1 - \sum_{\tau=1}^{\infty}\sum_{h^\tau\in\tilde{H}^{\omega,\tau}} \delta^\tau \Pr(h^\tau,\omega|s,\omega') \\
&\le 1 - \sum_{\tau=1}^{\infty} \delta^{\tau|\Omega|}(1-p_i)^{\tau-1}p_i \\
&= 1 - \frac{p_i\delta^{|\Omega|}}{1-\delta^{|\Omega|}(1-p_i)} \\
&= \frac{1-\delta^{|\Omega|}}{1-\delta^{|\Omega|}(1-p_i)} \\
&\le (1-\delta)\frac{|\Omega|}{p_i}
\end{aligned}$$

so that $1-x^{\omega'}$ is of order $(1-\delta)$. Also, letting $g_i^* = \max_{\omega' \in \Omega} \max_{a \in A} |g_i^{\omega'}(a)|$, we have

$$\begin{aligned} |x_i^{\omega'}| &\leq |g_i^{\omega'}(s(\omega'))| + \sum_{\tau=1}^{\infty} \sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}} \sum_{\omega'' \neq \omega} \delta^\tau \Pr(\tilde{h}^\tau, \omega'' | s, \omega') \left| g_i^{\omega''}(s(\omega'')) \right| \\ &\leq \sum_{\tau=0}^{\infty} (1-p_i)^\tau |\Omega| g_i^* = \frac{|\Omega| g_i^*}{p_i} \end{aligned}$$

so that $(1-\delta)x_i^{\omega'}$ is of order $(1-\delta)$.

Since $v_i^{\omega'} = (1-\delta)x_i^{\omega'} + x^{\omega'}\tilde{v}_i$ and both $(1-\delta)x_i^{\omega'}$ and $1-x^{\omega'}$ are of order $(1-\delta)$, it follows that $v_i^{\omega'} - \tilde{v}_i$ is of order $(1-\delta)$. Likewise, one can check that $v_i^{\omega'}(a_i) - \tilde{v}_i$ is of order $(1-\delta)$ for all $a_i \in A_i$. Then without loss of generality we can let $z_i^{\omega'}(y^1)$ be of order $(1-\delta)$. (This follows from the fact that $v_i^{\omega'} - v_i^{\omega'}(a_i)$ is of order $(1-\delta)$ and that the coefficient matrix corresponding to (2) and (3) is independent of $\delta$. If the coefficient matrix has many more columns than rows, then there are infinitely many solutions, but we can still select $z_i^{\omega'}(y^1)$ so that it is of order $(1-\delta)$.) Similarly, we can let $z_i^{\omega'}(y^1, \cdots, y^t)$ be of order $(1-\delta)$. Thus there is $K > 0$ such that $\sum_{i \in \boldsymbol{I}} |\sum_{\omega' \in \Omega(h^t)} z_i^{\omega'}(y^{t-T(\omega')+2}, \cdots, y^t)| < K(1-\delta)$ for any $t$ and $h^t$. Also, since $1-x^\omega$ is of order $(1-\delta)$, given any $\overline{\delta} \in (0,1)$, there is $\kappa > 0$ such that $\frac{1-x^\omega}{x^\omega} < (1-\delta)\kappa$ for any $\delta \in (\overline{\delta}, 1)$. Then it follows from (8) that $|\tilde{v} - v| < (1-\delta)\kappa |v^\omega(\delta, s) - v|$ for any $\delta \in (\overline{\delta}, 1)$ and $v$. This proves clause (iii). *Q.E.D.*

## A.4 Proof of Lemma 6

**Lemma 6.** *For each initial state $\omega$, Markov strategy $s_{-i} \in S_{-i}^{IFR}(i)$, and $\overline{\delta} \in (0,1)$, there is $\kappa > 0$ such that for each direction $\lambda$ such that $\lambda_i \neq 0$ and $\lambda_j = 0$ for all $j \neq i$, there is $K > 0$ such that for each $\delta \in (\overline{\delta}, 1)$ and for each $v \in \boldsymbol{R}^I$ such that $\lambda_i v_i \leq \lambda_i \max_{s_i'} v_i^\omega(\delta, s_i', s_{-i})$, there is player $i$'s pure Markov strategy $s_i$ and $w$ such that*

