Theoretical Economics 10 (2015), 131–173 1555-7561/20150131

A folk theorem for stochastic games with infrequent state changes

Marcin Pęski
Department of Economics, University of Toronto

Thomas Wiseman
Department of Economics, University of Texas at Austin

We characterize perfect public equilibrium payoffs in dynamic stochastic games in the case where the length of the period shrinks, but players' rate of time discounting and the transition rate between states remain fixed. We present a meaningful definition of the feasible and individually rational payoff sets for this environment, and we prove a folk theorem under imperfect monitoring. Our setting differs significantly from the case considered in previous literature (Dutta 1995, Fudenberg and Yamamoto 2011, and Hörner et al. 2011) where players become very patient. In particular, the set of equilibrium payoffs typically depends on the initial state.

Keywords. Stochastic games, folk theorem.

JEL classification. C72, C73.

1. Introduction

Stochastic games are generalizations of repeated games in which the payoffs in a period depend not only on the current action profile, but also on the value of a state variable, whose random evolution is itself influenced by players' actions. Stochastic games allow for dynamic interaction between players, but do not impose the strong restriction that the parameters of the interaction in one period are independent of outcomes in previous periods. Important economic examples are models with stock variables such as capital, savings, technology, brand awareness, or natural resource population; models with persistent shocks to demand, productivity, or income; models of durable goods markets; and political economy models where government policy changes at discrete intervals.

In many such settings, players get frequent opportunities to adjust their actions, while the state changes more rarely. For example, firms may vary research and development expenditure daily, but breakthroughs occur infrequently. Similarly, frequent price competition between oligopolists can be affected by less frequent events like aggregate shocks to the macroeconomy or the development of a new product in a related market.

Marcin Pęski: mpeski@gmail.com
Thomas Wiseman: wiseman@austin.utexas.edu
We thank Drew Fudenberg, Johannes Hörner, Chris Sleet, Nicolas Vieille, and Şevin Yeltekin for helpful conversations, and the referees for useful comments.

DOI: 10.3982/TE1512

So as to analyze such situations, we examine stochastic games as the length of a period shrinks, but the players' rate of time discounting and the state transition rate per unit of time (not per period) remain fixed. In the limit, the discounting between periods shrinks to zero, but the discounted time until a state transition does not. The set of equilibrium payoffs typically depends on the initial state even in the limit, because the discounted fraction of the game that players expect to spend in that state before the first transition is nonnegligible.

We consider such stochastic games with finitely or countably many states and imperfect public monitoring. (In each period, players observe the state and a noisy signal of the action profile just played.) Let $E^\delta(s)$ denote the set of perfect public equilibrium (PPE) payoffs given discount factor δ and initial state s. The main contribution of this paper is that we define two collections of payoff sets, $V_0 = \{V_0(s)\}_s$ and $V_{0+} = \{V_{0+}(s)\}_s$, such that (i) $V_{0+}(s) \subseteq \lim_{\delta\to 1} E^\delta(s) \subseteq V_0(s)$ for any state s, and (ii) if V_0 has nonempty interior, then generically $V_{0+} = V_0$. The first inclusion in (i) requires that Fudenberg et al.'s (1994) (FLM) conditions on public monitoring hold in each state: if the set of states is infinite, we require that the conditions hold uniformly, in a sense that we define. The definitions of V_0 and V_{0+} correspond to a particular notion of feasibility and individual rationality. One of the goals of the paper is to explain why other notions do not work, in that they do not deliver the set of PPE payoffs. Along the same lines, we demonstrate, through an example with finitely many states, that the equilibrium set may include payoffs that cannot be achieved through any stationary Markov strategy that is individually rational in each state. We also explore how the "right" notion of individual rationality may depend on the solution concept.

In our model, the set of feasible and individually rational payoffs depends on the initial state. Similarly, minmax payoffs vary with the current state. A strategy profile can be part of a PPE, then, only if it delivers continuation payoffs that are individually rational, given the current state, after each history: the payoffs thus achievable are *ex post individually rational* in Dutta's (1995) terminology. In fact, though, even a feasible, ex post individually rational payoff vector may not be achievable in equilibrium. From that fact, we conclude that the set of feasible, ex post individually rational payoffs is not the appropriate generalization of the feasible, individually rational set in standard repeated games.

Instead, we define for each player and each state the minmax payoff *relative* to a collection $F = \{F(s)\}_s$ of available continuation payoffs in each state, and we say that F is *self-individually rational* if for every state s, each payoff in $F(s)$ exceeds the state-s minmax payoff relative to F for each player. Similarly, we say that F is *self-feasible* if each payoff in $F(s)$ can be generated as the expected payoff from an action in state s followed, after a transition to any state s', by a continuation payoff in $F(s')$. Extending those definitions, for $\epsilon > 0$, we say that F is *self-ϵ-individually rational* if each payoff in $F(s)$ is at least ϵ greater than the state-s minmax payoff relative to F, and that F is *self-ϵ-feasible* if the ϵ-neighborhood around each payoff in $F(s)$ can be generated using continuation payoffs in F.

We define V_0 as the largest self-feasible and self-individually rational collection, and define V_{0+} as the union across $\epsilon > 0$ of the self-ϵ-feasible and self-ϵ-individually rational collections. We show (in Theorem 1) that under FLM's monitoring conditions, any

payoff vector $v \in V_{0+}(s)$ can be achieved in a PPE starting from initial state s, as long as the length of the period is short enough. Next (in Corollary 1 and Theorem 2), we show that V_0 contains all PPE payoffs, and that for generic (i.e., full Lebesgue measure), finite-state games that satisfy a full-dimensionality condition, V_0 is equal to V_{0+}. Thus, for such games, we obtain a complete characterization of PPE payoffs (as the length of the periods shrinks). Finally, even in games for which collection V_0 is strictly larger than the set of PPE payoffs, we argue (in Theorem 3) that it is a good approximation: every payoff in V_0 is a PPE payoff of a nearby game (again, given full dimensionality).

Previous work has focused on an alternative limiting case for stochastic games: fix the period length and let players become very patient. In that case, the discounted time until the state changes shrinks to zero. Dutta (1995) derives a folk theorem for that environment. Fudenberg and Yamamoto (2011) (FY) and Hörner et al. (2011) (HSTV) provide conditions on imperfect public monitoring under which Dutta's (1995) folk theorem extends. A key difference from our model is that all three of those results require that the set of PPE payoffs be independent of the initial state as the discount factor δ approaches 1.

To ensure that independence, Dutta (1995), FY, and HSTV use *irreducibility*, the condition that no single player's deviation can prevent the Markov process governing the state variable from being irreducible. We do not assume irreducibility, since in our model, $\lim_{\delta\to 1} E^{\delta}(s)$ typically varies with the state s in any case. We allow, for example, multiple absorbing states, each reachable from the initial state—the first firm to achieve a technological breakthrough might permanently capture the market, for instance.

The organization of the rest of the paper is as follows. In Section 2, we describe the model. In Section 3, we introduce the notions of self-feasible and self-individually rational collections of payoffs. In Sections 4 and 5, we present our main theorems. Section 6 contains an example of how to construct the maximal self-feasible and self-individually rational collection V_0. We discuss how using different equilibrium concepts affects our results in Section 7, and we describe the problems with using other notions of "feasible and individually rational" payoffs in analyzing stochastic games with infrequent state transitions. Section 8 is the conclusion.

2. Model

There are $N \geq 2$ expected-utility maximizing players playing an infinite-horizon stochastic game. The time between periods is given by $\Delta > 0$ and all players discount the future at rate $r > 0$, so that the per-period discount rate is $e^{-r\Delta} \equiv \delta$. There is a (finite or countably infinite) set S of states of the world. In each different state $s \in S$, there is a stage game $G(s)$ with a set of action profiles $A(s) = A_1(s) \times A_2(s) \times \cdots \times A_N(s)$, where $A_i(s)$ is the (finite) set of actions for player i. Let $m_i(s)$ denote the number of actions available to player i: $m_i(s) \equiv \#A_i(s)$. We assume that the total number of actions available in any state is bounded by $m^* \equiv \max_s \sum_i m_i(s) < \infty$. At the start of each period t, the state s is publicly observed. Each player i chooses an action $a_i \in A_i(s)$ and then all players observe a public signal y drawn from a finite set Y, which has m elements. The public signal is distributed according to $\rho(a, s)$, where a is the profile of actions of all players.

Player i's expected stage-game payoff in state s when action profile a is played is equal to $g_i(a,s)$. (As is typical in repeated games with imperfect monitoring, one can treat $g_i(a,s)$ as an expectation of a primitive utility that depends on the action profile and the public signal.) Denote by $g(a,s)$ the vector of expected payoffs for each player. We assume that the payoffs are uniformly bounded by $M \equiv \sup_{a,s} \|g(a,s)\| < \infty$ (where $\|\cdot\|$ denotes the Euclidean norm).

At the end of a period in which action profile a is played in state s, the probability that the state changes to state $s' \neq s$ is equal to $(1-\delta)\gamma(s';a,s)$. That is, the transition probability per *period* is proportional to the length of the period $\Delta \approx (1/r)(1-\delta)$, and the transition probability per unit of *time* (i.e., $(1-\delta)\gamma(s';a,s)/\Delta \approx r\gamma(s';a,s)$) has a nontrivial limit as $\delta \to 1$. With the remaining probability $1-(1-\delta)\gamma(a,s)$, where

$$\gamma(a,s) \equiv \sum_{s' \neq s} \gamma(s';a,s),$$

the state does not change. We assume that transition rates are uniformly bounded, i.e., $\gamma_{\max} \equiv \sup_{a,s} \gamma(a,s) < \infty$. We assume throughout that $(1-\delta)\gamma_{\max} < 1$, so that the maximum probability of transition in any period is less than 1.

All these definitions extend in a natural way to mixed and correlated action profiles α. This structure is common knowledge. A special case of a stochastic game is a standard repeated game, in which, for each state s and each action profile $a \in A(s)$, $\gamma(a,s) = 0$. Other (and less trivial) examples are given throughout the paper.

We assume that a public randomization device is available to the players. The set of *public histories* in period t is equal to $H_t \equiv Y^{t-1} \times S^t$, with element $h_t = (s_1, y_1, \ldots, y_{t-1}, s_t)$, where s_t denotes the state at the beginning of period t and y_t denotes the public signal realized at the end of period t. Player *i's private history* in period t is $h_t^i = (s_1, y_1, a_{i,1}, \ldots, y_{t-1}, a_{i,t-1}, s_t)$, where $a_{i,t}$ is player i's action in period t; $H_t^i \equiv (Y \times A_i)^{t-1} \times S^t$ is the set of such private histories. Define $H \equiv \bigcup_{t\geq 1} H_t$ and $H^i \equiv \bigcup_{t\geq 1} H_t^i$. For any history h_t, let $s(h_t) \equiv s_t$ denote the current state.

A strategy for player i is a mapping $\sigma_i : H^i \to \triangle A_i(s(h_t^i))$. A strategy is public if it depends only on public histories. Let Σ_i and Σ_i^P, respectively, denote the set of strategies and the set of public strategies for player i; let Σ and Σ^P denote the sets of strategy profiles and of public strategy profiles, respectively. Given discount factor $\delta < 1$, a profile of strategies σ, and an initial state $s \in S$, the vector of expected payoffs in the dynamic game is given by

$$v^\delta(\sigma,s) = (1-\delta)E\sum_{t=1}^{\infty} \delta^{t-1} g(a_t,s_t), \tag{1}$$

where the expectation is taken with respect to the distribution over actions and states induced by the profile σ and initial state s.[1] For each public strategy $\sigma \in \Sigma^P$ and public history $h \in H$, the continuation payoffs $v^\delta(\sigma,h)$ are calculated in the usual way.

[1] If the stage-game payoff is proportional to the period length, then (1) gives the total payoff of the dynamic game. On the other hand, if we interpret the stage-game payoff as constant (i.e., independent of the period length), then (1) represents the discounted average payoff of the game.

A *perfect public equilibrium* in a game with discount factor δ and initial state s is a public strategy profile σ such that for each public history h with $s_1 = s$, each player i, and each strategy σ_i' of player i, $v_i^\delta(\sigma, h) \geq v_i^\delta(\sigma_i', \sigma_{-i}, h)$. Let $E^\delta(s)$ be the set of payoffs obtained in perfect public equilibria of the game, given initial state s and discount factor δ. Note that, in contrast to the setting in FY and HSTV, the set of equilibrium payoffs $E^\delta(s)$ varies with the state even in the limit as δ approaches 1. This happens for two reasons: first, we do not assume irreducibility. More importantly, both the discount rate and the rate of transition between states are fixed per unit of time as δ grows, so the discounted time that the game spends in any given state before the next transition is nonnegligible.

In most of this paper, "equilibrium" means perfect public equilibrium. Section 7 explains how the results of this paper extend to other solution concepts.

In the case of a finite state space, we often assume that the monitoring structure satisfies the following condition.

IDENTIFIABILITY CONDITION. FLM's Conditions 6.2 and 6.3 hold in each state.[2]

For an infinite space S, we must modify the definition of the Identifiability Condition to ensure that FLM's requirements hold *uniformly* across all infinitely many states. To simplify the exposition, we delay the infinite-space Identifiability Condition until Appendix A. Unless specified otherwise, all results and proofs apply to the general, infinite-space case.

For future purposes, it is instructive to rewrite (1) as

$$v^\delta(\sigma, s) = (1-\delta)g(\sigma_1, s_1) + \delta\left(\sum_{s' \neq s}(1-\delta)\gamma(s'; \sigma_1, s)u_{s'}\right) + \delta(1-(1-\delta)\gamma(\sigma_1, s))u_s, \quad (2)$$

where $u_{s'}$ denotes the continuation payoff from period 2 onward, given period-2 state s'. Observe that the weights on both the period-1 payoffs and the continuation payoffs after a state change (the first and second terms) shrink to zero as $\delta \to 1$. It is going to be helpful to separate the terms of order $(1-\delta)$ from the continuation payoff in case of no state transition. Given a state $s \in S$, let $u = (u_{s'})_{s' \neq s}$ specify a vector of continuation payoffs in case of a state transition. For each such u and (possibly correlated) action profile α, we define

$$\psi^\delta(\alpha, s, u) \equiv \frac{1}{1+\delta\gamma(\alpha, s)}\left[g(\alpha, s) + \delta\sum_{s' \neq s}\gamma(s'; \alpha, s)u_{s'}\right] \quad (3)$$

and

$$\beta^\delta(\alpha, s) \equiv (1-\delta)(1+\delta\gamma(\alpha, s)) \in (0, 1).$$

[2]FLM's folk theorem (Theorem 6.2) requires either the pairwise full rank condition (Condition 6.2) or that every pure-action, Pareto-efficient profile is *pairwise identifiable* for all pairs of players. In our setting, the appropriate analog would be Pareto efficiency in terms of pseudo-instantaneous payoffs. Because pseudo-instantaneous payoffs are endogenous (they depend on continuation values after a state change), we do not focus on that second condition.

Then we can further rewrite the total payoff (2) as

$$v^\delta(\sigma, s) = \beta^\delta(\sigma_1, s)\psi^\delta(\sigma_1, s, u) + (1 - \beta^\delta(\sigma_1, s))u_s.$$

We refer to $\psi^\delta(a, s, u)$ as the *pseudo-instantaneous payoff* from playing action profile a in state s, given continuation payoffs u. The pseudo-instantaneous payoff combines two payoff effects of the same order: the instantaneous payoff g and the expected continuation payoff after transitions (which today's action affects through influencing transition rates).[3]

Notice that, typically, $\psi^\delta(\alpha, s, u)$ is not equal to $E_\alpha\psi^\delta(a, s, u)$. Nevertheless, $\psi^\delta(\alpha, s, u)$ lies in the convex hull of the set $\{\psi^\delta(a, s, u) : a \in \text{supp}(\alpha)\}$, since

$$\psi^\delta(\alpha, s, u) = \sum_a \frac{\alpha(a)\beta^\delta(a, s)}{E_\alpha\beta^\delta(a, s)}\psi^\delta(a, s, u).$$

Finally, observe that (3) is well defined for all $\delta \leq 1$. Recall that $\delta = 1$ corresponds not to infinite patience, but to the case of players who discount the future and can adjust their actions at infinitely brief intervals. The definitions above will be useful later.

3. Characterizing payoffs

In a standard repeated game, a folk theorem says that any payoff in the feasible and individually rational set of the stage game can be attained as an equilibrium payoff of the repeated game for δ close to 1. In our setting, where each state corresponds potentially to a different stage game, there is no obvious counterpart of that feasible and individually rational set. In this section, we present a particular notion, based on pseudo-instantaneous payoffs, of feasible and individually rational sets of payoffs for stochastic games. That notion has the advantage that using it leads to a folk theorem—in subsequent sections we give conditions under which these feasible and individually rational sets are exactly the limiting sets of PPE payoffs. We discuss alternative (possibly simpler and more intuitive) notions in Sections 7.1 and 7.2, and we explain why those alternatives may encompass payoffs outside the equilibrium set.

Take any collection $F = \{F(s)\}_s$ of sets $F(s) \subseteq [-M, M]^N$ for each $s \in S$. (Recall that M is the upper bound on the length of any stage-game payoff vector.) Say that collection F is *self-δ-feasible* if

$$F(s) \subseteq \text{co}\{\psi^\delta(a, s, u) : a \in A(s), u \in \times_{s' \neq s} F(s')\}.$$

Self-δ-feasibility means that each payoff in $F(s)$ can be generated as the expected payoff from some action profile in the state-s stage game followed (after a state transition) by continuation payoffs that belong to collection F. Because both sides of the above inclusion refer to collection F, the definition has a fixed point flavor.

[3]To help interpret this definition, observe that if we let τ denote the (random) number of periods until a transition to state s' from state s, given transition rate $\gamma(s'; a, s)$, then $E\delta^\tau = \delta\gamma(s'; a, s)/(1 + \delta\gamma(s'; a, s))$, the weight on $u_{s'}$ in (3).

For each player i, define the *δ-minmax payoff relative to F* for player i in state s as

$$e_i^\delta(s;F) = \inf_{\alpha_{-i}\in\times_{j\neq i}\Delta A_j(s),\, u\in\times_{s'\neq s}F(s')} \left\{ \max_{a_i\in A_i(s)} \psi_i^\delta(a_i,\alpha_{-i},s,u)\right\}.$$

Say that the collection F is *self-δ-individually rational* if for each state s, player i, and $v\in F(s)$, $v_i \geq e_i^\delta(s;F)$.

