PIER Working Paper 12-043

"A Foundation for Markov Equilibria in Infinite
Horizon Perfect Information Games"

by

V. Bhaskar, George J. Mailath, and Stephen Morris

http://ssrn.com/abstract=2171293

A Foundation for Markov Equilibria in Infinite Horizon Perfect Information Games[*]

V. Bhaskar[†], George J. Mailath[‡], and Stephen Morris[§]

October 29, 2012[¶]

Abstract

We study perfect information games with an infinite horizon played by an arbitrary number of players. This class of games includes infinitely repeated perfect information games, repeated games with asynchronous moves, games with long and short run players, games with overlapping generations of players, and canonical non-cooperative models of bargaining.

We consider two restrictions on equilibria. An equilibrium is *purifiable* if close by behavior is consistent with equilibrium when agents' payoffs at each node are perturbed additively and independently. An equilibrium has *bounded recall* if there exists K such that at most one player's strategy depends on what happened more than K periods earlier. We show that only Markov equilibria have bounded memory and are purifiable. Thus if a game has at most one long-run player, all purifiable equilibria are Markov.

1   Introduction

Repeated game theory has shown that *punishment strategies*, strategies contingent on payoff irrelevant histories, greatly expand the set of equilibrium outcomes. Yet in much applied analysis of dynamic games, researchers restrict attention to Markov equilibria, equilibria in which behavior does not depend on payoff irrelevant histories. Arguments for focussing on Markov equilibria include (i) their simplicity; (ii) their sharp predictions; (iii) their role in highlighting the key payoff relevant dynamic incentives; and (iv) their descriptive accuracy in settings where the coordination implicit in payoff irrelevant history dependence does not seem to occur. However, principled reasons for restricting attention to Markov equilibria are limited.[1]

This paper provides a foundation for Markov strategies for dynamic games with perfect information that rests on two assumptions. First, we make the restriction that all players (except possibly one) must use bounded-recall strategies, i.e., strategies that do not depend on the infinite past. Second, we require equilibrium strategies to be "purifiable," i.e., to also constitute an equilibrium of a perturbed game with independent private payoff shocks in the sense of Harsanyi (1973). Our main result is that Markov equilibria are the only bounded and purifiable equilibria.

The purifiability requirement reflects the view that our models are only an approximation of reality, and there is always some private payoff information. We make the modest requirement that there must be some continuous shock under which the equilibrium survives. The boundedness requirement is of interest for two distinct reasons. First, in many contexts, it is natural to assume that there do not exist two players who can observe the infinite past: consider, for example, games between a long-run player and a sequence of short-run players or in games with overlapping generations of players. Second, strategies that depend on what happens in the arbitrarily distant past do not seem robust to memory problems and/or noisy information. While we do not formally model the latter justification for focussing on bounded memory strategy profiles, we believe it makes them interesting objects of study.[2]

---

[1] For asynchronous choice games, Jehiel (1995) and Bhaskar and Vega-Redondo (2002) provide a rationale for Markov equilibria based on complexity costs. Maskin and Tirole (2001) discuss the notion of payoff relevance and the continuity properties of Markov equilibria; we discuss Maskin and Tirole (2001) in Section 3.2.

[2] In a different context (repeated games with imperfect public monitoring), Mailath and Morris (2002, 2006) show that strategies based on infinite recall are not "robust to private monitoring," i.e, they cease to constitute equilibrium with even an arbitrarily small amount of private noise added to public signals.
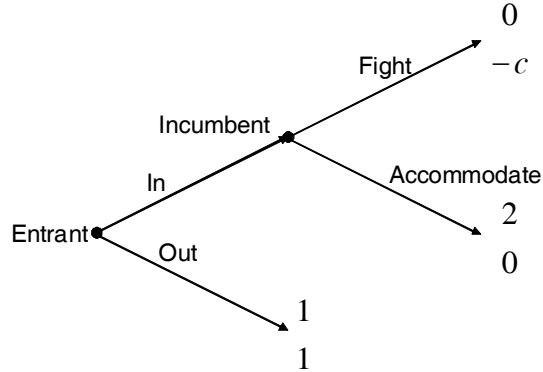
Figure 1: The stage game for the chain store. The top payoff is the payoff to the Entrant.

Our argument exploits special features of the games we study: only one player moves at a time and there is perfect information. Perfect information and the purifying payoff shocks imply that if a player conditions upon a past (payoff irrelevant) event at date t, then some future player must also condition upon this event. Such conditioning is possible in equilibrium only if the strategy profile exhibits infinite history dependence. We thus give the most general version of an argument first laid out by Bhaskar (1998) in the context of a particular (social security) overlapping generations game. This argument does not apply with simultaneous moves since two players may mutually reinforce such conditioning at the same instant, as we discuss in Section 6.

2 A Long-Run Player/Short-Run Player Example

Consider the following example of a repeated perfect information game, the chain store game, played between a long-run player and an infinite sequence of short-run players. In each period, an entrant (the short-run player) must decide whether to enter or stay out. If the entrant stays out, the stage game ends; if he enters, then the incumbent (the long-run player) must decide whether to accommodate or fight. The stage game is depicted in Figure 1.