*(i)* *$(s, v)$ is enforced by $w$ for $\delta$ at $\omega$,*

*(ii)* *$\lambda_i w_i(h^t)$ is independent of $t$ and $h^t$ and $\lambda_i w_i(h^t) \leq \lambda_i v_i - (1-\delta)\lambda_i(\max_{s_i'} v_i^\omega(\delta, s_i', s_{-i}) - v_i)$, and*

*(iii)* *$|v - w(h^t)| < (1-\delta)(\kappa |v^\omega(\delta, s) - v| + K)$ for each $t$ and $h^t$.*

*Proof.* Fix $s_{-i} \in S_{-i}^{\mathrm{IFR}}(i)$, $\lambda$, $\delta$, and $v$. we find $s_i$ and $w(h^t)$ satisfying the desired conditions.

We first specify $s_i$. Given any Markov strategy $s_i'$, let

$$\tilde{v}_i(s_i') = \frac{v_i - (1-\delta)\left[\begin{array}{l} g_i^{\omega}(s_i'(\omega), s_{-i}(\omega)) \\ +\sum_{\tau=1}^{\infty}\sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}}\sum_{\omega'' \neq \omega}\delta^\tau \Pr(\tilde{h}^\tau, \omega'' | s_i', s_{-i}, \omega) g_i^{\omega''}(s_i'(\omega''), s_{-i}(\omega'')) \end{array}\right]}{\sum_{\tau=1}^{\infty}\sum_{h^\tau \in \tilde{H}^{\omega,\tau}}\delta^\tau \Pr(h^\tau, \omega | s_i', s_{-i}, \omega)}.$$

That is, $\tilde{v}_i(s_i')$ is chosen in such a way that $v_i$ is achieved as player $i$'s payoff in the return game from $\omega$ to $\omega$ when players play $(s_i', s_{-i})$ and player $i$ receives the constant payment $\tilde{v}_i(s_i')$ at the end of the return game. Let $F(s_i)$ be the set of player $i$'s optimal Markov strategies against $s_{-i}$ in the return game from state $\omega$ with constant payment $\tilde{v}_i(s_i)$. The correspondence $F$ is non-empty-valued, convex valued, and has closed graph, so by Kakutani's fixed point theorem, there is $\tilde{s}_i$ such that $\tilde{s}_i \in F(\tilde{s}_i)$. We denote $\tilde{v}_i(\tilde{s}_i)$ by $\tilde{v}_i$, and let $s_i$ be player $i$'s pure Markov strategy such that $s_i$ is a best reply against $s_{-i}$ in the return game with constant payment $\tilde{v}_i$. By construction, player $i$'s payoff in the return game from state $\omega$ with constant payment $\tilde{v}_i$ is $v_i$, if players play $s$.

Next we construct $w$. For each $j \neq i$, let $\tilde{v}_j$, $v_j^{\omega'}$, and $v_j^{\omega'}(a_j)$ be as in the proof of Lemma 5. Since $s_{-i} \in S_{-i}^{\mathrm{IFR}}(\delta)$, for each $\omega'$ there is an integer $T(\omega') > 0$ and a sequence $(\omega^t(\omega'))_{t=1}^T$ of states such that $\omega^1(\omega') = \omega'$, $\omega^T(\omega') = \omega$, and for each $t < T(\omega')$, the profile $s(\omega^{t-1}(\omega'))$ has individual full rank for $(\omega^{t-1}(\omega'), \omega^t(\omega'))$ and for all $j \neq i$. Then as in the proof of Lemma 5, there are $\{z_j^{\omega'}(y^1, \cdots, y^{T(\omega')-1})\}$ such that