The (straightforward) proof of the following claim is in Appendix B:

Claim 1. *Given $\delta<1$, define the (convex hull of the) set of feasible payoffs in initial state s, $\hat{V}^\delta(s)$, as*

$$\hat{V}^\delta(s) \equiv \text{co}\{v^\delta(\sigma,s) : \sigma\in\Sigma^P\}.$$

The collection of feasible payoffs $\hat{V}^\delta \equiv \{\hat{V}^\delta(s)\}_s$ is self-δ-feasible, and the collection of equilibrium payoffs E^δ is both self-δ-feasible and self-δ-individually rational.

Note that self-δ-feasibility and self-δ-individually rationality together imply an ex post notion of individual rationality: each payoff above the minmax payoffs can be generated using continuation payoffs that are themselves above the minmax levels.

Next, we define stronger versions of these concepts, to be used in constructing equilibria. Let $B(v,\epsilon)$ denote the closed ball centered at v with radius $\epsilon\geq 0$. Say that collection F is *self-δ,ϵ-feasible* if for each $v\in F(s)$,

$$B(v,\epsilon)\subseteq \text{co}\{\psi^\delta(a,s,u) : a\in A(s),\, u\in\times_{s'\neq s}F(s')\}.$$

Say that collection F is *self-δ,ϵ-individually rational* if for each state s, player i, and $v\in F(s)$, $v_i\geq e_i^\delta(s;F)+\epsilon$.

Lemma 1. *For each $\delta\leq 1$ and $\epsilon\geq 0$, there exists the largest collection V_ϵ^δ such that $V_\epsilon^\delta(s)\subseteq [-M,M]^N$ for each $s\in S$, and V_ϵ^δ is self-δ,ϵ-feasible and self-δ,ϵ-individually rational. Moreover, each $V_\epsilon^\delta(s)$ is compact and convex, and $\limsup_{\delta\to 1} V_\epsilon^\delta(s)\subseteq V_\epsilon^1(s)$.*[4]

Proof. The first claim follows from the fact if F and G are any two self-δ,ϵ-feasible and self-δ,ϵ-individually rational collections, then their union $F\cup G$ is also self-δ,ϵ-feasible and self-δ,ϵ-individually rational. Since, further, the convex hull of F, $\text{co}\,F\equiv\{\text{co}\,F(s)\}_s$, and the closure of F, $\text{cl}\,F\equiv\{\text{cl}\,F(s)\}_s$, are self-$\delta,\epsilon$-feasible and self-$\delta,\epsilon$-individually rational collections, the second claim holds. For the third claim, notice that $\lim_{\delta\to 1}\psi^\delta(a,s,u)=\psi^1(a,s,u)$ for each a, s, and u, so the collection $\limsup_{\delta\to 1}V_\epsilon^\delta(s)$ is self-1, ϵ-feasible and self-1, ϵ-individually rational, and thus contained in $V_\epsilon^1(s)$. □

We will refer to elements of $V_\epsilon^\delta(s)$ as self-δ,ϵ-FIR (feasible and individually rational) payoffs in state s. We will refer to $V_0^1(s)$ as the set of self-FIR payoffs in state s. Recall

[4]We define $\limsup_{\delta\to 1}V_\epsilon^\delta(s)$ as the set of all limits $v=\lim_{n\to\infty}v_n$ of sequences $v_n\in V_\epsilon^{\delta_n}(s)$ such that $\delta_n\nearrow 1$. We also define $\liminf_{\delta\to 1}V_\epsilon^\delta(s)$ as the set of points v such that for all $\delta_n\to 1$, there exists $v_n\in V_\epsilon^{\delta_n}(s)$ such that $v=\lim_{n\to\infty}v_n$. If the sup and the inf limits are equal, we refer to them as $\lim_{\delta\to 1}V_\epsilon^\delta(s)$.

from Claim 1 that $V_0^\delta(s)$ is a subset of $\hat{V}^\delta(s)$, the set of feasible payoffs. Note that the collection of equilibrium payoffs E^δ is not necessarily self-δ, ϵ-feasible for $\epsilon > 0$.

Finally, it is also useful to define

$$V_{0+}^\delta(s) \equiv \bigcup_{\epsilon>0} V_\epsilon^\delta(s)$$

for each state s and $\delta \leq 1$.

In Appendix B, we show that the collection V_ϵ^δ can be constructed by iteratively applying a particular operator to a superset of the collection of feasible payoffs. (The argument is standard. See, for example, Mailath and Samuelson's 2006 Proposition 7.3.3, or Stokey et al.'s 1989 Theorem 17.7. Judd and Yeltekin 2011 provide techniques that might be used to implement such a computation in practice.)

Claim 1 and the last part of Lemma 1 lead to the following corollary.

COROLLARY 1. *For each stochastic game,* $\limsup_{\delta\to 1} E^\delta(s) \subseteq V_0^1(s)$.

Corollary 1 says that the set $V_0^1(s)$ is an upper bound on the limit set of equilibrium payoffs as the length of the period converges to 0. That result implies that each $V_0^1(s)$ is nonempty, because standard arguments ensure that a PPE exists.[5]

4. PARTIAL FOLK THEOREM

In this section, we show that when the Identifiability Condition holds, for sufficiently high δ, any self-δ, ϵ-FIR payoff at state s can be attained in a perfect public equilibrium from initial state s: V_ϵ^δ is contained in E^δ. The proof of that result is based on techniques in the proof of FLM's folk theorem for repeated games with imperfect public monitoring.

FLM's folk theorem requires that the set of feasible and individually rational payoffs have a nonempty interior. The role of that condition is to guarantee that after any history, it is possible to provide incentives by constructing continuation payoffs that lie in any direction from the target payoffs. Here that full dimensionality is implied by the definition of V_ϵ^δ: self-δ, ϵ-feasibility means that for any payoff $v \in V_\epsilon^\delta(s)$, every payoff within ϵ of v can be generated by some action profile and some continuation payoffs in V_ϵ^δ.

THEOREM 1. *Suppose that the Identifiability Condition holds. Then for each $\epsilon > 0$, there exists $\delta^* < 1$ such that $V_\epsilon^\delta(s) \subseteq E^\delta(s)$ for any initial state s and any $\delta \geq \delta^*$. In particular,* $V_{0+}^1(s) \subseteq \liminf_{\delta\to 1} E^\delta(s)$.

Theorem 1 demonstrates that the set $V_{0+}^1(s)$ is a lower bound on the limit set of equilibrium payoffs, as the length of the period converges to 0. By Corollary 1, the set $V_0^1(s)$ is an upper bound. We define the following property.

DEFINITION 1. Property A: For each state s, $V_0^1(s) = \mathrm{cl}\, V_{0+}^1(s)$.

[5]Note that even if the set of states is infinite, the uniform upper bound on transition rates allows us to compactify the set of strategies.

If the Identifiability Condition holds and the game satisfies Property A, then the set $V_0^1(s)$ is equal to the limit set of equilibrium payoffs.

COROLLARY 2. *Suppose that Property A and the Identifiability Condition hold. Then for each state s, $V_0^1(s) = \lim_{\delta\to 1} E^\delta(s)$.*

In general, not all games have Property A. (See Example 1 in Section 5.4.) That fact implies the possibility that Theorem 1 does not capture all equilibrium payoffs. We investigate that possibility in Section 5.

It is worth mentioning that standard repeated games satisfy Property A if and only if the set of feasible and individually rational payoffs has full dimension. (See Section 5.5.) Thus, our result extends FLM's folk theorem to stochastic games with countably many states.

4.1 *Proving Theorem 1*

FLM's proof shows that any smooth set of payoffs W strictly in the interior of the feasible and individually rational set can be attained in equilibrium. A key step is to show that any payoff on the boundary of W can be achieved as the weighted average of a stage-game payoff in the current period that lies outside W (thus the requirement that W is strictly in the interior of the feasible set) and expected continuation payoffs that lie in W. Here, we want to do something similar, with pseudo-instantaneous payoffs taking the place of the stage-game payoffs. The self-δ,ϵ-feasibility of V_ϵ^δ ensures that for each state s, there is a pseudo-instantaneous payoff outside $V_\epsilon^\delta(s)$ in each direction.

Given a state s, let $V \subseteq \mathbb{R}^N$ be a set of payoffs and let $W = \{W(s')\}_{s'\neq s}$, where each $W(s') \subseteq \mathbb{R}^N$, be a collection of payoff sets. Extending FLM and Abreu et al. (1986, 1990), we say that V is *decomposable with respect to δ and W in state s* if for each $v \in V$, there exist a mixed action profile α, payoffs $u = (u_{s'})_{s'\neq s}$ such that $u_{s'} \in W(s')$ for each $s' \neq s$, and a function $w\colon Y \to V$ such that for each player i and each action $a_i \in A_i(s)$,

$$\begin{aligned} v_i &= E_\alpha\left(\beta^\delta(a,s)\psi_i^\delta(a,s,u) + [1-\beta^\delta(a,s)]\sum_{y\in Y}\rho(a,s)[y]w_i(y)\right) \\ &\geq E_{a_i,\alpha_{-i}}\left(\beta^\delta(a,s)\psi_i^\delta(a,s,u) + [1-\beta^\delta(a,s)]\sum_{y\in Y}\rho(a,s)[y]w_i(y)\right). \end{aligned} \tag{4}$$

Expression (4) says that (i) playing profile α in state s, followed by continuation payoffs u (if the state changes) or by public-signal-contingent continuation payoffs $w(y)$ (if the state does not change) yields expected payoff v, and that (ii) given those continuation payoffs, playing α is optimal for all players. (Note that the continuation payoff after a state change does not depend on the public signal y; we discuss that point later.)

The proof of Theorem 1 relies on the following lemma (which is proven in Appendix C).

LEMMA 2. *Suppose that the Identifiability Condition holds. Then for each $\epsilon > 0$, there exists $\delta^* < 1$ such that for each state s, each $\delta \geq \delta^*$, and each $v^* \in V_{10\epsilon}^\delta(s)$, the set $B(v^*,\epsilon)$ is decomposable with respect to $V_{10\epsilon}^\delta$ and δ in state s.*

A key feature of Lemma 2 is that for any ϵ, there is a single δ^* that works in every state. That uniformity allows us to cover the case of an infinite state space. Using Lemma 2, we can complete the proof of Theorem 1.

PROOF OF THEOREM 1. For each state s, let $\bar{B}(V^\delta_{10\epsilon}(s),\epsilon) \equiv \bigcup_{v\in V^\delta_{10\epsilon}(s)} B(v,\epsilon)$ denote the closed ϵ-neighborhood of the set $V^\delta_{10\epsilon}(s)$. Lemma 2 shows that the collection of payoff sets $\{\bar{B}(V^\delta_{10\epsilon}(s),\epsilon)\}_{s\in S}$ is "self-decomposable" for high enough δ, in the sense that each $\bar{B}(V^\delta_{10\epsilon}(s),\epsilon)$ is decomposable with respect to δ and the collection $\{\bar{B}(V^\delta_{10\epsilon}(s),\epsilon)\}_{s'\in S}$. Lemma 1 shows that each $V^\delta_{10\epsilon}(s)$ is compact and convex, so each $\bar{B}(V^\delta_{10\epsilon}(s),\epsilon)$ is as well. An argument analogous to the second paragraph of FLM's proof of Lemma 4.2 establishes the result. The last claim in the theorem follows from the fact that for each $\delta \le 1$, $V^1_{0+}(s) \subseteq \lim_{\delta\to 1} V^\delta_{0+}(s)$ (since for each $0<\epsilon<\epsilon'$, $V^1_{\epsilon'}$ is self-δ, ϵ-FIR for sufficiently high δ). □

5. FULL AND APPROXIMATE FOLK THEOREMS

Theorem 1 provides a partial folk theorem for all stochastic games, in the sense that it describes a subset of the limit set of PPE payoffs. Corollary 2 shows that if Property A holds, then in fact Theorem 1 is a full folk theorem: it completely characterizes the set of PPE payoffs as the period length shrinks. However, it is possible that Property A fails, in which case, Theorem 1 may not capture all equilibrium payoffs. The purpose of this section is to investigate the possibility.

We argue that there are two different reasons why that possibility is not a significant problem. First, we show that Property A is generic among all games for which the sets $V^1_0(s)$ have nonempty interior.

Second, even if Property A does not hold for a given game, we argue that every payoff in $V^1_0(s)$ is a PPE payoff of a nearby game, i.e., a game with very similar payoffs and transition rates. Furthermore, any payoff outside $V^1_0(s)$ is *not* a PPE payoff of *any* nearby game. We conclude that if there is any uncertainty about the true description of the game, then $V^1_0(s)$ is a tight approximation of the set of PPE payoffs.

The results of this section apply to games for which the sets $V^1_0(s)$ have nonempty interior. In standard repeated games, Property A is equivalent to the condition that V^1_0 has a nonempty interior: V^1_0 has full dimension if and only if V^1_ϵ does for small enough ϵ. (See Section 5.5.) For stochastic games, that equivalence breaks down. Property A still implies that each $V^1_0(s)$ has nonempty interior (since $V^1_0(s) = \operatorname{cl} V^1_{0+}(s)$ and $V^1_0(s)$ is nonempty, $V^1_{0+}(s)$ also is nonempty and thus has nonempty interior by definition), but the converse fails. (See Example 1 in Section 5.4.) This difference is another illustration that stochastic games are qualitatively more complicated than repeated games.

We use the following notation. Given the sets of players, states, and actions available in each state and the monitoring structure, a game G can be identified as a tuple (g,γ) of stage-game payoffs and (nonnegative) transition rates. Let $\mathcal{G} = \times_{s\in S}(\mathbb{R}^{N\times \#A(s)} \times \mathbb{R}_+^{(\#S-1)\times \#A(s)})$ be the space of games. Let $\mathcal{G}_0 \subseteq \mathcal{G}$ be the class of games that satisfy the following uniform version of the "interiority" condition: there exists $\epsilon>0$ such that for each state s, $B(v,\epsilon)\subseteq V^1_0(s)$ for some $v\in\mathbb{R}^N$. The uniformity has bite only

if the state space is infinite. If there is a finite number of states, then $\mathcal{G}_0$ consists of all games such that $V_0^1(s)$ has nonempty interior for each state s.

5.1 *Equilibrium payoffs in generic games*

We assume throughout this subsection that the state space is finite, so $\mathcal{G}$ is a subset of finitely dimensional space that is convex and has nonempty interior. Hence, it can be equipped with a Lebesgue measure Λ. We say that a claim holds *for generic games* $G \in \mathcal{G}_0$ if there exists a (measurable) subset $\mathcal{G}' \subseteq \mathcal{G}_0$ such that (i) $\Lambda(\mathcal{G}_0 \setminus \mathcal{G}') = 0$ and (ii) the claim holds for each game $G \in \mathcal{G}'$.[6] (Observe that class $\mathcal{G}_0$ has nonzero Λ-measure: any game G such that the transition rate is zero for all actions in all states, and such that the stage game in each state has a feasible and individually rational payoff set with full dimension, satisfies the condition, as does an open set around G.) The proof of the next result can be found below.

THEOREM 2 (Folk theorem for generic games). *If the set of states is finite, then Property A holds for generic games $G \in \mathcal{G}_0$. In particular, if the Identifiability Condition holds, then, for each state s, $\lim_{\delta\to 1} E^\delta(s; G) = V_0^1(s; G)$ for generic games $G \in \mathcal{G}_0$.*

5.2 *Equilibrium payoffs in nearby games*

Next consider the perspective of a researcher who must specify the set of all possible equilibrium payoffs and who is not certain whether the game G is specified correctly. Specifically, she believes that the true game G' is within distance ϵ of G for some $\epsilon > 0$, where the distance between two games G, and G' is measured as a supremum norm in space $\mathcal{G}$:

$$D(G, G') = \sup_{a,s,s',i} \max\bigl(|\gamma(s'; a, s) - \gamma'(s'; a, s)|, |g_i(a, s) - g_i'(a, s)|\bigr).$$

If the researcher is cautious and does not want to exclude any potential equilibrium payoffs, then her prediction should be somewhere between

$$\bigcup_{G':D(G',G)\le\epsilon} V_{0+}^1(s; G') \quad \text{and} \quad \bigcup_{G':D(G',G)\le\epsilon} V_0^1(s; G').$$

Theorem 3 shows that when the uncertainty disappears ($\epsilon \to 0$), then so does the gap between the lower and upper bounds on the cautious researcher's prediction.

THEOREM 3 (Approximate folk theorem). *For all games $G \in \mathcal{G}_0$,*

$$\bigcap_{\epsilon>0} \bigcup_{G':D(G',G)\le\epsilon} V_{0+}^1(s; G') = V_0^1(s; G) = \bigcap_{\epsilon>0} \bigcup_{G':D(G',G)\le\epsilon} V_0^1(s; G').$$

[6] We assume that the state space is finite so as to use the Lebesgue notion of genericity.

In other words, $V_0^1(s;G)$ is equal to the union of the limit equilibrium payoffs sets for games G' that approximate game G.

The second equality in Theorem 3 follows from the fact that the self-FIR correspondence is upper hemicontinuous with respect to the game G. The proof of the first equality can be found below.

5.3 *Sketch of the proofs of Theorems 2 and 3*

The proofs of Theorem 2 and the first equality of Theorem 3 follow from the same observation. For each game $G \in \mathcal{G}_0$, we define a class of games $\{G^\eta\}$ indexed by a one-dimensional parameter η such that $G^1 = G$ and the parametrization is continuous with respect to the supremum metric. We show (Lemma 8 in Appendix D) that for any $\eta > \eta'$ in the domain of parametrization, there exists $\epsilon > 0$ such that the self-FIR correspondence in game $G^{\eta'}$ is contained in the self-1, ϵ-FIR correspondence of game G^η. It follows that

$$V_0^1(s;G) = V_0^1(s;G^1) \subseteq \lim \inf_{\eta \searrow 1} V_{0+}^1(s;G^\eta).$$

That observation finishes the proof of Theorem 3.

To complete the proof of Theorem 2, we show (in Lemma 10, Appendix D) that there are at most countably many $\eta \geq 1$ such that game G^η does *not* have Property A, so such games have zero measure in the one-dimensional space $\{G^\eta\}$. We expand on this observation to show that the subset of games without Property A has zero measure. (The rest of the proof is in Appendix D.)

5.4 *A game without Property A*

Theorem 2 established that Property A holds for generic games. Here, we present an example where Property A fails. In this "reciprocal effort" game, in state s_i, player i can exert either low effort L or high effort H; the latter is costly, but yields a benefit to the other player that exceeds the cost.