$$v_j^{\omega'} = v_i^{\omega'}(a_j^1) + \delta^{T(\omega')} \sum_{(y^1,\cdots y^{T(\omega')-1})} z_j^{\omega'}(y^1, \cdots, y^{T(\omega')-1}) \prod_{t=1}^{T(\omega')-1} \pi^{\omega^t(\omega')}(y^t, \omega^{t+1}(\omega') | a_j^t, s_{-j}(\omega^t(\omega')))$$

for all $j \neq i$ and $(a_j^1, \cdots, a_j^{T(\omega')-1})$. Let $z_i^{\omega'}(y^1, \cdots, y^{T(\omega')-1}) = 0$. For each history $h^t$, let $\Omega(h^t)$ be the set of all $\omega' \in \Omega$ such that the sequence of states for the last $T(\omega') - 1$ periods is identical to $(\omega^1(\omega'), \cdots, \omega^{T(\omega')-1}(\omega'))$. Then let

$$w_j(h^t) = \tilde{v}_j + \sum_{\omega' \in \Omega(h^t)} z_j^{\omega'}(y^{t-T(\omega')+2}, \cdots, y^t)$$

for each $j \in \boldsymbol{I}$.

Now we show that the above $(s,w)$ satisfies all the desired conditions. As in the proof of Lemma 5, player $j \neq i$ is indifferent over all actions in every period so that his incentive compatibility is satisfied. Also, player $i$'s incentive compatibility is satisfied as well, because $s_i$ is defined to be a best reply in the return game. In addition, $v$ is generated using $(s,w)$, as $\tilde{v}$ is chosen in such a way that the payoff of the return game is $v$. Therefore, clause (i) follows.

Let $s_i^*$ be a Markov strategy such that $s_i^* \in \max_{s_i'} w_i^\omega(\delta, s_i', s_{-i})$. Let

$$x^* = \sum_{\tau=1}^{\infty} \sum_{h^\tau \in \tilde{H}^{\omega,\tau}} \delta^\tau \Pr(h^\tau, \omega | s_i^*, s_{-i}, \omega)$$

and

$$x_i^* = g_i^{\omega'}(s_i^*(\omega), s_{-i}(\omega)) + \sum_{\tau=1}^{\infty} \sum_{\tilde{h}^\tau \in \tilde{H}^{\omega,\tau}} \sum_{\omega'' \neq \omega} \delta^\tau \Pr(\tilde{h}^\tau, \omega'' | s_i^*, s_{-i}, \omega) g_i^{\omega''}(s_i^*(\omega''), s_{-i}(\omega'')).$$

Note that player $i$'s payoff in the return game from state $\omega$ with constant payment $v_i^\omega(\delta, s_i^*, s_{-i})$ is equal to $v_i^\omega(\delta, s)$ when players play $(s_i^*, s_{-i})$; that is,

$$v_i^\omega(\delta, s_i^*, s_{-i}) = (1-\delta)x_i^* + x^* v_i^\omega(\delta, s_i^*, s_{-i}).$$

Arranging,

$$v_i^\omega(\delta, s_i^*, s_{-i}) = \frac{(1-\delta)x_i^*}{1-x^*}. \tag{9}$$

Recall that playing $s_i$ is optimal in the return game from state $\omega$ with constant payment $\tilde{v}_i$, and its payoff is $v_i$. Therefore playing $s_i^*$ can yield at most $v_i$ in this return game, i.e.,

$$v_i \geq (1-\delta)x_i^* + x^* \tilde{v}_i.$$

Plugging (9) to this,

$$\tilde{v}_i \leq v_i - \frac{1-x^*}{x^*}(v_i^\omega(\delta, s_i^*, s_{-i}) - v_i).$$

Then for $\lambda_i > 0$,

$$\lambda_i w_i(h^t) = \lambda_i \tilde{v}_i \leq \lambda_i v_i - \frac{1-x^*}{x^*}\lambda_i(v_i^\omega(\delta, s_i^*, s_{-i}) - v_i).$$