Example 1. There are two players and two states, $S = \{s_1, s_2\}$. In each state s_i, player i has two actions. The payoffs are

	H	L
H	$-1, 2$	
L	$0, 0$	

State s_1

H	L
$2, -1$	$0, 0$

State s_2

The transition rates in each state do not depend on actions and are equal to 1. ◊

It is easy to see that the vector of minmax payoffs in each state is $(0,0)$ for any discount factor. At $\delta = 1$, the sets of feasible payoffs in each state are

$$\hat{V}^1(s_1) = \text{co}\left\{(0,0), (0,1), \left(\tfrac{2}{3}, \tfrac{-1}{3}\right), \left(\tfrac{-2}{3}, \tfrac{4}{3}\right)\right\}$$

$$\hat{V}^1(s_2) = \text{co}\left\{(0,0), (1,0), \left(\tfrac{-1}{3}, \tfrac{2}{3}\right), \left(\tfrac{4}{3}, \tfrac{-2}{3}\right)\right\}.$$

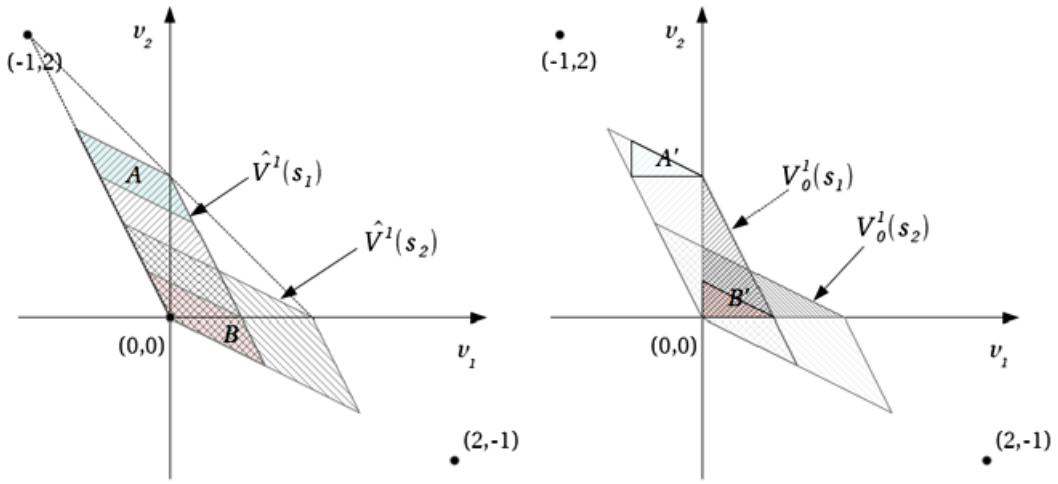

FIGURE 1. Payoff sets in Example 1.

(See the left side of Figure 1.)[7] For example, playing a_1 in state s_1 and a_2 in state s_2 yields $v^1((a_1,a_2),s_i)$ in state s_i, where $v^1((a_1,a_2),s_i) = \psi^1(a_i,s_i,v^1((a_1,a_2),s_j)) = \frac{1}{2}g(a_i,s_i) + \frac{1}{2}v^1((a_1,a_2),s_j)$. Solving yields $v^1((a_1,a_2),s_i) = \frac{2}{3}g(a_i,s_i) + \frac{1}{3}g(a_j,s_j)$.

The following proposition shows that this game violates Property A. That is, there is a discontinuity of the largest self-1, ϵ-FIR set with respect to ϵ at $\epsilon = 0$. In particular, at $\epsilon = 0$, the set has full dimension, but for any positive ϵ, the set is empty.

PROPOSITION 1. *For the game in Example 1,*

$$V_0^1(s_1) = \operatorname{co}\left\{(0,1),(0,0),\left(\tfrac{1}{2},0\right)\right\}$$
$$V_0^1(s_2) = \operatorname{co}\left\{(1,0),(0,0),\left(0,\tfrac{1}{2}\right)\right\}$$
$$V_\epsilon^1(s_1) = V_\epsilon^1(s_2) = \varnothing \quad \forall \epsilon > 0.$$

PROOF. The self-1, 0-FIR payoffs in each state s must be feasible and individually rational, and so they must belong to the intersection of the set $\hat{V}(s)$ and the positive orthant. The collection $V_0^1(s)$ as defined above is the set of all such payoffs. The collection V_0^1 so defined is trivially self-1, 0-individually rational. Moreover, V_0^1 is self-1, 0-feasible. For

[7] We define $\hat{V}^1$ as the largest self-1-feasible collection. Recall that $\hat{V}^\delta$ was defined for $\delta < 1$ in Section 3; $\hat{V}^1$ contains the limit of the $\hat{V}^\delta$'s. It is easy to verify that the collection described above is self-1-feasible. For example, $\hat{V}^1(s_1)$ is equal to the convex hull of the sets

$$A \equiv \{\psi^1(H,s_1,u) : u \in \hat{V}^1(s_2)\} = \tfrac{1}{2}\{(-1,2)\} + \tfrac{1}{2}\hat{V}^1(s_2)$$

and

$$B \equiv \{\psi^1(L,s_1,u) : u \in \hat{V}^1(s_2)\} = \tfrac{1}{2}\{(0,0)\} + \tfrac{1}{2}\hat{V}^1(s_2),$$

as shown on the left side of Figure 1.

example, note that $V_0^1(s_1)$ is contained in the convex hull of the sets

$$A' \equiv \{\psi^1(H, s_1, u) : u \in V_0^1(s_2)\} = \tfrac{1}{2}\{(-1, 2)\} + \tfrac{1}{2}V_0^1(s_2)$$

and

$$B' \equiv \{\psi^1(L, s_1, u) : u \in V_0^1(s_2)\} = \tfrac{1}{2}\{(0, 0)\} + \tfrac{1}{2}V_0^1(s_2).$$

(See the right side of Figure 1.) It follows that V_0^1 is the largest self-1, 0-FIR collection. □

The proof that each $V_\epsilon^1(s) = \varnothing$ is in Appendix E.

For the game in Example 1, the conclusion of Theorem 2 fails: many payoffs in V_0^1 are not achievable in equilibrium. In fact, for any $\delta < 1$, the only PPE is to play L after every history, yielding the minmax payoffs $(0, 0)$ in each state, as the following proposition shows. (The proof is in Appendix E.)

PROPOSITION 2. *For the game in Example 1, $E^\delta(s) = \{(0, 0)\}$ for any $\delta < 1$ and each state s, regardless of the monitoring structure.*

The nature of the nongenericity that allows Property A to fail in Example 1 is that at $\delta = 1$, the strategy profile of playing H in both states after every history gives player i a payoff, starting from state s_i, exactly equal to his minmax payoff, 0. The reason is that the expected discounted fraction of time spent in the current state is $(1 + \delta)/(1 + 2\delta) = \frac{2}{3}$. For any $\delta < 1$, though, that profile gives player i a negative payoff, $[(1+\delta)/(1+2\delta)](-1)+[\delta/(1+2\delta)]2 < 0$, because he exerts effort up front and starts receiving benefits later. If, for instance, we increase the payoff to player j when action H is played in state s_i to any value strictly greater than 2, then Property A will hold, and both $V_0^1(s)$ and $\operatorname{cl} V_{0+}^1(s)$ will be equal to the intersection of the feasible payoff set $\hat{V}^1(s)$ (computed at the new parameter value) with the set of individually rational payoffs. (See Example 6 in Section 7.3.) If we decrease that payoff, on the other hand, then the set $V_0^1(s)$ collapses to the singleton $\{(0, 0)\}$.

5.5 *Standard repeated games*

As mentioned above, for a standard repeated game, Property A is trivially satisfied if the set of feasible and individually rational payoffs (in the stage game) has nonempty interior.

REMARK 1. For each game $G = (g, \gamma) \in \mathcal{G}_0$ (i.e., G that satisfies the nonempty interior condition), if $\gamma(a, s) = 0$ for each state s and each action profile $a \in A(s)$, then G has Property A.

The result follows from the fact that for such games, any ϵ-interior of $V_0^1(s)$ is self-1, ϵ-FIR. Moreover, because $G \in \mathcal{G}_0$, $V_0^1(s)$ is the limit of such ϵ-interiors.

6. Example: Oligopoly in an evolving world

Here, we describe a class of oligopoly games and we show how to compute the self-FIR collection. In this class, the computations are especially easy. In particular, the limiting (as $\delta \to 1$) set of equilibrium payoffs in state s can be found as the solution to a simple dynamic fixed point problem.

Example 2. There are N firms. In each period, firm i chooses an action $a_i \in A_i$ (quantities, prices, investments, the number of advertisements, etc.); each A_i includes an *inactive* action 0_i. The "inactive" action corresponds, for instance, to quantity zero in Cournot competition or to a very high price in Bertrand competition.

The state is a vector $s = (e, s_1, \ldots, s_N)$, where the state of the economy e represents aggregate market conditions, and the firm-specific state s_i reflects the level of technology, the patent pool, costs of production, etc. The state of the economy evolves independently of the actions and private states of the firms, and a firm's private state evolves independently of the actions and private states of other firms. (Simultaneous transitions are ruled out.) That is, the transition rate from state $s = (e, s_1, \ldots, s_N)$ to state $s' = (e', s'_1, \ldots, s'_N)$, given action profile a, is equal to

$$\gamma(s', a, s) = \begin{cases} \gamma_E(e', e) & \text{if } e' \neq e \text{ and } s'_i = s_i \text{ for all } i \\ \gamma_i(s'_i, a_i, e, s_i) & \text{if } s'_i \neq s_i,\ e' = e, \text{ and } s'_j = s_j \text{ for all } i \neq j \\ 0 & \text{otherwise.} \end{cases}$$

A firm's instantaneous payoff $g_i(a, e, s_i)$ depends on the actions of all firms, the aggregate conditions, and the firm-specific state. The payoff from the inactive action, $g_i((0_i, a_{-i}), e, s_i)$, is 0 for all states and all actions by the other firms, and 0 is the (stage-game) minmax payoff. In any state, the payoffs of all players $j \neq i$ (weakly) increase if player i is inactive: $g_j((a_i, a_{-i}), e, s_j) \leq g_j((0_i, a_{-i}), e, s_j)$. Finally, there exists $\eta > 0$ such that in any state, each player can get a payoff of at least η if the other players are inactive: $\max_{a_i} g_i((a_i, 0_{-i}), e, s_i) \geq \eta$. ◊

The key feature that simplifies the analysis of this class of examples is that any player i can maximize the (dynamic) payoffs of any other players by choosing to be inactive, and doing so gives player i exactly his minimum total payoff given that i best-responds to the strategies of other players; in Section 7.1, we denote that value as the *Nash minmax*.

It is straightforward to see that this Nash minmax for each player in each state is 0. We will describe the set of all feasible payoffs in the limit as $\delta \to 1$. For each unit vector λ, we characterize the largest possible reach of hyperplane λ that can be attained among all feasible payoffs. Specifically, let $(c^\lambda(s))_{s \in S}$ be the vector of solutions to the system of equations

$$\begin{aligned} c(s) &= \max_{\alpha \in \Delta A} \frac{\lambda \cdot g(\alpha, s) + \sum_{s' \neq s} \gamma(s', \alpha, s) c(s')}{1 + \gamma(\alpha, s)} \\ &= \max_{\alpha \in \Delta A,\, u : \lambda \cdot u_{s'} \leq c(s') \text{ for each } s' \neq s} \lambda \cdot \psi^1(\alpha, s, u). \end{aligned}$$

Because the transition rate $\gamma(a,s)$ is bounded, one can use the contraction mapping theorem to show that the solution is unique.

Note that $c^\lambda(s)$ is a tight upper bound on the value in a dynamic problem, in which players receive an instantaneous payoff $\lambda \cdot g(\alpha, s)$ as long as they stay in state s and then receive $c^\lambda(s')$ following a transition to state s'. It is easy to see that if v is a feasible payoff (in the limit as $\delta \to 1$) in state s, then for each unit vector λ, $\lambda \cdot v \leq c^\lambda(s)$. From self-generation arguments, the opposite claim holds as well: the (limit) set of feasible payoffs in state s is equal to the set of all payoffs v such that $\lambda \cdot v \leq c^\lambda(s)$ for all unit vectors λ.

The next result shows that in the game of Example 2, the set of feasible payoffs that lie above Nash minmaxes is self-feasible and individually rational, and that it is equal to the (limit) set of equilibrium payoffs.

PROPOSITION 3. *In the game in Example 2, Property A holds, and for each state s,*

$$V_0^1(s) = \{v : v_i \geq 0 \text{ for each } i \text{ and } \lambda \cdot v \leq c^\lambda(s) \text{ for each unit vector } \lambda\}.$$

The proof is in Appendix F.

7. Comments and other examples

The focus of this paper has been on characterizing the set of PPE payoffs under the Identifiability Condition. An important component of that effort was defining the appropriate notion of feasible and individually rational payoffs in Section 3. The goal of this section is to examine how important our assumptions and definitions are for the results. In Sections 7.1 and 7.2, we discuss two alternative, perhaps more natural, definitions of individually rational and feasible payoffs. We explain why these definitions do not lead to a folk theorem type of result; that is, why payoffs that satisfy these alternative definitions may not be achievable in perfect public equilibrium. Similarly, in Section 7.3, we show that a stochastic game may have equilibrium payoffs that cannot be achieved by any Markov strategy that gives all players at least their minmax payoffs after every history. In Section 7.4, we show that we cannot relax the identifiability conditions in a way analogous to HSTV without sacrificing some PPE payoffs. Finally, in Sections 7.5 and 7.6, we discuss briefly how our results change if, instead of looking at PPEs, we consider other solution concepts: Nash and sequential equilibrium.

7.1 *The "Nash" minmax*

A player's minmax payoff in a given state typically depends on the set of continuation payoffs available in other states. In Sections 4 and 5.1, we use the minmax payoff given that those continuation payoffs must lie in the equilibrium set. The more obvious parallel to standard repeated games might be to consider all feasible continuation payoffs. We consider the consequences of using that definition here and in the following section. (See also the discussion in Section 7.5.)

Recall that $\hat{V}^\delta(s)$ is the set of payoffs attainable through some strategy profile starting from state s:

$$\hat{V}^\delta(s) \equiv \text{co}\{v^\delta(\sigma, s) : \sigma \in \Sigma^P\}.$$

Define the *Nash minmax* $e_i^{\text{Nash},\delta}(s)$ in state s as the minimum payoff obtained by player i given that i best-responds to the strategies of other players:

$$e_i^{\text{Nash},\delta}(s) \equiv \min_{\sigma_{-i}\in\times_{j\neq i}\Sigma_j^P} \max_{\sigma_i} v_i^\delta(\sigma_i, \sigma_{-i}, s).$$

It is easy to check that the Nash minmax satisfies the recursive equation

$$e_i^{\text{Nash},\delta}(s) = \min_{\alpha_{-i}\in\times_{j\neq i}\Delta A_j(s)} \Big\{ \max_{a_i\in A_i(s)} \psi_i^\delta\big((a_i,\alpha_{-i}), s, (e^{\text{Nash},\delta}(s'))_{s'\neq s}\big)\Big\}, \tag{5}$$

where $e^{\text{Nash},\delta}(s')$ is the vector of Nash minmax payoffs in state s'. In other words, the Nash minmax is the δ-minmax payoff relative to the collection of feasible payoffs $\hat{V}^\delta$. Since $V_0^\delta(s)\subseteq\hat{V}^\delta(s)$, $e_i^{\text{Nash},\delta}(s)$ is weakly lower than $e_i^\delta(s;V_0^\delta)$.[8] Finally, define the set of feasible and Nash-individually rational payoffs, i.e., payoffs that can be obtained in some strategy profile and that are higher than the minmaxes, as

$$V^{\text{Nash},\delta}(s) = \{v : v\in\hat{V}^\delta(s) \text{ and } v_i \geq e_i^{\text{Nash},\delta}(s)\ \forall i\}.$$

The set of PPE payoffs in the stochastic game may be strictly smaller than $V^{\text{Nash},\delta}(s)$. One of the reasons is that some of the payoffs $v\in V^{\text{Nash},\delta}(s)$ might be obtainable only in strategy profiles that lead to continuation payoffs that are not Nash-individually rational. In that case, v cannot be a payoff in any PPE. The following example illustrates the point. (We set the transition rate equal to $1/\delta$ merely to simplify algebra; setting $\gamma=1$ would yield similar results.)

EXAMPLE 3. There are two players and two states, $S=\{s_1,s_2\}$. In each state s_i, player i has two actions. The payoffs are

H	3, 1
L	0, 0

State s_1

H	L
3, −2	0, 0

State s_2

The transition rates in each state do not depend on actions and are equal to $1/\delta$: $\gamma(s_j;a,s_i)=1/\delta$ for all $i,j\in\{1,2\}$, $i\neq j$, and all $a\in\{H,L\}$. ◊

The Nash minmax value for player 2 is equal to 0 in both states. The Nash minmax value for player 1, obtained by his playing H in state s_1 and player 2's playing L in state s_2, can be computed using the recursive formula (5). It depends on the state: $e_1^{\text{Nash},\delta}(s_1)=2$ and $e_1^{\text{Nash},\delta}(s_2)=1$.

Payoff vector $v=(3,0)$ is feasible in state s_1, as it can be obtained from a profile in which H is played in each period (and only by that profile). (Note that solving $v^\delta(HH,s_1)=\psi^\delta(H,s_1,v^\delta(HH,s_2))$ and $v^\delta(HH,s_2)=\psi^\delta(H,s_2,v^\delta(HH,s_1))$ yields $v^\delta(HH,s_1)=(3,0)$ and $v^\delta(HH,s_2)=(3,-1)$.) The payoff v is also Nash-individually rational. Nevertheless, it cannot be an equilibrium payoff. The reason is that playing H in

[8] Recall that the simplifying feature of the class of games in Example 2 was that $e_i^{\text{Nash},\delta}(s)=e_i^\delta(s;V_0^\delta)$ for each player in each state.

each period, starting from state s_2, yields payoff $(3,-1)$: that payoff cannot result from any equilibrium, since player 2 can get a payoff of at least 0 by always playing L.[9]

7.2 *"Ex post Nash" individual rationality*

To eliminate payoffs like the one in Example 3, one can require an *ex post* notion of individual rationality;[10] that is, each payoff v must be generated using strategies with continuation payoffs above Nash minmaxes:

$$V^{\text{ex post Nash},\delta}(s) \equiv \text{co}\{v^\delta(\sigma,s) : \sigma \in \Sigma^P \text{ and } v_i^\delta(\sigma,h) \geq e_i^{\text{Nash},\delta}(s_t)\ \forall i \text{ and } \forall h \in H_t\}.$$

Here we show that even a strategy that delivers each player at least his Nash minmax payoff after every history may yield a payoff that cannot be achieved in equilibrium. That is, the set $V^{\text{ex post Nash},\delta}(s)$ may be strictly larger than the set of PPE payoffs $E^\delta(s)$.