Also, letting $x^\omega$ be as in the proof of Lemma 5, we have (8) so that for $\lambda_i < 0$,

$$\lambda_i w_i(h^t) = \lambda_i \tilde{v}_i = \lambda_i v_i - \frac{1-x^\omega}{x^\omega}\lambda_i(v_i^\omega(\delta, s) - v_i) \leq \lambda_i v_i - \frac{1-x^\omega}{x^\omega}\lambda_i(v_i^\omega(\delta, s_i^*, s_{-i}) - v_i).$$

Then clause (ii) follows for either case, since $0 < x^* < \delta$ and $0 < x^\omega < \delta$. Also, clause (iii) follows as in the proof of Lemma 5. Note in particular that we can choose $\kappa$ and $K$ independently of the specification of $s_i$, as the number of player $i$'s pure Markov strategies is finite. *Q.E.D.*

## A.5 Proof of Lemma 7

**Lemma 7.** *Suppose that (IFR) and (PFR) hold. Then for any initial state $\omega \in \Omega$ and for any smooth subset $W$ of the interior of $V^*$, there are $\varepsilon > 0$ and $\overline{\delta} \in (0,1)$ such that the following properties hold:*

*(a) There is a finite set of Markov strategy profiles $\{s^1, \cdots, s^N\}$ such that $s^n \in S^{PFR}$ for all $n$, and for every $\delta \in (\overline{\delta}, 1)$ and for every regular $\lambda \in \Lambda$, $\max_{n \in \{1,\cdots,N\}} \lambda \cdot v^n > \max_{v \in W} \lambda \cdot v + \varepsilon$, where $v^n = v^\omega(\delta, s^n)$.*

*(b) For each $i \in \boldsymbol{I}$ and for each $\lambda$ such that $|\lambda_i| = 1$ and $\lambda_j = 0$ for all $j \neq i$, there is $s_{-i} \in S_{-i}^{IFR}(i)$ such that for every $\delta \in (\overline{\delta}, 1)$, there is player $i$'s Markov strategy $s_i$ such that $s = (s_i, s_{-i}) \in S^{IFR}(\delta, i)$ and $\lambda \cdot v > \max_{v' \in W} \lambda \cdot v' + \varepsilon$, where $v = v^\omega(\delta, s)$.*

*Proof.* Since $W$ is in the interior of $V^*$, there is $\varepsilon > 0$ such that $\max_{v \in V^*} \lambda \cdot v > \max_{v \in W} \lambda \cdot v + 6\varepsilon$ for all $\lambda \in \Lambda$. In what follows, we consider a fixed initial state $\omega$.

Part (a). Let $\{\tilde{v}^1, \cdots, \tilde{v}^N\}$ be the finite set of all extreme points of the feasible set $V$. As Dutta (1995) shows, each extreme point of $V$ is attainable by a pure Markov strategy profile in the no-discounting case. Let $\{\tilde{s}^1, \cdots, \tilde{s}^N\}$ be the set of pure Markov strategy profiles that achieve these extreme points. Then from Lemma 3, there is a nearby Markov strategy profile $s^n$ such that $s^n \in S^{\mathrm{PFR}}$ and $\lambda \cdot v^n$ is within $\varepsilon$ of $\lambda \cdot \tilde{v}^n$ for sufficiently large $\delta$, where $v^n = v^\omega(\delta, s^n)$. (The latter condition holds from Abel's theorem: the limit payoff of a Markov strategy profile is equal to its time-average payoff. See also (A.2) of Dutta (1995).) This proves part (a).

Part (b). We first prove the following claim: There is $\overline{\delta} \in (0,1)$ such that for any $\delta \in (\overline{\delta}, 1)$, for any $i \in I$, and for any Markov strategy profile $s$, $|v_i^\omega(\delta, s) - \lim_{\delta \to 1} v_i^\omega(\delta, s)| < \varepsilon$. The key here is that $\overline{\delta} \in (0,1)$ is cohesions independently

of $i$ and $s$. To prove the claim, note first that for a given pure Markov strategy profile $s$, there is $\overline{\delta} \in (0,1)$ such that $|v_i^\omega(\delta,s) - \lim_{\delta' \to 1} v_i^\omega(\delta',s)| < \varepsilon$ for any $i \in I$ and $\delta \in (\overline{\delta},1)$. Since the action space and the state space are finite, it follows that there is $\overline{\delta} \in (0,1)$ such that $|v_i^\omega(\delta,s) - \lim_{\delta' \to 1} v_i^\omega(\delta',s)| < \varepsilon$ for any $i \in I$, for any $\delta \in (\overline{\delta},1)$, for any pure Markov strategy profile $s$. One can check that this inequality remains true even for a mixed Markov strategy profile $s$, since the payoff from a mixed Markov strategy profile is a convex combination of those from pure Markov strategy profiles. This shows the claim. In what follows, let $\overline{\delta}$ be as in the claim.

Consider $\lambda$ such that $\lambda_i = 1$ and $\lambda_j = 0$ for all $j \neq i$. Given a $\delta$, let $s(\delta)$ be a strategy profile that gives the highest discounted average payoff to player $i$ when the initial state is $\omega$. Without loss of generality we assume that $s(\delta)$ is pure and Markov for all $\delta$. It is easy to see that for $\delta$ sufficiently close to one, player $i$'s discounted average payoff from $s(\delta)$ is within $\varepsilon$ of his highest payoff of the limit feasible set $V$. This, together with the above claim, shows that there is $\delta^* \in (\overline{\delta},1)$ such that for every $\delta \in (\delta^*,1)$,

$$\left| v_i^\omega(\delta, s(\delta^*)) - \max_{v' \in V} \lambda \cdot v' \right| < 2\varepsilon. \tag{10}$$

Choose such $\delta^*$ and let $s^*$ denote $s(\delta^*)$. For each $\delta \in (\delta^*,1)$, let $s_i^\delta$ be player $i$'s pure Markov strategy that is optimal against $s_{-i}^*$ for $\delta$. We will show that for each $\delta \in (\delta^*,1)$, the profile $(s_i^\delta, s_{-i}^*)$ satisfies the desired properties.

Since $s^*$ gives the highest payoff to player $i$ for $\delta^*$, we have $v_i^\omega(\delta^*,s^*) \geq v_i^\omega(\delta^*, s_i^\delta, s_{-i}^*)$. Likewise, since $s_i^\delta$ is a best reply for $\delta$, we have $v_i^\omega(\delta,s^*) \leq v_i^\omega(\delta, s_i^\delta, s_{-i}^*)$. Also, from the above claim, we obtain $|v_i^\omega(\delta^*,s^*) - v_i^\omega(\delta,s^*)| < 2\varepsilon$ and $|v_i^\omega(\delta^*, s_i^\delta, s_{-i}^*) - v_i^\omega(\delta, s_i^\delta, s_{-i}^*)| < 2\varepsilon$ for every $\delta \in (\delta^*,1)$. These inequalities yield $|v_i^\omega(\delta^*,s^*) - v_i^\omega(\delta, s_i^\delta, s_{-i}^*)| < 2\varepsilon$. Plugging this into (10), it follows that

$$\left| v_i^\omega(\delta, s_i^\delta, s_{-i}^*) - \max_{v' \in V} \lambda \cdot v' \right| < 4\varepsilon$$

for every $\delta \in (\delta^*,1)$. Since $\max_{v \in V} \lambda \cdot v > \max_{v \in V^*} \lambda \cdot v > \max_{v \in W} \lambda \cdot v + 6\varepsilon$, it follows that $\lambda \cdot v^\omega(\delta, s_i^\delta, s_{-i}^*) > \max_{v \in W} \lambda \cdot v + \varepsilon$, as desired. Also, by definition, $(s_i^\delta, s_{-i}^*) \in S^{\mathrm{IFR}}(\delta, i)$.