EXAMPLE 4. There are two players and two states, $S=\{s_1,s_2\}$. State s_1 is the initial state and state s_2 is absorbing. The payoffs are

State s_1

	$A2$	$B2$
$A1$	$1,1$	$1,0$
$B1$	$0,1$	$2,2$

State s_2

	$C2$	$D2$
$C1$	$4,0$	$-1,-2$
$D1$	$9,2$	$0,-2$

In state s_1, the transition rate is $1/\delta$ for every action profile. In state s_2, the transition rate for every profile is 0. That is, $\gamma(s_2;a,s_1)=1/\delta$ for all a, and $\gamma(s_1;a,s_2)=0$ for all a. ◊

In state s_2, the stochastic game is reduced to a standard repeated game, and it is easy to see that player 1's minmax payoff is 0. We can then use that value to calculate his Nash minmax payoff in state s_1. Since

$$\begin{aligned}
\psi_1^\delta((A1,A2),s_1,0) &= \psi_1^\delta((A1,B2),s_1,0) = \tfrac{1}{2}\cdot 1+\tfrac{1}{2}\cdot 0=\tfrac{1}{2}\\
\psi_1^\delta((B1,A2),s_1,0) &= \tfrac{1}{2}\cdot 0+\tfrac{1}{2}\cdot 0=0\\
\psi_1^\delta((B1,B2),s_1,0) &= \tfrac{1}{2}\cdot 2+\tfrac{1}{2}\cdot 0=1,
\end{aligned}$$

that minmax value is $\frac{1}{2}$. Similarly, player 2's minmax payoff in state s_2 is also 0, and her Nash minmax value in state s_1 is also $\frac{1}{2}$.

Thus, the Markov strategy profile where player 1 plays $B1$ in state s_1 and randomizes with probability $\frac{1}{10}$ on $D1$ in state s_2, and player 2 plays $A2$ in state s_1 and $C2$ in state s_2, yielding payoffs $(\frac{9}{4},\frac{3}{5})$ in state s_1 and $(\frac{9}{2},\frac{1}{5})$ in state s_2, gives both players more than their Nash minmax values after every history. There is no PPE, however, that gives a payoff close to $(\frac{9}{4},\frac{3}{5})$ in state s_1. In fact, any PPE must give player 1 a payoff of at least

[9]Similarly, for small η, the *strictly* Nash-individually rational payoff $(3-\eta,\frac{2}{3}\eta)$ can be obtained in state s_1 by playing H with probability 1 in state s_1 and with probability $1-\eta$ in state s_2, but it is not an equilibrium payoff.

[10]The idea of ex post individual rationality appears in Dutta (1995).

2.5 in state s_1. The reason is that in state s_2, any feasible payoff that gives player 2 at least her minmax value of 0 gives player 1 a payoff of at least 4. Thus, any continuation equilibrium once state s_2 is reached must give player 1 a payoff of at least 4. By playing action $A1$ in state s_1 until a transition occurs, player 1 can assure himself an expected payoff of at least $\frac{1}{2}\cdot 1+\frac{1}{2}\cdot 4=\frac{5}{2}$ against any equilibrium strategy of player 2.

In a standard repeated game, the fact that player 1 must get strictly more than his minmax payoff in any equilibrium (because, as in the game in state s_2, individual rationality for player 2 requires a high payoff for player 1) does not affect the ability of player 2 to minmax player 1 for a *single period*. In a stochastic game, however, the relevant single-period minmax is the one that corresponds to pseudo-instantaneous payoffs (an action that gives a low payoff today is not an effective threat if it is likely to lead to another state with a high continuation payoff), and pseudo-instantaneous payoffs depend on the available continuation values in other states. Dutta's (1995) Example 2 provides a similar intuition for the case where transition probabilities are independent of δ.[11]

Note that the result does not depend on the fact that state s_2 is absorbing. Adding a small positive transition rate out of state s_2 would not qualitatively change the result.

7.3 *Markov strategies are not enough*

Another interesting property of stochastic games is that the set of payoffs generated by stationary Markov strategies that are individually rational in each state may be strictly smaller than the set of equilibrium payoffs. Given the widespread use of Markov strategies in the applied literature on stochastic games, it is important to note that if we restrict attention to such strategies, then we may not be able to describe all possible equilibrium payoffs (leaving aside the question of whether or not those payoffs can be achieved in a Markov perfect equilibrium). We illustrate this possibility in Appendix G, where we modify the reciprocal effort game in Example 1 slightly, so that Property A is satisfied.

7.4 *Incentives after state transitions*

In the strategies constructed in the proof of Theorem 1, continuation payoffs after a state transition are independent of the public signal about actions. In principle, we could, like

[11]The details of that example are not quite correct, however. In fact, in state σ, any feasible payoff vector x that gives both players more than 0 can be achieved in a subgame perfect equilibrium (SPE) if players are patient. First, note that in the absorbing state, s, patient players can get any payoff in $\text{co}\{(0,2),(3,3)\}$ in equilibrium. The payoff x can be written as $x=\alpha w+(1-\alpha)v$ for some $\alpha\in[0,1]$, where $v\in\text{co}\{(0,2),(3,3)\}$ and $w\in\text{co}\{(0,-1),(3,0)\}$. (That is, w is the result of player 1 playing a_1 and player 2 randomizing between her actions.) Define T as $\delta^T\equiv\alpha$. Then there is a subgame perfect strategy profile that achieves (approximately) x: play the profile that yields w for the first T periods, and then switch to state s (by playing (a_2,b_1)) and play the SPE that yields v. After any unilateral deviation by player 2 during the first T periods, restart. After any unilateral deviation by player 1 during the first T periods, switch to state s (if the deviation did not already result in a switch) and play a SPE that gives player 1 a payoff below x_1. (To achieve x exactly, the post-transition continuation payoff v would need to be adjusted slightly to compensate for the one period of payoff $(1,0)$ when the players switch to state s, as well as for the fact that T may not be an integer.) Dutta's (1995) Example 1 has a similar problem.

HSTV, strengthen players' incentives to choose the specified actions by allowing post-transition payoffs to depend on the public signal as well. Such a strengthening is not necessary to get a folk theorem as long as the Identifiability Condition holds. HSTV, though, are able to provide a weaker sufficient condition on the monitoring structure by taking advantage of the property that transition rates vary with the action profile and using the informativeness of state transitions about actions. In our setting, however, transitions do not occur frequently enough to allow such a weakening, in general. Because the per-period probability of a transition is proportional to $1-\delta$, the expected value of changing the infinite stream of future payoffs only if a transition occurs in the current period is on the same scale as the instantaneous payoff. Post-transition incentives, therefore, are only as effective as punishing a deviation for a single period would be in a standard repeated game. The following example, which satisfies HSTV's monitoring conditions (Assumptions F1 and F2) but not our Identifiability Condition, and in which the folk theorem fails, illustrates.

EXAMPLE 5. The stage game is a symmetric two-player prisoners' dilemma with payoffs that are independent of the state,

	C	D
C	$1,1$	$-L,1+G$
D	$1+G,-L$	$0,0$

where $L>G>0$. The set of public signals Y is a singleton: the public signal is uninformative. There are eight states: $S=\{CC,DD,CD,DC\}\times\{-1,1\}$. The first component of a state reflects the action played during the period of the most recent state transition; the ±1 component allows the players to identify when a state change occurs even if the action played during the transition is the same as in the previous transition. That is, state transitions are given by

$$\gamma((a',j');\hat{a},(a,j))=\begin{cases}\gamma & \text{if } a'=\hat{a} \text{ and } j'=-j\\ 0 & \text{otherwise,}\end{cases}$$

where $\gamma>0$. ◊

We claim that when G is large, then there is no PPE for any value of δ that supports the efficient symmetric payoff $(1,1)$. To see why, note that in such an equilibrium, (C,C) must be played in every period. A player would prefer to deviate to D unless the following condition holds:

$$1\geq(1-\delta)(1+G)+\delta\big[(1-\delta)\gamma\cdot0+[1-(1-\delta)\gamma]\cdot1\big]. \tag{6}$$

Expression (6) says that the payoff of 1 from cooperating must be at least as high as the expected payoff from deviating for one period. The deviation can be punished only if it is publicly revealed—that is, only if the state changes—and the punishment payoff can be no worse than the minmax payoff 0. But (6) is equivalent to the condition that $\delta\gamma\geq G$. Thus, if $G>\gamma$, then payoffs $(1,1)$ cannot be sustained.

In fact, a similar argument rules out *any* payoffs other than $(0,0)$. The one-period gain from playing D rather than C outweighs the greatest possible loss in continuation value $(1-0=1)$ times the probability that the deviation is publicly detected $((1-\delta)\gamma)$.

7.5 *Nash equilibrium*

Let $E^{\text{Nash},\delta}$ denote the set of Nash equilibria in public strategies. In standard repeated games, the folk theorem tells us that for high values of δ, the set of PPE payoffs E^δ is approximately equal to $E^{\text{Nash},\delta}$. In stochastic games, that equivalence need not hold: $E^{\text{Nash},\delta}$ may be strictly larger. The reason is that a different notion of "individually rational" applies to the two different solution concepts. Recall Example 4, and that the Markov strategy profile where player 1 plays $B1$ in state s_1 and randomizes with probability $\frac{1}{10}$ on $D1$ in state s_2, and player 2 plays $A2$ in state s_1 and $C2$ in state s_2, yielding payoffs $(\frac{9}{4},\frac{3}{5})$ in state s_1 and $(\frac{9}{2},\frac{1}{5})$ in state s_2, gives both players more than their Nash minmax values after every history. In Section 7.2, we showed that there is no PPE that gives a payoff close to $(\frac{9}{4},\frac{3}{5})$ in state s_1. However, with perfect monitoring, that payoff can be attained in a Nash equilibrium, using the strategy above, supported by the threat of switching to the Nash minmax strategy for player i if player i deviates. Note, though, that a continuation strategy in state s_2 that gave a player less than his or her state s_2 Nash minmax could *not* be part of a Nash equilibrium, because state s_2 is reached with probability 1 under any strategy.

As that example suggests, the "right" set of feasible and individually rational payoffs for games with perfect monitoring when the solution concept is Nash equilibrium is $V^{\text{ex post Nash},\delta}$, the collection of payoffs that can be generated using strategies with continuation payoffs above Nash minmaxes (defined in Section 7.2). Clearly, a strategy that gives some player a continuation payoff below his Nash minmax after some history on the path of play cannot be a Nash equilibrium: that player has a profitable deviation. So consider any strategy σ that gives continuation payoffs at least $\epsilon > 0$ above Nash minmaxes after every history on the path of play, and let v be the associated payoff. We claim that the payoff from that strategy, $v^\delta(\sigma, s_0)$, lies in $V^{\text{ex post Nash},\delta}(s_0)$, and that there exists a Nash equilibrium strategy σ^* that attains payoff $v^\delta(\sigma, s_0)$ if δ is high enough.

The intuition is as follows. Suppose that σ is a pure strategy. (If not, then for high δ, there exists a pure strategy that yields a payoff arbitrarily close to $v^\delta(\sigma, s_0)$.) We use σ to construct two other strategies: $\hat{\sigma}$ is the same as σ on the path, it specifies player i's Nash minmax strategy after any unilateral deviation by player i, and it specifies some continuation that gives at least Nash minmax payoffs after multilateral deviations. The strategy $\tilde{\sigma}$ is the same as σ on the path and it specifies some continuation that gives at least Nash minmax payoffs at every history off the path. Strategy $\hat{\sigma}$ gives the same payoff as σ and it is a Nash equilibrium if players are patient. Strategy $\tilde{\sigma}$ features continuation payoffs above Nash minmaxes after every history and it also gives the same payoff as σ, so we conclude that $v^\delta(\sigma, s_0) \in V^{\text{ex post Nash},\delta}(s_0)$. Thus, the set of Nash equilibrium payoffs is approximately equal to $V^{\text{ex post Nash},\delta}(s_0)$ for high δ.

7.6 *Sequential equilibrium*

Suppose that we use sequential equilibrium rather than PPE as our solution concept. (That is, we drop the requirement that players use public strategies.) The key difference that arises, as in standard repeated games, is that in games with imperfect monitoring, two or more players may be able to use information in their private histories as a correlation device and, thus, hold another player below his minmax payoff. Potentially, then, any feasible vector of payoffs such that each player gets at least his *correlated* minmax payoff may be sustainable in sequential equilibrium. (See Fudenberg and Tirole's 1991 Example 5.10, and see Gossner and Hörner 2010 for a discussion.) Thus, allowing private correlation (either exogenously through a private randomization device or endogenously through private strategies) can expand the set of equilibrium payoffs, and so the limiting set (as players become patient) of sequential equilibrium payoffs may be strictly greater than the limiting set of PPE payoffs. In a standard repeated game, however, if the public monitoring structure is such that a folk theorem holds (that is, if every feasible payoff that gives each player at least his uncorrelated minmax can be achieved in a perfect public equilibrium), then that expansion can only be "downward": any sequential equilibrium payoff that is not a PPE payoff must give at least one player a payoff below his uncorrelated minmax. (Otherwise the payoff would be in the set of PPE payoffs.)

In stochastic games, on the other hand, allowing private correlation may yield an equilibrium that Pareto dominates any PPE. We demonstrate with an example in Appendix H. The construction of the example uses the fact that a standard repeated game may have sequential equilibrium payoffs that are *not* Pareto dominated by any PPE payoff.

8. Summary and discussion

This paper examines properties of the set of PPE payoffs for stochastic games with imperfect public monitoring in the limit as the time between periods shrinks to zero, holding the time rate of discounting and the time rate of state transitions fixed. There are important qualitative differences between the results in this model and the results in the models studied by Dutta (1995), FY, and HSTV. In particular, in our environment the limiting set of PPE payoffs typically varies with the initial state.

Roughly speaking, we do two things in the paper. First, we characterize the set of payoffs that could potentially be achieved in equilibrium: that is, the analog for stochastic games of a repeated game's feasible and individually rational set of stage-game payoffs. Second, we show that all such payoffs can in fact be attained in equilibrium if monitoring is sufficiently informative and the length of the period is sufficiently small. We provide upper and lower bounds on the equilibrium set, and we show that if the number of states is finite, then those bounds coincide for generic games that satisfy a full-dimensionality condition. Additionally, we define pseudo-instantaneous payoffs, and demonstrate that they are a useful tool.

We note that the results of this paper can be easily generalized to sequences of stochastic games in which the payoff function g_δ and the transition function γ_δ are parametrized with discount factor $\delta < 1$, and $g_\delta \to g_1 = g$ and $\gamma_\delta \to \gamma_1 = \gamma$ for $\delta \to 1$

and some functions g and γ. We can redefine pseudo-instantaneous payoffs ψ^δ and collections V_ϵ^δ of self-(δ,ϵ)-FIR payoffs using the functions g_δ and γ_δ in place of g and γ. The statements of Theorems 1–3 and their proofs apply with straightforward changes. In particular, the definition of Property A is not changed.

The notion of self-feasible and self-individually-rational collections of payoff sets may potentially be useful in deriving equilibrium payoff sets in Dutta (1995), FY, and HSTV's environment for the case where irreducibility is not satisfied. In that setting, the pseudo-instantaneous payoff would be redefined to reflect the fact that transition probabilities are no longer proportional to $1-\delta$:

$$\psi^\delta(a,s,u) \equiv \frac{1}{1-\delta+\delta\gamma(a,s)}\left[(1-\delta)g(a,s)+\delta\sum_{s'\neq s}\gamma(s';a,s)u_{s'}\right].$$

Because we define the equilibrium set that we derive (V_0^1) implicitly as the largest fixed point of a given correspondence, the question of how to derive its properties from the primitives of the stage-game payoffs and transition probabilities for a given stochastic game is open. The approaches of Judd and Yeltekin (2011) and of Kitti (2013) may be useful. Other interesting topics for further research include introducing imperfect private monitoring of state transitions, and tying the precision of monitoring to the period length (as in Abreu et al. 1991, Fudenberg and Levine 2007, 2009, and Hellwig and Schmidt 2002, among others). In Pęski and Wiseman (2014), we study a variation on our model with a continuous state space, where as the period length decreases, the probability of state changes stays fixed, but the magnitude of the change shrinks.

Appendix A: Identifiability

For each state s, player i, and (mixed) action profile α, let $\Pi_i(\alpha,s)$ be the $m_i(s)\times m$ matrix whose rows correspond to the probability distribution over public signals induced by each of player i's actions, given s and α_{-i}: $\Pi_i(\alpha,s)\equiv\rho((\cdot,\alpha_{-i}),s)$. Similarly, for each state s and action profile α, let $\Pi_{ij}(\alpha,s)$ be the $(m_i(s)+m_j(s))\times m$ matrix whose first $m_i(s)$ rows are $\Pi_i(\alpha,s)$ and whose last $m_j(s)$ rows are $\Pi_j(\alpha,s)$.

Action profile α has *individual full rank* in state s if $\Pi_i(\alpha,s)$ has rank $m_i(s)$ for each player i. Action profile α has *pairwise full rank* for players i and j in state s if $\Pi_{ij}(\alpha,s)$ has rank $m_i(s)+m_j(s)-1$. FLM's identifiability requirements are based on those full-rank conditions. To ensure that the rank conditions hold uniformly over the possibly infinite number of states, we require some more notation.

Let $\mathcal{M}_{kl}$ be the set of $k\times l$ matrices and let $\mathcal{M}_l$ be the set of square l-matrices. Given $j\leq k,l$ and matrices $A\in\mathcal{M}_{kl}$ and $B\in\mathcal{M}_j$, we write $B\subseteq A$ if matrix B can be obtained from A by crossing out $k-j$ rows and $l-j$ columns. Let

$$d_j(A)=\max_{\{B\in\mathcal{M}_j:B\subseteq A\}}|\det B|.$$

Thus, $d_j(A)>0$ if and only if the rank of matrix A is not smaller than j. Individual full rank for action α in state s is equivalent to the condition $d_{m_i(s)}(\Pi_i(\alpha,s))>0$ for

each player i, and pairwise full rank for players i and j is equivalent to the condition $d_{m_i(s)+m_j(s)-1}(\Pi_{ij}(\alpha,s))>0$.

Given scalar $d>0$, say that action profile α has *individual d-rank* in state s if $d_{m_i(s)}(\Pi_i(\alpha,s))\geq d$ for each player i. Similarly, say that α has *pairwise d-rank* for players i and j in state s if $d_{m_i(s)+m_j(s)-1}(\Pi_{ij}(\alpha,s))\geq d$.

With those definitions, we can state the identifiability condition on the monitoring structure.

Definition 2. *Identifiability Condition.* There exists $d>0$ such that for each state s, the following statements hold:

(i) Every pure action profile has individual d-rank in state s.

(ii) For each pair of players i and j, there exists a profile $\alpha(s)$ that has pairwise d-rank for i and j in state s.

If the number of states is finite, then our identifiability condition requires only that FLM's Conditions 6.2 and 6.3 hold in each state.

Appendix B: Proof of Claim 1 and an iterative procedure for constructing V^δ_ϵ

Proof of Claim 1. Pick any nonzero vector $\lambda\in\mathbb{R}^N$. For each state s, let $v^\lambda(s)\in \arg\max_{v\in\hat V^\delta(s)}\lambda\cdot v$ and let $\sigma^\lambda(s)\in\Sigma^P$ be a strategy that yields payoff $v^\lambda(s)$ from initial state s. Strategy $\sigma^\lambda(s)$ induces mappings $w_{s'}:Y\to\hat V^\delta(s')$ that specify, for each state s' and each public signal y, the continuation payoff if public signal y is observed in period 1 and the state in period 2 is s'. Strategy $\sigma^\lambda(s)$ also specifies the (mixed) action profile α^λ to be played in the first period. Define $v^\lambda\in\times_{s'\neq s}\hat V^\delta(s')$ as $\{v^\lambda(s')\}_{s'}$. Then

$$\begin{aligned}\lambda\cdot v^\lambda(s)=\lambda\cdot E_{\alpha^\lambda}(1-\delta)g(a,s)&\\ +\lambda\cdot E_{\alpha^\lambda}\Bigg\{\delta\sum_{y\in Y}\rho(a,s)[y]&\\ \times\Bigg[\sum_{s'\neq s}(1-\delta)\gamma(s';a,s)w_{s'}(y)+\delta[1-(1-\delta)\gamma(a,s)]w_s(y)\Bigg]\Bigg\}&\\ \leq E_{\alpha^\lambda}\beta^\delta(a,s)\{\lambda\cdot\psi^\delta(a,s,v^\lambda)\}+[1-E_{\alpha^\lambda}\beta^\delta(a,s)]\lambda\cdot v^\lambda(s)&\\ \leq E_{\alpha^\lambda}\beta^\delta(a,s)\Big\{\max_{a\in A(s)}\lambda\cdot\psi^\delta(a,s,v^\lambda)\Big\}+[1-E_{\alpha^\lambda}\beta^\delta(a,s)]\lambda\cdot v^\lambda(s),&\end{aligned}$$

where the first inequality follows from the definition of $v^\lambda(\cdot)$ and the second inequality follows from the fact that $\psi^\delta(\alpha^\lambda,s,v^\lambda)$ lies in the convex hull of the set $\{\psi^\delta(a,s,v^\lambda):a\in\mathrm{supp}(\alpha^\lambda)\}$. Thus, $\lambda\cdot v^\lambda(s)\leq\max_{a\in A(s)}\lambda\cdot\psi^\delta(a,s,v^\lambda)$ for all λ. We conclude that

$$\hat V^\delta(s)\subseteq\mathrm{co}\{\psi^\delta(a,s,u):a\in A(s),u\in\times_{s'\neq s}\hat V^\delta(s')\};$$

that is, $\hat{V}^\delta$ is self-δ-feasible.

An analogous argument establishes that E^δ is self-δ-feasible. To see that E^δ is self-δ-individually rational, note that because a PPE strategy must specify, after any deviation, continuation payoffs that are themselves PPE payoffs, player i would have a profitable deviation from a strategy that did not give him a payoff of at least $e_i^\delta(s; E^\delta)$ starting from state s. □

An iterative procedure for constructing V_ϵ^δ

Let $\mathscr{F}$ denote the set of collections F, where $F = \{F(s)\}_s$ and $F(s) \subseteq [-M, M]^N$ for each $s \in S$. Given scalar $\epsilon \geq 0$ and set $X \subseteq \mathbb{R}^N$, let $\text{int}(X, \epsilon) \equiv \{x \in X : B(x, \epsilon) \subseteq X\}$ denote the ϵ-interior of X. Then for each $\delta \leq 1$ and $\epsilon \geq 0$, define the transformation $T^{\delta,\epsilon} : \mathscr{F} \to \mathscr{F}$ as follows.

DEFINITION 3. We have $T^{\delta,\epsilon}(F) \equiv \{T_s^{\delta,\epsilon}(F)\}_s$, where

$$T_s^{\delta,\epsilon}(F) \equiv \left\{v \in \text{int}(\hat{T}_s^\delta(F), \epsilon) : v_i \geq e_i^\delta(s; F) + \epsilon \ \forall i\right\}$$

and

$$\hat{T}_s^\delta(F) \equiv \text{co}\{\psi^\delta(a, s, u) : a \in A(s), u \in \times_{s' \neq s} F(s')\}.$$

That is, $T^{\delta,\epsilon}(F)$ is the largest collection that is δ, ϵ-feasible and δ, ϵ-individually rational relative to F.

Let $\bar{F} \equiv ([-M, M]^N)^S$ and let $\bar{F}_\infty^{\delta,\epsilon} \equiv \bigcap_m (T^{\delta,\epsilon})^m(\bar{F})$.

PROPOSITION 4. *We have $V_\epsilon^\delta = \bar{F}_\infty^{\delta,\epsilon} = \lim_{m\to\infty}(T^{\delta,\epsilon})^m(\bar{F})$.*

PROOF. Because (i) V_ϵ^δ is a fixed point of $T^{\delta,\epsilon}$, (ii) $V_\epsilon^\delta \subseteq \bar{F}$, and (iii) $T^{\delta,\epsilon}$ is monotone (in the sense that $F' \subseteq F \Rightarrow T^{\delta,\epsilon}(F') \subseteq T^{\delta,\epsilon}(F)$), we have that

$$V_\epsilon^\delta \subseteq (T^{\delta,\epsilon})^{m+1}(\bar{F}) \subseteq (T^{\delta,\epsilon})^m(\bar{F}) \subseteq \bar{F} \quad \forall m \geq 1.$$

Thus, $V_\epsilon^\delta \subseteq \bar{F}_\infty^{\delta,\epsilon}$. Also, note that because $\bar{F}$ is compact (by Tychonoff's theorem), so is $(T^{\delta,\epsilon})^m(\bar{F})$ for each m. (The argument is similar to the one in the proof of the second claim in Lemma 1.) Because the set of compact subsets of $\bar{F}$ is compact, the decreasing sequence $((T^{\delta,\epsilon})^m(\bar{F}))_m$ has a limit:

$$\bar{F}_\infty^{\delta,\epsilon} = \lim_{m\to\infty}(T^{\delta,\epsilon})^m(\bar{F}).$$

Finally, to show that $\bar{F}_\infty^{\delta,\epsilon} \subseteq V_\epsilon^\delta$, we argue that $\bar{F}_\infty^{\delta,\epsilon}$ is self-δ, ϵ-FIR (that is, that $\bar{F}_\infty^{\delta,\epsilon} \subseteq T^{\delta,\epsilon}(\bar{F}_\infty^{\delta,\epsilon})$) and thus must lie in the largest self-δ, ϵ-FIR collection V_ϵ^δ. Because the operator $T^{\delta,\epsilon}$ is upper hemicontinuous,

$$\begin{aligned}\bar{F}_\infty^{\delta,\epsilon} &= \lim_{m\to\infty}(T^{\delta,\epsilon})^m(\bar{F}) \\ &= \lim_{m\to\infty} T^{\delta,\epsilon}[(T^{\delta,\epsilon})^m(\bar{F})]\end{aligned}$$

$$\subseteq T^{\delta,\epsilon}\Big[\lim_{m\to\infty}(T^{\delta,\epsilon})^m(\bar{F})\Big]$$
$$= T^{\delta,\epsilon}(\bar{F}^{\delta,\epsilon}_{\infty}). \qquad \square$$

Appendix C: Preliminaries for and proof of Lemma 2

The proof of Lemma 2 follows the approach in FLM: for each state s, we identify a subset of payoffs and show that it can be decomposed on tangent hyperplanes. There are two major differences. First, we use pseudo-instantaneous payoffs instead of stage-game payoffs in state s to take into account continuation payoffs after the transition out of state s. Second, we need to ensure that we can choose the same discount factor *uniformly* for all, possibly infinitely many, states s. We therefore require tighter bounds on the ranges of the supporting in-state continuation payoffs than are needed in FLM. Section C.1 contains some preliminary results and the estimates that bound the sizes of the solutions to the systems of linear equations. Section C.2 ensures that every (correlated) action profile can be approximated by profiles that satisfy appropriate rank conditions. Section C.3 discusses enforceability, and the last section contains the proof of Lemma 2.

C.1 *Preliminary results*

We need two preliminary results. The first result provides a bound on the size of solutions to a system of linear equations. For any vector $x \in \mathbb{R}^n$, let $\|x\|_\infty \equiv \max_i |x_i|$ denote the sup norm. (Recall that $\|x\|$ denotes the Euclidean norm and notice that $\|x\|_\infty \le \|x\| \le n\|x\|_\infty$.) For each matrix A with generic element a_{ij}, let $\|A\|_\infty = \max_{ij}|a_{ij}|$.

Lemma 3. *Let positive integers $j \le n$, matrix $A \in \mathcal{M}_{jn}$, and vector $b \in \mathbb{R}^j$ be given.*

Case 1. We have $d_j(A) > 0$.

Case 2. We have $d_{j-1}(A) > 0$ and there exists a nonzero vector $a \in \mathbb{R}^j$ such that $a'b = 0$ and $a'A = \mathbf{0}$.

If either Case 1 or 2, then there exists $w \in \mathbb{R}^n$ such that

$$Aw = b$$

and

$$\|w\|_\infty \le \frac{1}{d_k(A)}\|A\|^n_\infty \|b\|_\infty,$$

where $k = j$ in Case 1 and $k = j-1$ in Case 2.

Proof. By the definition of $d_k(A)$, there exists matrix $B \in \mathcal{M}_k$ such that $B \subseteq A$ and $|\det B| = d_k(A)$. Define vector $\tilde{b} \in \mathbb{R}^k$ as $\tilde{b} = b$ in Case 1, and in Case 2, define it as the vector obtained from b by crossing out the same row that is crossed out from matrix A so as to obtain B. Let $\tilde{w} = B^{-1}\tilde{b} \in \mathbb{R}^k$. Then, by Cramer's rule, $\|\tilde{w}\|_\infty \le \|A\|^n_\infty \|b\|_\infty / d_k(A)$. Let $w_i = \tilde{w}_i$ for each column i that is not crossed out in matrix B and let $w_i = 0$ for all

other columns. Then vector w satisfies the required bound and $Aw = b$. (In Case 1, that equality is immediate; in Case 2 it follows from the existence of the vector a.) □

The second result, which will be used for applying the Identifiability Condition, provides a lower bound on the local "variability" of nonzero polynomials. Recall that m^* is the maximum number of actions available in any state. For each positive integer n, let $\mathcal{F}_{n,m*}$ be the space of polynomial functions $f : \mathbb{R}^{m^*} \to \mathbb{R}$ with m^* variables and of order not higher than n. We consider restrictions of such polynomials to the simplex Δ_{m^*} of probability distributions α over action profiles. For each $c \in (0, 1)$, let $\mathcal{F}^*_{n,m^*}(c) \subseteq \mathcal{F}_{n,m^*}$ be the subspace of polynomials f such that $\sup_{\alpha\in\Delta_{m^*}} |f(\alpha)| \in [c, 1]$.

LEMMA 4. *For each n, $c > 0$, and $\epsilon > 0$, there exists a constant $\bar{c} > 0$ such that for each polynomial $f \in \mathcal{F}^*_{n,m^*}(c)$ and each profile $\alpha \in \Delta_{m^*}$, there exists a profile $\alpha' \in \Delta_{m^*}$ such that $\|\alpha - \alpha'\| \le \epsilon$ and $|f(\alpha')| \ge \bar{c}$.*

PROOF. On the contrary, suppose that there is a sequence of polynomials $f_k \in \mathcal{F}^*_{n,m^*}(c)$ and profiles α_k such that $\sup_{\alpha':\|\alpha_k-\alpha'\|\le\epsilon} |f_k(\alpha')| \to 0$. Because $\mathcal{F}^*_{n,m^*}(c)$ and Δ_{m^*} are compact, we can choose a subsequence so that $f_k \to f^* \in \mathcal{F}^*_{n,m^*}(c)$ and $\alpha_k \to \alpha^*$, and $\sup_{\alpha':\|\alpha^*-\alpha'\|\le\epsilon} |f^*(\alpha')| = 0$. Because f^* is a polynomial, this implies that $f^* \equiv 0$, which contradicts the fact that $\sup_{\alpha\in\Delta_{m^*}} |f^*(\alpha)| \ge c$. □

C.2 *Results on identifiability*

For each state s, player i, (mixed) action profile α, and $\delta \le 1$, define the $m_i(s) \times m$ matrix

$$\Gamma^\delta_i(\alpha, s) = \begin{bmatrix} E_{(a_i^{(1)},\alpha_{-i})}[(1-\beta^\delta(a,s))\rho(a,s)] \\ E_{(a_i^{(2)},\alpha_{-i})}[(1-\beta^\delta(a,s))\rho(a,s)] \\ \vdots \\ E_{(a_i^{(m_i(s))},\alpha_{-i})}[(1-\beta^\delta(a,s))\rho(a,s)] \end{bmatrix},$$

where $a_i^{(1)}, \ldots, a_i^{(m_i(s))}$ is an enumeration of all actions in $A_i(s)$. For any two players i and j, let

$$\Gamma^\delta_{ij}(\alpha, s) = \begin{bmatrix} \Gamma^\delta_i(\alpha, s) \\ \Gamma^\delta_j(\alpha, s) \end{bmatrix}.$$

Observe that $\Gamma^1_i(\alpha, s) = \Pi_i(\alpha, s)$ and $\Gamma^1_{ij}(\alpha, s) = \Pi_{ij}(\alpha, s)$.

LEMMA 5. *If the Identifiability Condition holds, then for each $\epsilon > 0$, there exist $\delta_\epsilon < 1$ and $d_\epsilon > 0$ such that for each state s, each action profile α, and player i, the following statements hold:*

(i) There exists an action profile α'_{-i} for players other than i such that $\|\alpha_{-i} - \alpha'_{-i}\| \le \epsilon/M$ and for each action $a_i \in A_i(s)$, each $\delta \ge \delta_\epsilon$, and each player $j \ne i$, $d_{m_j(s)}(\Gamma^\delta_j(a_i, \alpha'_{-i})) \ge d_\epsilon$.

(ii) There exists an action profile α' such that $\|\alpha - \alpha'\| \leq \epsilon/M$ and for each $\delta \geq \delta_\epsilon$, $d_{m_i(s)+m_j(s)-1}(\Gamma_{ij}^\delta(\alpha')) \geq d_\epsilon$.

PROOF. If $\delta = 1$, then the existence of a $d'_\epsilon > 0$ satisfying the two conditions follows from the fact that the determinant of any of the relevant matrices is a nonzero polynomial in the mixed strategies of the players, Lemma 4, and the proofs of Lemmas 6.2 and 6.3 from FLM.

For square matrices A and X of the same dimension, the following formula holds: $\det(A+\epsilon X) - \det A = \det(A)\operatorname{tr}(A^{-1}X)\epsilon + O(\epsilon^2)$. The error term $O(\epsilon^2)$ can be bounded by a $C\epsilon^2$, where C depends on m^* and upper bounds on $\|A\|_\infty$, $\|X\|_\infty$, and $\det(A)$. Since the monitoring probabilities lie between 0 and 1, then we conclude that there exists a constant $C' > 0$ such that

$$d_{m_i}(\Gamma_i^\delta(\alpha, s)) \geq d_{m_i}(\Gamma_i^1(\alpha, s)) - (1-\delta)C'$$

for each profile α, each state s, each $\delta < 1$, and all players i and j. That conclusion implies that we can take $d_\epsilon = d'_\epsilon/2$ and $\delta_\epsilon = 1 - d_\epsilon/2C'$. □

C.3 *Identifiability and enforceability*

Let U be the set of unit vectors in $\mathbb{R}^N$: $U \equiv \{\lambda \in \mathbb{R}^N | \|\lambda\| = 1\}$. For each unit vector $\lambda \in U$, let $N(\lambda) = \{i : \lambda_i \neq 0\}$ and $b(\lambda) = \min_{i \in N(\lambda)} |\lambda_i|$. Say that vector $\lambda \in \mathbb{R}^N$ is *regular* if $\#N(\lambda) \geq 2$. Recall that m is the number of public signals. The following result follows directly from Case 1 of Lemma 3.

COROLLARY 3. *For each $\delta < 1$, $d > 0$, player i, state s, profile α^* such that $d_{m_i(s)}(\Gamma_i^\delta(\alpha^*, s)) \geq d$, and vector $x \in \mathbb{R}^{m_i(s)}$ such that $\|x\|_\infty \leq M$, there exists $w \in \mathbb{R}^m$ such that $\Gamma_i^\delta(\alpha^*, s)w = x$ and $\|w\|_\infty \leq d^{-1}M$.*

The next result requires only slightly more work.

LEMMA 6. *For each $\delta < 1$, $d > 0$, state s, profile α^* such that $d_{m_i(s)+m_j(s)-1}(\Gamma_{ij}^\delta(\alpha^*)) \geq d$ for all players i and j, each regular unit vector $\lambda \in \mathbb{R}^N$, and each collection of vectors $\{x_i\}_{i=1}^N \in \times_i \mathbb{R}^{m_i(s)}$ such that $\|x_i\|_\infty \leq M$ and $\alpha_i^* \cdot x_i = 0$ for all i, there exists a mapping $w : Y \to \mathbb{R}^N$ such that for each player i,*

$$\Gamma_i^\delta(\alpha^*, s)w_i = x_i,$$

and, for each y, $\lambda \cdot w(y) = 0$ and $\|w(y)\|_\infty \leq NM/db(\lambda)$.