Next we consider $\lambda$ such that $\lambda_i = -1$ and $\lambda_j = 0$ for all $j \neq i$. Lemma 4 says that for any $\delta$, there is a Markov strategy profile $s(\delta)$ such that $s(\delta) \in S^{\mathrm{IFR}}(\delta, i)$

and player $i$'s payoff is within $\varepsilon$ of $\underline{v}_i(\delta)$. Since $\underline{v}_i(\delta)$ approximates $\max_{v' \in V^*} \lambda \cdot v'$ as $\delta$ goes to one, the above claim implies that there is $\delta^* \in (\overline{\delta}, 1)$ such that for every $\delta \in (\delta^*, 1)$,

$$\left| v_i^{\omega}(\delta, s(\delta^*)) - \max_{v' \in V^*} \lambda \cdot v' \right| < 3\varepsilon.$$

Choose such $\delta^*$, and let $s^*$ denote $s(\delta^*)$. For each $\delta \in (\delta^*, 1)$, let $s_i^{\delta}$ be player $i$'s pure Markov strategy that is optimal against $s_{-i}^*$ for $\delta$. Then as in the previous case, we can show that this $(s_i^{\delta}, s_{-i}^*)$ satisfies the desired properties. *Q.E.D.*

# References


Abreu, D., D. Pearce, and E. Stacchetti (1990): "Toward a Theory of Discounted Repeated Games with Imperfect Monitoring," *Econometrica* 58, 1041-1063.

Athey, S., and K. Bagwell (2008): "Collusion with Persistent Cost Shocks," *Econometrica* 76,493-540.

Besanko, D., U. Doraszelski, Y. Kryukov, and M. Satterthwaite (2008): "Learning by Doing, Organizational Forgetting, and Industry Dynamics," forthcoming in *Econometrica*.

Bewley, T., and E. Kohlberg (1976): "The Asymptotic Theory of Stochastic Games," *Mathematics of Operations Research* 1, 200-208.

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

Ellison, G. (1994): "Theories of Cartel Stability and the Joint Executive Committee," *RAND Journal of Economics* 25, 37-57.

Escobar, J., and J. Toikka (2009) "A Folk Theorem with Markovian Private Information," *mimeo*.

Fudenberg, D., and D.K. Levine (1994): "Efficiency and Observability in Games with Long-Run and Short-Run Players," *Journal of Economic Theory* 62, 103-135.

Fudenberg, D., D.K. Levine, and E. Maskin (1994): "The Folk Theorem with Imperfect Public Information," *Econometrica* 62, 997-1040.

Fudenberg. D., and J. Tirole (1991): *Game Theory*, MIT Press, Cambridge, MA.

Hörner, J., and W. Olszewski (2009): "How Robust is the Folk Theorem?" *Quarterly Journal of Economics* 124, 1773-1814.

Hörner, J., T. Sugaya, S. Takahashi, and N. Vieille (2009): "Recursive Methods in Discounted Stochastic Games: an Algorithm for $\delta \to 1$ and a Folk Theorem," *mimeo*.

Kandori, M., and H. Matsushima (1998): "Private Observation, Communication and Collusion," *Econometrica* 66, 627-652.

Mertens, J.F., and A. Neyman (1981): "Stochastic Games," *International Journal of Game Theory* 10, 53-66.

Rotemberg, J., and G. Saloner (1986): "A Supergame-Theoretic Model of Price Wars during Booms," *American Economic Review* 76, 390-407.

Shapley, L. (1953): "Stochastic Games," *Proceedings of the National Academy of Sciences of the United States of America* 39, 1095-1100.

Sobel, M. (1971): "Noncooperative Stochastic Games," *The Annals of Mathematical Statistics* 42, 1930-1935.