PROOF. Pick any $i, j \in N(\lambda)$ such that $i \neq j$. By definition of the matrices Γ^δ, $\alpha_i^* \cdot \Gamma_i^\delta(\alpha^*, s) = \alpha_j^* \cdot \Gamma_j^\delta(\alpha^*, s)$. Case 2 of Lemma 3 implies that there exists $w^{i,j} \in \mathbb{R}^m$ such that

$$\Gamma_{ij}^\delta(\alpha^*, s)w^{i,j} = \begin{bmatrix} \frac{1}{\#N(\lambda)-1}x_i \\ -\frac{\lambda_j}{\lambda_i}x_j \end{bmatrix}$$

and $\|w^{i,j}\|_\infty \leq M/db(\lambda)$. Using Case 1 of Lemma 3, for each player $i \notin N(\lambda)$, there exists $w^i \in \mathbb{R}^m$ such that

$$\Gamma_i^\delta(\alpha^*, s)w^i = x_i$$

and $\|w^i\|_\infty \leq M/d$. Fix $i \in N(\lambda)$ and define $w(y)$ as

$$\begin{aligned} w_i(y) &= \sum_{j \in N(\lambda)\setminus\{i\}} w^{ij}(y) \\ w_j(y) &= -\frac{\lambda_i}{\lambda_j} w^{ij}(y) \quad \text{for each } j \in N(\lambda)\setminus\{i\} \\ w_j(y) &= w^j(y) \quad \text{for each } j \notin N(\lambda). \end{aligned}$$

The result follows. □

C.4 *Proving Lemma 2*

Lemma 2 follows easily from the following lemma:

LEMMA 7. *Suppose that the Identifiability Condition holds. Then for each $z \in B(0,\epsilon)$, there exist $\eta_z > 0$ and $\delta_z < 1$ such that for each $\delta \geq \delta_z$, each state s, each payoff vector $v^* \in V^\delta_{10\epsilon}(s)$, and each $v \in B(v^*,\epsilon) \cap B(v^* + z, \eta_z)$, there exist continuation payoffs $u \in \times_{s' \neq s} V^\delta_{10\epsilon}(s')$, profile α, and a function $w : Y \to B(v^*,\epsilon)$ such that (4) holds for each player i.*

PROOF. We first observe that the definition of collection V^δ_ϵ implies that for all states s and $v^* \in V^\delta_{10\epsilon}(s)$, the following conditions hold:

- For each unit vector λ, there exist continuation payoffs $u \in \times_{s' \neq s} V^\delta_{10\epsilon}(s')$ and profile α, such that $\lambda \cdot \psi^\delta(\alpha, s, u) \geq \lambda \cdot v^* + 10\epsilon$.
- For each player i, there exist continuation payoffs $u \in \times_{s' \neq s} V^\delta_{10\epsilon}(s')$ and profile α_{-i}, such that for each a_i, $v_i^* \geq \psi_i^\delta(a_i, \alpha_{-i}, s, u) + 10\epsilon$.

We consider four cases separately:

Case 1: $z = \epsilon\lambda$ for some regular $\lambda \in U$. Notice that $z \in \operatorname{bd} B(0,\epsilon)$ and λ is the normal vector to the boundary at z. Using the observation from the beginning of the proof, for each state s and $v^* \in V^\delta_{10\epsilon}(s)$, we can find continuation payoffs $u \in \times_{s' \neq s} V^\delta_{10\epsilon}(s')$ and profile α, such that

$$\lambda \cdot \psi^\delta(\alpha, s, u) \geq \lambda \cdot v^* + 10\epsilon \geq \lambda \cdot (v^* + z) + 9\epsilon.$$

(Note that $\lambda \cdot z \leq \epsilon$.) Using Lemma 5, we can always replace strategy profile α by profile α^* so that $\lambda \cdot \psi^\delta(\alpha^*, s, u) \geq \lambda \cdot (v^* + z) + 5\epsilon$ and $d_{m_i(s)+m_j(s)-1}(\Gamma^\delta_{ij}(\alpha^*)) \geq d_\epsilon > 0$ for all i, j and all $\delta \geq \delta_\epsilon$.

Take any $v \in B(v^*, \epsilon)$ and define

$$x_i(a_i) = \frac{1}{1-\delta} E_{(a_i,\alpha^*_{-i})}\beta^\delta(a,s)(v_i - \psi_i^\delta(a,s,u)) - \frac{1}{1-\delta}\frac{1-E_{(a_i,\alpha^*_{-i})}\beta^\delta(a,s)}{1-E_{\alpha^*}\beta^\delta(a,s)}\big((E_{\alpha^*}\beta^\delta(a,s))v_i - \big(E_{\alpha^*}[\beta^\delta(a,s)\psi_i^\delta(a,s,u)]\big)\big).$$

Note that $\|x_i\| \le 2(1+\gamma_{\max})M$ and $\alpha_i^* \cdot x_i = 0$. By Lemma 6, then, there exists $\hat{w}: Y \to \mathbb{R}^N$ such that $\lambda \cdot \hat{w}(y) = 0$ for each y, $\|\hat{w}(y)\|_\infty \le 2(1+\gamma_{\max})NM/d_\epsilon b(\lambda)$, and for each player i,

$$\Gamma_i^\delta(\alpha^*, s)\hat{w}_i = x_i.$$

Let

$$w(y) = v + \frac{1}{1-E_{\alpha^*}\beta^\delta(a,s)}\big((E_{\alpha^*}\beta^\delta(a,s))v - (E_{\alpha^*}\beta^\delta(a,s)\psi^\delta(a,s,u))\big) + (1-\delta)\hat{w}(y).$$

Then simple computations show that

$$\begin{aligned}
&E_{(a_i,\alpha^*_{-i})}\left(\beta^\delta(a,s)\psi_i^\delta(a,s,u) + [1-\beta^\delta(a,s)]\sum_{y\in Y}\rho(a,s)[y]w_i(y)\right)\\
&= E_{(a_i,\alpha^*_{-i})}\beta^\delta(a,s)\psi_i^\delta(a,s,u) + (1-E_{(a_i,\alpha^*_{-i})}\beta^\delta(a,s))v_i\\
&\qquad + \frac{1-E_{(a_i,\alpha^*_{-i})}\beta^\delta(a,s)}{1-E_{\alpha^*}\beta^\delta(a,s)}\big((E_{\alpha^*}\beta^\delta(a,s))v_i - (E_{\alpha^*}\beta^\delta(a,s)\psi_i^\delta(a,s,u))\big)\\
&\qquad + (1-\delta)E_{(a_i,\alpha^*_{-i})}[1-\beta^\delta(a,s)]\sum_{y\in Y}\rho(a,s)[y]\hat{w}(y)\\
&= (1-E_{(a_i,\alpha^*_{-i})}\beta^\delta(a,s))v_i + E_{(a_i,\alpha^*_{-i})}\beta^\delta(a,s)v_i = v_i.
\end{aligned}$$

Thus, (4) holds with equality for all players and all actions.

Additionally, observe that

$$\begin{aligned}
\lambda\cdot(v-w(y)) &= \frac{1}{1-E_{\alpha^*}\beta^\delta(a,s)}\lambda\cdot\big(E_{\alpha^*}\beta^\delta(a,s)\psi^\delta(a,s,u) - (E_{\alpha^*}\beta^\delta(a,s))v\big)\\
&= \frac{E_{\alpha^*}\beta^\delta(a,s)}{1-E_{\alpha^*}\beta^\delta(a,s)}\lambda\cdot(\psi^\delta(\alpha^*,s,u)-v)\\
&\ge \frac{E_{\alpha^*}\beta^\delta(a,s)}{1-E_{\alpha^*}\beta^\delta(a,s)}\epsilon \ge (1-\delta)\epsilon.
\end{aligned}$$

The first inequality holds because $v \in B(v^*, \epsilon)$, which then implies $\lambda \cdot \psi^\delta(\alpha^*, s, u) \ge \lambda \cdot v + 5\epsilon$. The second inequality holds because $E_{\alpha^*}\beta^\delta(a,s)/[1-E_{\alpha^*}\beta^\delta(a,s)] \ge 1-\delta$. Furthermore,

$$\begin{aligned}
\|v-w(y)\| &\le \left\|\frac{E_{\alpha^*}\beta^\delta(a,s)}{1-E_{\alpha^*}\beta^\delta(a,s)}v\right\| + \left\|\frac{E_{\alpha^*}\beta^\delta(a,s)\psi^\delta(a,s,u)}{1-E_{\alpha^*}\beta^\delta(a,s)}\right\| + (1-\delta)\|\hat{w}(y)\|\\
&\le (1-\delta)C^{\lambda,\epsilon}
\end{aligned}$$

for some constant $C^{\lambda,\epsilon}$ that depends on λ and ϵ, and on the constants $\gamma_{\max}$, M, and N.

Notice that for any two vectors a and b, $\|a+b\|^2 = \|a\|^2 + \|b\|^2 + 2a\cdot b$. Also, notice that for each $v \in B(v^*+z,\eta)\cap B(v^*,\epsilon)$, for any vector θ, $(v-v^*)\cdot\theta \le \epsilon\theta\cdot\lambda + \eta\|\theta\|$. It follows that for each $y\in Y$ and each $v\in B(v^*+z,\eta)\cap B(v^*,\epsilon)$,

$$\begin{aligned}\|w(y)-v^*\|^2 &= \|v-v^*+w(y)-v\|^2\\ &= \|v-v^*\|^2 + \|w(y)-v\|^2 + 2(v-v^*)\cdot(w(y)-v)\\ &\le \epsilon^2 + (C^{\lambda,\epsilon})^2(1-\delta)^2 - 2\epsilon^2(1-\delta) + 2\eta(1-\delta)C^{\lambda,\epsilon}.\end{aligned}$$

Let $\eta_z = \epsilon^2/2C^{\lambda,\epsilon}$. Then there exists δ_z such that for all $\delta\ge\delta_z$, $\|w(y)-v^*\|\le\epsilon$.

Case 2: $z_i = -\epsilon$ for some i and $z_j = 0$ for all $j\neq i$. For each state s and $v^*\in V^\delta_{10\epsilon}(s)$, we can find continuation payoffs $u\in\times_{s'\neq s}V^\delta_{10\epsilon}(s')$ and profile α_{-i} such that for each a_i, $v_i^*\ge \psi_i^\delta(a_i,\alpha_{-i},s,u)+10\epsilon$. Using Lemma 5, we can replace α_{-i} by profile α^*_{-i} such that for each a_i, $v_i^*\ge\psi_i^\delta(a_i,\alpha^*_{-i},s,u)+5\epsilon$ and $d_{m_j(s)}(\Gamma_j^\delta(a_i,\alpha^*_{-i}))\ge d_\epsilon>0$ for all $\delta\ge\delta_\epsilon$. Let a_i^* be an action that maximizes $\psi_i^\delta(a_i,\alpha^*_{-i},s,u)$ and let $\alpha^*=(a_i^*,\alpha^*_{-i})$.

Take any $v\in B(v^*,\epsilon)$ and for each $j\neq i$, let

$$x_j(a_j) = \frac{1}{1-\delta}E_{(a_j,\alpha^*_{-j})}\beta^\delta(a,s)(v_j-\psi_j^\delta(a,s,u)).$$

Then, since $\|x_j\|_\infty\le 2(1+\gamma_{\max})M$, by Corollary 3, there exists $\hat w: Y\to\mathbb{R}^N$ such that for each y, $\|\hat w(y)\|_\infty\le 2(1+\gamma_{\max})NM/d_\epsilon b(\lambda)$, and for all players $j\neq i$ and all actions a_j,

$$x_j(a_i) = E_{(a_j,\alpha^*_{-j})}(1-\beta^\delta(a,s))\sum_{y\in Y}\rho(a,s)[y]\hat w_j(y).$$

Let

$$w_j(y) = v + (1-\delta)\hat w(y)$$

for $j\neq i$ and let

$$w_i(y) = v_i + (1-\delta)E_{\alpha^*}\beta^\delta(a,s)(v_i-\psi_i^\delta(a,s,u)).$$

Then for all players $j\neq i$ and all actions a_j,

$$v_j = E_{(a_j,\alpha^*_{-j})}\left(\beta^\delta(a,s)\psi_j^\delta(a,s,u) + [1-\beta^\delta(a,s)]\sum_{y\in Y}\rho(a,s)[y]w_j(y)\right),$$

and for player i, a_i^* is a best response and it satisfies

$$v_i = E_\alpha\left(\beta^\delta(a,s)\psi_i^\delta(a,s,u) + [1-\beta^\delta(a,s)]\sum_{y\in Y}\rho(\alpha,s)[y]w_i(y)\right).$$

Moreover,

$$\begin{aligned}w_i(y)-v_i &= (1-\delta)E_{\alpha^*}\beta^\delta(a,s)(v_i-\psi_i^\delta(a,s,u))\\ &\ge (1-\delta)(v_i-\psi_i^\delta(\alpha^*,s,u)) \ge (1-\delta)\epsilon,\end{aligned}$$

where the last inequality holds because $v \in B(v^*, \epsilon)$ and $\psi_i^\delta(\alpha^*, s, u) \leq v^* - 5\epsilon$. Second,

$$\begin{aligned}\|v - w(y)\| &\leq (1-\delta)\|\hat{w}(y)\| \\ &\leq (1-\delta)C^{\lambda,\epsilon}\end{aligned}$$

for some constant $C^{\lambda,\epsilon}$ that depends on ϵ. This case is concluded by the same argument as Case 1.

Case 3: $z_i = \epsilon$ for some i and $z_j = 0$ for all $j \neq i$. The proof of this case is analogous to that of Case 2.

Case 4: $z \in \operatorname{int} B(0, \epsilon)$. Fix state s and continuation payoffs u. Find profile α^δ that is a Nash equilibrium of a one-shot game with payoffs $\psi^\delta(a, s, u)$. For each $v \in B(v^*, \epsilon)$, let

$$w^\delta(y) = v + \frac{E_{\alpha^*}\beta^\delta(a,s)}{1 - (E_{\alpha^*}\beta^\delta(a,s))}(v - \psi^\delta(\alpha^\delta, s, u)).$$

Let $\eta_z = \frac{1}{2}(\epsilon - \|z\|)$ and $\delta_z = 1 - (\epsilon - \|z\|)/4(1+\gamma_{\max})M$. Then, for each $v \in B(v^*, \epsilon) \cap B(v^* + z, \eta_z)$ and for each $\delta \geq \delta_z$,

$$\begin{aligned}\|w^\delta(y) - v^*\| &\leq \|v - (v^* + z)\| + \|z\| + \|w^\delta(y) - v\| \\ &\leq \eta_z + \|z\| + 2(1-\delta)(1+\gamma_{\max})M \\ &\leq \epsilon.\end{aligned}$$

This concludes the proof of the lemma. □

Now we can complete the proof of Lemma 2.

Proof of Lemma 2. Lemma 7 shows that for each $z \in B(0, \epsilon)$, there exist $\delta_z < 1$ and η_z such that for each $\delta \geq \delta_z$, each state s, each payoff vector $v^* \in V_{10\epsilon}^\delta(s)$, and each $v \in B(v^*, \epsilon) \cap B(v^* + z, \eta_z)$, there exist continuation payoffs $u \in \times_{s' \neq s} V_{10\epsilon}^\delta(s')$, profile α, and a function $w : Y \to B(v^*, \epsilon)$ such that (4) holds for each player i. Because a closed ball in $\mathbb{R}^N$ is compact, there is a finite collection Z of z's such that $B(v^*, \epsilon) \subseteq \bigcup_Z B(v^* + z, \eta_z)$. Setting $\delta^* = \max_Z \delta_z$ yields the required $\delta^* < 1$ for Lemma 2. □

Appendix D: Proof of Theorem 2

D.1 *One-dimensional class of games*

For each $d \in \mathbb{R}^{\#S \times N}$, let $\mathcal{G}_0(d) \subseteq \mathcal{G}_0$ be a class of games such that there exists $\epsilon > 0$ such that for each state s, $B(d(s), \epsilon) \subseteq \operatorname{int} V_0^1(s)$. For each game G, let $\gamma_{\max}(G) = \max_{a,s} \gamma(a, s)$. For each game $G \in \mathcal{G}_0(d)$, we define a one-dimensional class of games indexed by $\eta \in (0, 1 + (\gamma_{\max}(G))^{-1})$, i.e., $G^{\eta;d} = (g^{\eta;d}, \gamma^{\eta;d})$, where

$$\begin{aligned}g^{\eta;d}(a,s) &= \frac{\eta}{1-(\eta-1)\gamma(a,s)}g(a,s) - \frac{(\eta-1)(1+\gamma(a,s))}{1-(\eta-1)\gamma(a,s)}d(s) \\ \gamma^{\eta;d}(s';a,s) &= \frac{\eta}{1-(\eta-1)\gamma(a,s)}\gamma(s';a,s).\end{aligned}$$

Notice that $G^{1,d} = G$. We choose the parametrization so that pseudo-instantaneous payoffs $\psi^1(a, s, u; G^{\eta;d})$ expand radially from $d(s)$ relative to payoffs $\psi^1(a, s, u; G)$. The next result summarizes that property and other properties of the parametrization.

LEMMA 8. *For each $G \in \mathcal{G}_0(d)$ and each $\eta, \nu > 1$ such that $\eta\nu < 1 + (\gamma_{\max}(G))^{-1}$, the following statements hold:*

(i) We have $(G^{\eta;d})^{\nu;d} = G^{\eta\nu;d}$.

(ii) For each action profile a, each state s, and all continuation payoffs $u \in \times_{s' \neq s} \mathbb{R}^N$, $\psi^1(a, s, u; G^{\eta;d}) - d(s) = \eta(\psi^1(a, s, u; G) - d(s))$.

(iii) There exists $\epsilon > 0$ such that for each state s, $V_0^1(s) \subseteq V_\epsilon^1(s; G^{\eta,d})$.

(iv) We have $G^{\eta;d} \in \mathcal{G}_0(d)$.

PROOF. Part (i). Notice that

$$\begin{aligned}\frac{\eta}{1-(\eta-1)\frac{\nu}{1-(\nu-1)\gamma(a,s)}\gamma(a,s)}\frac{\nu}{1-(\nu-1)\gamma(a,s)} &= \frac{\eta\nu}{1-(\nu-1)\gamma(a,s)-(\eta-1)\nu\gamma(a,s)}\\ &= \frac{\eta\nu}{1-(\eta\nu-1)\gamma(a,s)}.\end{aligned}$$

This implies that $(\gamma^{\eta;d})^{\nu;d} = \gamma^{\eta\nu;d}$. In a similar way, we show that $(g^{\eta;d})^{\nu;d} = g^{\eta\nu;d}$.

Part (ii). For each $\eta > 0$,

$$\begin{aligned}\psi^1(a, s, u; G^{\eta;d}) - d(s) &= \frac{g^{\eta;d}(a, s) + \sum_{s' \neq s} \gamma^{\eta;d}(s'; a, s)u(s')}{1 + \gamma^{\eta;d}(a, s)} - d(s)\\ &= \frac{\eta}{1+\gamma(a,s)}g(a, s) - (\eta - 1)d(s)\\ &\qquad + \frac{\sum_{s' \neq s} \eta\gamma(s'; a, s)u(s')}{1 + \gamma(a, s)} - d(s)\\ &= \eta\left(\frac{g(a, s) + \sum_{s' \neq s} \gamma(s'; a, s)u(s')}{1 + \gamma(a, s)} - d(s)\right)\\ &= \eta(\psi^1(a, s, u) - d(s)).\end{aligned}$$

Part (iii). Choose $\epsilon' > 0$ such that for each s, $B(d(s), \epsilon') \subseteq V_0^1(s)$. Let $\epsilon = (\eta - 1)\epsilon'$. We show that collection $V_0^1(\cdot; G)$ is self-$(1, \epsilon)$-individually rational in game $G^{\eta;d}$. Using part (ii), we get

$$\begin{aligned}&e_i^1(s; V_0^1(\cdot; G), G^{\eta,d})\\ &\quad= \inf_{\alpha_{-i} \in \times_{j \neq i} \Delta A_j(s), u \in \times_{s' \neq s} V_0^1(s'; G)} \left\{\max_{\alpha_i \in \Delta A_i(s)} \psi_i^1(\alpha_i, \alpha_{-i}, s, u; G^{\eta;d})\right\}\\ &\quad= \inf_{\alpha_{-i} \in \times_{j \neq i} \Delta A_j(s), u \in \times_{s' \neq s} V_0^1(s'; G)} \left\{\max_{\alpha_i \in \Delta A_i(s)} \left[d_i(s) + (\psi_i^1(\alpha_i, \alpha_{-i}, s, u; G^{\eta;d}) - d_i(s))\right]\right\}\end{aligned}$$

$$= \inf_{\alpha_{-i}\in\times_{j\neq i}\Delta A_j(s),u\in\times_{s'\neq s}V_0^1(s';G)}\left\{\max_{\alpha_i\in\Delta A_i(s)}\left[d_i(s)+\eta(\psi_i^1(\alpha_i,\alpha_{-i},s,u;G)-d_i(s))\right]\right\}$$

$$= -(\eta-1)d_i(s)+\eta\inf_{\alpha_{-i}\in\times_{j\neq i}\Delta A_j(s),u\in\times_{s'\neq s}V_0^1(s';G)}\left\{\max_{\alpha_i\in\Delta A_i(s)}\psi_i^1(\alpha_i,\alpha_{-i},s,u;G)\right\}$$

$$= d_i(s)+\eta\big(e_i^1(s;V_0^1(\cdot;G),G)-d_i(s)\big).$$

Because $e_i^1(s;V_0^1(\cdot;G),G)\leq d_i(s)-\epsilon'$, it must be that $e_i^1(s;V_0^1(\cdot;G),G^{\eta,d})\leq e_i^1(s;V_0^1(\cdot;G),G)-(\eta-1)\epsilon'$. This implies that $V_0^1(\cdot;G)$ is self-$(1,\epsilon)$-individually rational in game $G^{\eta;d}$.

Next, we show that collection $V_0^1(\cdot;G)$ is self-$(1,\epsilon)$-feasible. Take any $v_0\in V_0^1(s)$ and $v\in B(v_0,\epsilon)$. For each $\lambda\in\mathbb{R}^N$ such that $\max_i|\lambda_i|=1$,

$$\lambda\cdot v\leq\epsilon+\lambda\cdot v_0\leq(\eta-1)\epsilon'+\sup_{v'\in V_0^1(s)}\lambda\cdot v'$$

$$=(\eta-1)\epsilon'+\lambda\cdot\psi^1(a^\lambda,s,u^\lambda;G)$$

for some action profile a^λ and continuation payoffs $u^\lambda\in\times_{s'\neq s}V_0^1(s';G)$, where

$$(a^\lambda,u^\lambda)\in\operatorname*{arg\,max}_{a\in A(s),u\in\times_{s'\neq s}V_0^1(s')}\lambda\cdot\psi^1(a,s,u;G).$$

Because $V_0^1(\cdot;G)$ is self-1, 0-feasible in game G and $B(d(s),\epsilon')\subseteq V_0^1(s)$, it must be that $\epsilon'\leq\lambda\cdot(\psi^1(a^\lambda,s,u^\lambda;G)-d(s))$ and

$$\begin{aligned}\lambda\cdot v&\leq(\eta-1)\epsilon'+\lambda\cdot\psi^1(a^\lambda,s,u^\lambda;G)\\&\leq(\eta-1)\lambda\cdot(\psi^1(a^\lambda,s,u^\lambda;G)-d(s))+\lambda\cdot(\psi^1(a^\lambda,s,u^\lambda;G)-d(s))+\lambda\cdot d(s)\\&=\eta\lambda\cdot(\psi^1(a^\lambda,s,u^\lambda;G)-d(s))+\lambda\cdot d(s)\\&=\lambda\cdot(\psi^1(a^\lambda,s,u^\lambda;G^{\eta;d})-d(s))+\lambda\cdot d(s)\\&=\lambda\cdot\psi^1(a^\lambda,s,u^\lambda;G^{\eta;d}),\end{aligned}$$

where the second-to-last equality comes from part (ii). The above implies that

$$v\in\operatorname{co}\{\psi^1(a,s,u;G^{\eta;d}):a\in A(s)\text{ and }u\in\times_{s'\neq s}V_0^1(s')\}$$

and that V_0^1 is self-$(1,\epsilon)$-feasible in game $G^{\eta;d}$.

Part (iv) follows from part (iii) and the fact that $V_\epsilon^1(s;G^{\eta;d})\subseteq V_0^1(s;G^{\eta;d})$. □

D.2 *Intermediate results*

In the next two parts of this appendix, we assume that the space state is finite. The proof of Theorem 2 requires three intermediate results. Let $\mathcal{A}\subseteq\mathcal{G}_0$ be the set of all games with Property A. Recall that Λ is a Lebesgue measure on the space of games. Let Λ_N be a Lebesgue measure on $\mathbb{R}^N$.

LEMMA 9. *The set $\mathcal{A}$ is measurable.*

PROOF. For each $\epsilon \geq 0$ and each $s \geq 0$, define function $f_{\epsilon,s} : \mathcal{G} \to \mathbb{R}$ as

$$f_{\epsilon,s}(G) = \Lambda_N(V^1_\epsilon(s; G)).$$

Because the self ϵ-FIR correspondence is upper hemicontinuous, function $f_{\epsilon,s}$ is upper semicontinuous and hence measurable. Define $f_{0+,s} = \lim_{\epsilon \to 0} f_{\epsilon,s}$. Then function $f_{0+,s}$ is measurable. Finally, notice that $\mathcal{A} = \bigcap_s \{G : f_{0,s}(G) = f_{0+,s}(G)\}$. □

LEMMA 10. *For each game $G \in \mathcal{G}_0(d)$, the set*

$$E = \left\{\eta \in \left(0, 1 + (\gamma_{\max}(G))^{-1}\right) : G^{\eta,d} \in \mathcal{G}(d) \setminus \mathcal{A}\right\}$$

is countable.

PROOF. We will prove a weaker claim: for each game $G \in \mathcal{G}_0(d)$, there exist at most countably many $1 \leq \eta \leq 1 + (\gamma_{\max}(G))^{-1}$ such that $G^{\eta;d} \notin \mathcal{A}$. It turns out that the weaker claim delivers the lemma. On the contrary, if the set E is uncountable, then there exists $\eta^* \in E$ such that the set $\{\eta \in E : \eta > \eta^*\}$ is uncountable. But because of Lemma 8, this contradicts our claim applied to $G^{\eta^*,d}$.

We move to the proof of the claim. Define two functions of $\eta \in [1, 1 + (\gamma_{\max}(G))^{-1})$:

$$\mu_0(\eta) = \sum_s \Lambda_N(V^1_0(s; G^{\eta;d}))$$

$$\mu(\eta) = \sum_s \Lambda_N\left(\bigcup_{\epsilon>0} V^1_\epsilon(s; G^{\eta;d})\right).$$

Because $V^1_\epsilon(s) \subseteq V^1_0(s)$, we have $\mu_0(\eta) \geq \mu(\eta)$ for each η, and game $G^{\eta;d}$ has Property A if and only if $\mu_0(\eta) = \mu(\eta)$. Also, notice that part (iii) of Lemma 8 implies that for each $\eta > 1$, $\mu(\eta) > \mu_0(1)$. That fact, together with part (i) of the lemma, implies that the functions μ_0 and μ are strictly increasing. Part (iii) implies that $\lim_{\eta' \searrow \eta} \mu(\eta) = \mu_0(\eta)$ for each η. It follows that

$$\mu(\eta) \geq \mu(1) + \sum_{\eta' \in E, \eta' < \eta} (\mu_0(\eta') - \mu(\eta')).$$

If there are uncountably many elements in E, then there exists η such that the right-hand side of the above inequality is unbounded. But this contradicts the fact that $\mu(\eta)$ is a well defined function for each $\eta \in [1, 1 + (\gamma_{\max}(G))^{-1})$. □

By Lemma 9, set $\mathcal{A}$ is measurable, and we can state the following result.

LEMMA 11. *For each $d \in \mathbb{R}^{\#S \times N}$, $\Lambda(\mathcal{G}_0(d) \setminus \mathcal{A}) = 0$.*

PROOF. Let $\mathcal{G}_{00} \subseteq \mathcal{G}_0$ be the subclass of all games such that $\gamma_{\max}(G) > 0$ and let $\mathcal{G}^*_{00} \subseteq \mathcal{G}_{00}$ be the subclass of all games such that $\gamma_{\max}(G) = 1$. Then there exists

a continuous mapping $h:\mathcal{G}_{00}\to\mathcal{G}_{00}^*\times(0,2)$: for each $G=(g,\gamma)\in\mathcal{G}_{00}$, let $h(G)=(G^{\eta(\gamma_{\max}(G));d},(\eta(\gamma_{\max}(G)))^{-1})$, where

$$\eta(\gamma)=\frac{1+\gamma}{2\gamma}.$$

Note that $\mathcal{G}_{00}^*$ is an open subset of the union of finitely many affine subspaces of $\mathcal{G}$ (i.e., the subspace that contains all games with the maximum transition rate equal to 1). Hence, it can be equipped with a Lebesgue measure. This implies that $\mathcal{G}_{00}^*\times(0,2)$ can be equipped with Lebesgue measure Λ^* as well. By Fubini's theorem and Lemma 10, $\Lambda^*(h((\mathcal{G}_0(d)\cap\mathcal{G}_{00})\setminus\mathcal{A}))=0$. (Notice that all sets involved here are measurable.[12]) Because the mapping h is differentiable and has a differentiable inverse, it follows that $\Lambda((\mathcal{G}_0(d)\cap\mathcal{G}_{00})\setminus\mathcal{A})=0$.

Finally, Remark 1 shows that all games in $\mathcal{G}_0\setminus\mathcal{G}_{00}$ have Property A. □

D.3 *Proof of Theorem 2*

Lemma 11 says that the Lebesgue measure of $\mathcal{G}_0(d)\setminus\mathcal{A}$ is equal to 0. Notice that

$$\mathcal{G}_0\setminus\mathcal{A}=\bigcup_{d\in\mathbf{Q}^{\#S\times N}}\mathcal{G}_0(d)\setminus\mathcal{A},$$

where $\mathbf{Q}$ is the set of rational numbers. Thus, $\mathcal{G}_0\setminus\mathcal{A}$ has zero Lebesgue measure: it is the union of countably many sets with zero Lebesgue measure.

Appendix E: Proving Propositions 1 and 2

Rest of the proof of Proposition 1. We want to show that $V_\epsilon^1(s_1)=V_\epsilon^1(s_2)=\varnothing$ $\forall\epsilon>0$. Consider $\epsilon>0$ and suppose that $V_\epsilon^1(s_1)$ is nonempty. It follows by the symmetry of the game that $V_\epsilon^1(s_2)$ is nonempty as well. Let

$$l=\max_{(v_1,v_2)\in V_\epsilon^1(s_1)}(2v_1+v_2)$$

and note that (by symmetry again) l also satisfies

$$l=\max_{(v_1,v_2)\in V_\epsilon^1(s_2)}(v_1+2v_2).$$

Thus, any $(v_1,v_2)\in V_\epsilon^1(s_2)$ satisfies $v_2\le\frac{1}{2}l-\frac{1}{2}v_1$. Further, since $B((v_1,v_2),\epsilon)$ must lie in $\hat{V}^1(s)\cap\{(v_1,v_2)\in\mathbb{R}^2:v_1,v_2\ge0\}$, it must be that $v_2\ge\epsilon$. Therefore, $v_1\le l-2\epsilon$, since $v_1+2v_2\le l$. Finally, define $V_\epsilon^1(s_2|1)\equiv\{v_1\in\mathbb{R}:\exists v_2\text{ s.t. }(v_1,v_2)\in V_\epsilon^1(s_2)\}$ and note that

$$\begin{aligned}l&\le\max_{a\in\{H,L\},(v_1,v_2)\in V_\epsilon^1(s_2)}\left\{2\psi_1^1(a,s_1,(v_1,v_2))+\psi_2^1(a,s_1,(v_1,v_2))\right\}\\&=\tfrac{1}{2}\max_{a\in\{H,L\}}\{2g_1(a,s_1)+g_2(a,s_1)\}+\tfrac{1}{2}\max_{(v_1,v_2)\in V_\epsilon^1(s_2)}\{2v_1+v_2\}\end{aligned}$$

[12] We are grateful to an anonymous referee for pointing out an omission in our earlier proof.

$$\begin{aligned}
&\leq 0 + \tfrac{1}{2} \max_{v_1 \in V_\epsilon^1(s_2|1)} \left\{ 2v_1 + \tfrac{1}{2}l - \tfrac{1}{2}v_1 \right\} \\
&\leq \max_{v_1 \in V_\epsilon^1(s_2|1)} \left\{ \tfrac{3}{4}v_1 + \tfrac{1}{4}l \right\} \\
&\leq \tfrac{3}{4}(l - 2\epsilon) + \tfrac{1}{4}l \\
&\leq l - \tfrac{3}{2}\epsilon.
\end{aligned}$$

But this is a contradiction, and so both $V_\epsilon^1(s_1)$ and $V_\epsilon^1(s_2)$ must be empty. □

The proof of Proposition 2 is very similar to the proof of Proposition 1.

Proof of Proposition 2. First, note that playing L after every history is a perfect public equilibrium, so $(0,0) \in E^\delta(s)$. Next, let

$$l = \max_{(v_1,v_2) \in E^\delta(s_1)} (2v_1 + v_2)$$

and note that (by symmetry) l also satisfies

$$l = \max_{(v_1,v_2) \in E^\delta(s_2)} (v_1 + 2v_2).$$

Thus, any $(v_1, v_2) \in E^\delta(s_2)$ satisfies $v_2 \leq \frac{1}{2}l - \frac{1}{2}v_1$. Since individual rationality requires that $v_1, v_2 \geq 0$, we must have $l \geq v_1 \geq 0$. Furthermore, the condition that $v_1, v_2 \geq 0$ implies that if $E^\delta(s_2)$ contains any point other than $(0,0)$, then $l > 0$. Define $E^\delta(s_2|1) \equiv \{v_1 \in \mathbb{R} : \exists v_2 \text{ s.t. } (v_1, v_2) \in E^\delta(s_2)\}$ and note that

$$\begin{aligned}
l &\leq \max_{a \in \{H,L\}, (v_1,v_2) \in E^\delta(s_2)} \left\{ 2\psi_1^\delta(a, s_1, (v_1, v_2)) + \psi_2^\delta(a, s_1, (v_1, v_2)) \right\} \\
&= \frac{1}{1+\delta} \max_{a \in \{H,L\}} \{2g_1(a, s_1) + g_2(a, s_1)\} + \frac{\delta}{1+\delta} \max_{(v_1,v_2) \in E^\delta(s_2)} \{2v_1 + v_2\} \\
&\leq 0 + \frac{\delta}{1+\delta} \max_{v_1 \in E^\delta(s_2|1)} \left\{ 2v_1 + \tfrac{1}{2}l - \tfrac{1}{2}v_1 \right\} \\
&\leq \frac{\delta}{1+\delta} 2l \\
&\leq l.
\end{aligned}$$

The final inequality is strict if $l > 0$, so we conclude that $l = 0$, and that $E^\delta(s_2)$ and (by symmetry) $E^\delta(s_1)$ are equal to $\{(0,0)\}$. □

Appendix F: Proof of Proposition 3

Let U be the set of all unit vectors and let U_+ be the subset of vectors with nonnegative coordinates. For each state s, let

$$F(s) \equiv \{v : v_i \geq 0 \text{ and } \lambda \cdot v \leq c^\lambda(s) \text{ for each } \lambda \in U\}.$$

We divide the proof of the proposition into four steps.

Step 1: $V_0^1(s) \subseteq F(s)$ for each state s. To see this, notice that the (limit) set of all feasible payoffs is equal to

$$\hat{V}^1(s) = \limsup_{\delta \to 1} \hat{V}^\delta(s) = \{v : \lambda \cdot v \leq c^\lambda(s) \text{ for each } \lambda \in U\}.$$

Moreover, the Nash minmax for each player is equal to 0. The claim follows from the fact that the self-FIR payoffs must be feasible and Nash-individually rational; see Section 7.1.

Step 2: F is self-$(1,0)$-feasible. Fix state s, and take any unit vector $\lambda \in U$ and $v \in F(s)$. It is enough to show that there exists an action profile $a \in A$ and a continuation payoffs $u \in \times_{s' \neq s} F(s')$ such that $\lambda \cdot v \leq \lambda \cdot \psi^1(a, s, u)$. When λ has only nonpositive coordinates, the claim follows from the payoff assumptions and the fact that $(0, \dots, 0) \in F(s)$ for each state s. From now on, suppose that λ has some strictly positive entries, and define $\lambda^+ \in U_+$ by $\lambda_i^+ \equiv c \max(\lambda_i, 0)$ for all i, where $c \equiv 1/\sqrt{\sum_j (\max(\lambda_j, 0))^2} > 0$ is a constant of proportionality.

As an intermediate step, notice that for each state s', each $x \in \hat{V}^1(s')$, and each set of players $N_0 \subsetneq \{1, \dots, N\}$, there exists $x^+ \in F(s')$ such that $x_i^+ \geq \max(x_i, 0)$ for each player $i \notin N_0$ and $x_i^+ = 0$ for each player $i \in N_0$. The claim follows from the fact that each payoff vector x is the expected payoff from a strategy profile σ, given initial state s'. Consider a strategy profile σ' in which, after each history, each action of players $j \in N_0 \cup \{j : x_j < 0\}$ is replaced by the inactive action 0_j. Then, by the assumptions on the properties of the inactive actions, the payoff x^+ from such a profile has the required property.

Because $v \in \hat{V}^1(s)$, there exist action profile a' and continuation payoffs $u' \in \times_{s' \neq s} \hat{V}^1(s')$ such that $\lambda^+ \cdot v \leq \lambda^+ \cdot \psi^1(a', s, u')$. Let a be an action profile obtained from a' by replacing the actions of players $j \in N_0 = \{i : \lambda_i \leq 0\}$ by the inactive action 0_j. Let $u(s') = (u'(s'))^+ \in F(s)$, where $(u'(s'))^+$ is chosen as in the intermediate step. Then, because $v_i \geq 0$ for each player i,

$$\begin{aligned} \lambda \cdot v \leq \sum_{i \notin N_0} \lambda_i v_i = \frac{1}{c} \lambda^+ \cdot v \leq \frac{1}{c} \lambda^+ \cdot \psi^1(a', s, u') \\ \leq \frac{1}{c} \lambda^+ \cdot \psi^1(a, s, u) = \lambda \cdot \psi^1(a, s, u). \end{aligned}$$

The last inequality follows from the fact that $\psi_j^1(a', s, u') \leq \psi_j^1(a, s, u)$ for each player $j \notin N_0$, and the last equality follows from the fact that $\psi_j^1(a, s, u) = 0$ for each player $j \in N_0$.

Step 3: F is self-$(1,0)$-individually rational. The claim follows from the assumptions and the fact that $(0, \dots, 0) \in F(s)$ for each state s.

Steps 1, 2, and 3 establish that $V_0^1(s) = F(s)$ for each state s.

Step 4: Property A holds. By the assumption, there exists $v^* = (\eta/3N, \dots, \eta/3N)$ and, for each state s, a (possibly mixed) action profile α_s^* such that for each state s, $g(\alpha_s^*, s) = v^*$,

$$B\left(v^*, \frac{\eta}{3N}\right) \subseteq V_0^1, \quad \text{and} \quad B\left(v^*, \frac{\eta}{3N}\right) \subseteq \text{co}\{g(a, s) : a \in A\}. \tag{F.1}$$

For each $\epsilon > 0$ and each state s, define

$$F_\epsilon(s) = \epsilon v^* + (1-\epsilon)V_0^1(s).$$

Let $C = \eta/(3N(1+\gamma_{\max}(G)))$. We show that the collection F_ϵ is self-1, $C\epsilon$-FIR. To see self-1, $C\epsilon$-feasibility, notice that for each s,

$$\begin{aligned} F_\epsilon(s) &\subseteq \operatorname{co}\{\epsilon v^* + (1-\epsilon)\psi^1(a,s,u) : a \in A, u \in \times_{s'\neq s} V_0^1(s')\} \\ &= \operatorname{co}\left\{ \frac{\epsilon v^* + (1-\epsilon)g(a,s) + \sum_{s'\neq s}\gamma(s';a,s)(\epsilon v^* + (1-\epsilon)u_{s'})}{1+\gamma(a,s)} : \right. \\ &\qquad\qquad \left. a \in A, u \in \times_{s'\neq s} V_0^1(s') \right\} \\ &= \operatorname{co}\left\{ \frac{g((1-\epsilon)a + \epsilon\alpha_s^*, s) + \sum_{s'\neq s}\gamma(s';a,s)u_{s'}}{1+\gamma(a,s)} : a \in A, u \in \times_{s'\neq s} F_\epsilon(s') \right\}, \end{aligned}$$

so (F.1) implies that

$$B(F_\epsilon(s), C\epsilon) \subseteq \operatorname{co}\{\psi^1(a,s,u) : a \in A, u \in \times_{s'\neq s} F_\epsilon(s')\}.$$

Self-1, $C\epsilon$-rationality is proven using a similar argument. Thus, for each $v \in V_0^1(s)$ and each $\epsilon > 0$, there exists $v' \in F_\epsilon(s)$ such that $\|v - v'\| \leq \epsilon M$. It follows that $V_0^1(s) = \operatorname{cl} V_{0+}^1(s)$.

Appendix G: Markov strategies are not enough: Example

Example 6. The players and states are as in Example 1. The payoffs are

	H	L
H	$-1,3$	
L	$0,0$	

State s_1

H	L
$3,-1$	$0,0$

State s_2

As before, the transition rates in each state do not depend on actions and are equal to 1. ◊

The vector of Nash minmax payoffs in each state is $(0,0)$ for any discount factor. At $\delta = 1$, the sets of feasible payoffs in each state are

$$\begin{aligned} \hat{V}^1(s_1) &= \operatorname{co}\left\{(0,0), \left(\tfrac{1}{3}, \tfrac{5}{3}\right), \left(1, \tfrac{-1}{3}\right), \left(\tfrac{-2}{3}, 2\right)\right\} \\ \hat{V}^1(s_2) &= \operatorname{co}\left\{(0,0), \left(\tfrac{5}{3}, \tfrac{1}{3}\right), \left(\tfrac{-1}{3}, 1\right), \left(2, \tfrac{-2}{3}\right)\right\}. \end{aligned}$$

Recall that $V^{\text{Nash},1}(s)$ is the set of feasible payoffs, starting from state s, that give both players at least their Nash minmax payoffs:

$$V^{\text{Nash},1}(s) \equiv \hat{V}^1(s) \cap \{(v_1, v_2) \in \mathbb{R}^2 : v_1, v_2 \geq 0\}.$$

Those sets are given by

$$V^{\text{Nash},1}(s_1) = \text{co}\left\{(0,0), \left(\tfrac{1}{3}, \tfrac{5}{3}\right), \left(\tfrac{8}{9}, 0\right), \left(0, \tfrac{16}{9}\right)\right\}$$
$$V^{\text{Nash},1}(s_2) = \text{co}\left\{(0,0), \left(\tfrac{5}{3}, \tfrac{1}{3}\right), \left(0, \tfrac{8}{9}\right), \left(\tfrac{16}{9}, 0\right)\right\}.$$

We can use Corollary 2 to show that all of these payoffs can be achieved (approximately) in a PPE when players are patient.

PROPOSITION 5. *Suppose that the Identifiability Condition holds for the game in Example 6. Then for each $\eta > 0$, there exists $\delta^* < 1$ such that for each $v \in V^{\text{Nash},1}(s)$ and each $\delta \geq \delta^*$, there exists $v' \in E^\delta(s)$ such that $\|v - v'\| \leq \eta$.*

PROOF. Without loss of generality, consider initial state s_1. First, we show that $V_0^1(s_1) = V^{\text{Nash},1}(s_1)$: any self-1, 0-FIR set must lie in $(V^{\text{Nash},1}(s))_s$, so it is sufficient to show that $(V^{\text{Nash},1}(s))_s$ is self-1, 0-FIR. It is clearly self-1, 0-individually rational, since all payoffs are weakly positive. It is also self-1, 0-feasible: $(0,0) = \psi^1(L, s_1, (0,0))$, $(\frac{1}{3}, \frac{5}{3}) = \psi^1(L, s_1, (\frac{16}{9}, 0))$, $(\frac{8}{9}, 0) = \psi^1(H, s_1, (\frac{5}{3}, \frac{1}{3}))$, and $(0, \frac{16}{9}) = \psi^1(H, s_1, (1, \frac{5}{9}))$. (Note that $(1, \frac{5}{9}) = \frac{2}{5}(0, \frac{8}{9}) + \frac{3}{5}(\frac{5}{3}, \frac{1}{3})$, so $(1, \frac{5}{9}) \in \tilde{V}^1(s_2)$.) It is straightforward to verify that the collection $(V_\epsilon^1(s))_s$ is continuous in ϵ at $\epsilon = 0$. Thus, Property A holds, so Corollary 2 implies the result. □

Next, consider a stationary Markov strategy $\alpha^M = (\alpha_1^M, \alpha_2^M)$, where $\alpha_i^M \in [0,1]$ is the probability that player i plays action H in state i. Denote by $M^\delta(s)$ the set of payoffs in initial state s that are generated by some stationary Markov strategy that yields both players at least their minmax payoffs in both states:

$$M^\delta(s_i) \equiv \left\{v^\delta((\alpha_1^M, \alpha_2^M), s_i) : (\alpha_1^M, \alpha_2^M) \in [0,1]^2, v^\delta((\alpha_1^M, \alpha_2^M), s_j) \geq 0 \text{ for } j \in \{1,2\}\right\}.$$

The highest payoff for player 1 in $M^\delta(s_1)$ is strictly lower than in $V^{\text{Nash},1}(s_1)$. The intuition is that any stationary Markov strategy in which both players exert high effort H with positive probability is "biased" in favor of player 2: because s_1 is the initial state, player 1 incurs effort costs and player 2 gets the benefit "up front," while player 1 must wait for player 2 to reciprocate. A (non-Markov) strategy in which player 1 initially exerts low effort and after the first state transition, both players exert high effort, yields a higher payoff for player 1. In particular, we can state the following proposition.

PROPOSITION 6. *For the game in Example 6,* $\max\{v_i : (v_1, v_2) \in M^\delta(s_i)\} \leq \frac{5}{9}$ *for* $i \in \{1,2\}$ *and all* $\delta \leq 1$.

PROOF. Without loss of generality, consider $i = 1$. First, note that

$$\begin{aligned} v^\delta((\alpha_1^M, \alpha_2^M), s_1) &= \psi^\delta\big(\alpha_1^M, s_1, v^\delta((\alpha_1^M, \alpha_2^M), s_2)\big) \\ &= \frac{1}{1+\delta}(-\alpha_1^M, 3\alpha_1^M) + \frac{\delta}{1+\delta} v^\delta((\alpha_1^M, \alpha_2^M), s_2). \end{aligned}$$

Symmetrically,

$$v^\delta((\alpha_1^M,\alpha_2^M),s_2)=\frac{1}{1+\delta}(3\alpha_2^M,-\alpha_2^M)+\frac{\delta}{1+\delta}v^\delta((\alpha_1^M,\alpha_2^M),s_1).$$

Solving yields

$$v^\delta((\alpha_1^M,\alpha_2^M),s_1)=\frac{1+\delta}{1+2\delta}(-\alpha_1^M,3\alpha_1^M)+\frac{\delta}{1+2\delta}(3\alpha_2^M,-\alpha_2^M)$$
$$v^\delta((\alpha_1^M,\alpha_2^M),s_2)=\frac{1+\delta}{1+2\delta}(3\alpha_2^M,-\alpha_2^M)+\frac{\delta}{1+2\delta}(-\alpha_1^M,3\alpha_1^M).$$

Individuality rationality requires that $v^\delta((\alpha_1^M,\alpha_2^M),s)\geq 0$ for each state s. The necessary and sufficient condition is that

$$\alpha_2^M\in\left[\frac{1+\delta}{1+2\delta}\alpha_1^M,\frac{1+2\delta}{1+\delta}\alpha_1^M\right].$$

Thus,

$$\begin{aligned}v_1^\delta((\alpha_1^M,\alpha_2^M),s_1)&=\frac{1+\delta}{1+2\delta}(-\alpha_1^M)+\frac{\delta}{1+2\delta}(3\alpha_2^M)\\&\leq\max_{\alpha\in[0,1]}\left\{\frac{1+\delta}{1+2\delta}(-\alpha)+\frac{\delta}{1+2\delta}3\min\left\{1,\frac{1+2\delta}{1+\delta}\alpha\right\}\right\}\\&\leq\frac{5\delta^2+\delta-1}{4\delta^2+4\delta+1}\\&\leq\tfrac{5}{9}.\end{aligned}$$

□

Note that for any payoff $v\in V^{\text{Nash},1}(s)$, there does exist a stationary Markov strategy that delivers payoff v from initial state s. For example, the strategy $(\frac{1}{6},1)$ yields $\frac{2}{3}(-\frac{1}{6},\frac{1}{2})+\frac{1}{3}(3,-1)=(\frac{8}{9},0)$ at initial state s_1 and $\delta=1$. Consistent with Dutta's (1995) Lemma 1, any feasible payoff can be achieved by a stationary Markov strategy. The point of the preceding example is that those Markov strategies may not be individually rational after a state transition.

Appendix H: Sequential equilibrium: Example

Consider a version of Gossner and Hörner's (2010) duenna game with public monitoring: there are three players, and two actions for each player, B and C. The common payoff to players 1 and 2 (the "lovers") is 1 if they choose the same action and player 3 (the "duenna") chooses the other action, and is 0 otherwise. The duenna gets the negative of the lovers' payoff. The duenna's action is publicly observed and, with probability $\alpha\in(0,1)$, the actions of both lovers are observed. Otherwise, no information about their actions is revealed. (Note that this monitoring structure satisfies the Identifiability Condition.) Players observe their own payoffs each period. In the infinitely repeated game,

the best perfect public equilibrium payoff for players 1 and 2 is $\frac{1}{4}$, attainable from independent, uniform randomization by all players. However, it is straightforward to construct a sequential equilibrium that yields the lovers an expected payoff of $v^*(\alpha) \in (\frac{1}{4}, \frac{5}{16})$ (and the duenna a payoff of $-v^*(\alpha)$).[13]

EXAMPLE 7. There are two states and three players. In state s_2, which is absorbing, the duenna game described above is played. In initial state s_1, each player has at least two actions, and every action profile gives instantaneous payoffs $g_L \in (\frac{1}{4}, v^*(\alpha))$ to the lovers and -1 to the duenna. There is exactly one profile a^* with a positive transition rate; the transition rate for all other profiles is 0. ◊

In any PPE of the stochastic game in Example 7, play remains in state s_1 forever, yielding g_L for the lovers and -1 for the duenna. The reason is that any continuation PPE in state s_2 gives the lovers at most $\frac{1}{4} < g_L$, and each player can unilaterally prevent a transition out of state s_1. There is a sequential equilibrium of the stochastic game, however, with payoffs that Pareto dominate that PPE payoff: in state s_1, action a^* is played until the state changes (deviations in state s_2 are ignored), and in state s_2, players play the sequential equilibrium of the duenna game that yields continuation payoffs $v^*(\alpha) > g_L$ for the lovers and $-v^*(\alpha) > -1$ for the duenna.

REFERENCES

Abreu, Dilip, Paul Milgrom, and David Pearce (1991), "Information and timing in repeated partnerships." *Econometrica*, 59, 1713–1733. [153]

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1986), "Optimal cartel equilibria with imperfect monitoring." *Journal of Economic Theory*, 39, 251–269. [139]

Abreu, Dilip, David Pearce, and Ennio Stacchetti (1990), "Toward a theory of discounted repeated games with imperfect monitoring." *Econometrica*, 58, 1041–1063. [139]

Dutta, Prajit K. (1995), "A folk theorem for stochastic games." *Journal of Economic Theory*, 66, 1–32. [131, 132, 133, 148, 149, 152, 153, 171]

[13]The following strategy profile is a sequential equilibrium and it yields the lovers an expected payoff of

$$v^*(\alpha) \equiv \frac{1-\delta}{1-\delta^2}\frac{1}{4} + \frac{\delta(1-\delta)}{1-\delta^2}\frac{5-\alpha}{16} \in \left(\tfrac{1}{4}, \tfrac{5}{16}\right).$$

- In odd periods, players randomize independently and uniformly (yielding payoff $\frac{1}{4}$).
- In even periods, the duenna randomizes uniformly and the behavior of the lovers depends on the outcome in the previous period:
  - If the public signal revealed their actions in the previous period or if they got payoff 1 in the previous period, then they randomize independently and uniformly (yielding payoff $\frac{1}{4}$).
  - Otherwise, player 1 plays the same action as in the previous period and player 2 switches.

Conditional on their actions not being revealed in the previous period, the lovers get expected payoff $\frac{5}{16}$ in an even period. Their unconditional expected payoff in an even period, then, is $(1-\alpha)\frac{5}{16} + \alpha\frac{1}{4}$, so their overall payoff is $v^*(\alpha)$.

Fudenberg, Drew, David K. Levine, and Eric Maskin (1994), "The folk theorem with imperfect public information." *Econometrica*, 62, 997–1039. [132]

Fudenberg, Drew and David K. Levine (2007), "Continuous time limits of repeated games with imperfect public monitoring." *Review of Economic Dynamics*, 10, 173–192. [153]

Fudenberg, Drew and David K. Levine (2009), "Repeated games with frequent signals." *Quarterly Journal of Economics*, 124, 233–265. [153]

Fudenberg, Drew and Jean Tirole (1991), *Game Theory*. MIT Press, Cambridge, Massachusetts. [152]

Fudenberg, Drew and Yuichi Yamamoto (2011), "The folk theorem for irreducible stochastic games with imperfect public monitoring." *Journal of Economic Theory*, 146, 1664–1683. [131, 133]

Gossner, Olivier and Johannes Hörner (2010), "When is the lowest equilibrium payoff in a repeated game equal to the minmax payoff?" *Journal of Economic Theory*, 145, 63–84. [152, 171]

Hellwig, Martin and Klaus Schmidt (2002), "Discrete-time approximations of the Holmström–Milgrom Brownian-motion model of intertemporal incentive provision." *Econometrica*, 70, 2225–2264. [153]

Hörner, Johannes, Takuo Sugaya, Satoru Takahashi, and Nicolas Vieille (2011), "Recursive methods in discounted stochastic games: An algorithm for $\delta \to 1$ and a folk theorem." *Econometrica*, 79, 1277–1318. [131, 133]

Judd, Kenneth and Şevin Yeltekin (2011), "Computing equilibria of dynamic games." Working paper, Tepper School of Business, Carnegie Mellon University. [138, 153]

Kitti, Mitri (2013), "Subgame perfect equilibria in discounted stochastic games." Discussion Paper 87, Aboa Centre for Economics. [153]

Mailath, George J. and Larry Samuelson (2006), *Repeated Games and Reputations: Long-Run Relationships*. Oxford University Press, Oxford. [138]

Pęski, Marcin and Thomas Wiseman (2014), "Equilibrium payoffs in stochastic games with gradual state changes." Working paper. [153]

Stokey, Nancy, Robert Lucas Jr., and Edward Prescott (1989), *Recursive Methods in Economic Dynamics*. Harvard University Press, Cambridge, Massachusetts. [138]

Submitted 2013-4-17. Final version accepted 2014-2-19. Available online 2014-2-26.