Each entrant maximizes his stage game payoff, only observing and thus only conditioning on what happened in the previous period. The incumbent maximizes the discounted sum of payoffs, observing the entire history. The

incumbent's discount factor δ is less than but close to 1. We require equilibria to satisfy sequential rationality—each player is choosing optimally at every possible history.

Ahn (1997, Chapter 3) shows that there is no pure strategy equilibrium where entry is deterred (for generic values of the discount factor). To provide some intuition, restrict attention to stationary strategies. Since the entrant only observes the outcome of the previous period, the entrant's history is an element of $A = \{\mathsf{Out}, \mathsf{A}, \mathsf{F}\}$. Consider a trigger strategy equilibrium where the entrant enters after accommodation in the previous period, and stays out otherwise. For this to be optimal, the incumbent must play a strategy of the form: F as long as he has not played A in the previous period; A otherwise. Such a strategy is not sequentially rational, because it is not optimal to play A when A had been played in the previous period. In this case, playing A secures a payoff of zero, while a one step deviation to F earns $-(1 - \delta)c + \delta$, which is strictly positive for high enough δ.

There is however a class of mixed strategy equilibria in which entry is deterred with positive probability in each period. In any equilibrium in this class, the incumbent plays F with probability $\frac{1}{2}$, independent of history. The entrant is indifferent between In and Out at any information set, given the incumbent's strategy. He plays In with probability p at $t = 1$. At $t > 1$ he plays In with probability p after $a_{t-1} \in \{\mathsf{Out}, \mathsf{F}\}$; if $a_{t-1} = \mathsf{A}$, he plays In with probability q, where $q = p + c/[\delta(1 + c)]$. That is, the difference in entry probabilities across histories, $q - p$, is chosen to make the incumbent indifferent between accommodating and fighting. If we choose $p = 0$, then no entry takes place on the equilibrium path. Note that we have a one-dimensional manifold of equilibria in this class. In any such equilibrium, the entrant's beliefs about the incumbent's response is *identical* after the two one-period histories $a_{t-1} = \mathsf{A}$ and $a_{t-1} \in \{\mathsf{Out}, \mathsf{F}\}$. Nevertheless, the entrant plays differently.

We now establish that none of these mixed strategy equilibria can be purified if we add small shocks to the game's payoffs. So suppose that the entrant gets a payoff shock $\varepsilon \tilde{z}_1^t$ from choosing Out while the incumbent gets a payoff shock $\varepsilon \tilde{z}_2^t$ from choosing F. We suppose each $\tilde{z}_i^t$ is drawn independently across players and across time according to some known density with support $[0, 1]$. The shocks are observed only by the player making the choice at the time he is about to make it. A strategy for the entrant is

$$\rho_t : \{\mathsf{Out}, \mathsf{A}, \mathsf{F}\} \times [0, 1] \to \Delta(A_1),$$

while a strategy for the incumbent is

$$\sigma_t : A^t \times [0,1] \to \Delta(A_2)$$

(in principle, it could condition on the history of past payoff shocks, but this turns out not to matter). Note that ρ_{t+1} does not condition on what happened at $t-1$. Fix a history $h^t = (a_1, a_2, \ldots, a_t) \in A^t$ with $a_t = \mathsf{In}$ (entry at date t) and z_2^t (payoff realization for incumbent). For almost all z_2^t, the incumbent has a unique pure best response. Since ρ_{t+1} does not condition on h^{t-1},

$$\sigma_t((h^{t-1}, \mathsf{In}), z_2^t) = \sigma_t((\hat{h}^{t-1}, \mathsf{In}), z_2^t)$$

for almost all z_2^t and any $\hat{h}^{t-1} \in A^{t-1}$. So the incumbent does not condition on h^{t-1}. Since the entrant at t also has a payoff shock, it has a unique pure best response for almost all payoff shock realizations, and so

$$\rho_t(h^{t-1}, z_1^t) = \rho_t(\tilde{h}^{t-1}, z_1^t)$$

for almost all z_1^t.

We conclude that for any $\varepsilon > 0$, only equilibria in Markov strategies exist. If ε is sufficiently small, the incumbent accommodates for all realizations of his payoff shock, and therefore, with probability one. So the entrant enters with probability one. Thus, in any purifiable equilibrium of the unperturbed game, the backwards induction outcome of the stage game must be played in every period.

3 The Model

3.1 The Perfect Information Game

We consider a potentially infinite dynamic game of perfect information, Γ. The game has a recursive structure and may also have public moves by nature. The set of players is denoted by $\mathcal{N}$ and the set of states by S, both of which are countable. Only one player can move at any state, and we denote the assignment of players to states by $\iota : S \to \mathcal{N}$. This assignment induces a partition $\{S(i) \mid i \in \mathcal{N}\}$ of S, where $S(i) = \{s \in S \mid \iota(s) = i\}$ is the set of states at which i moves. Let A denote the countable set of actions available at any state; since payoffs are state dependent, it is without loss of generality to assume that the set of actions is state independent. Let $q(s'|s, a)$ denote the probability of state s' following state s when action a is

played; thus $q : S \times A \to \Delta(S)$. The initial distribution over states is given by $q(\varnothing)$. Player i has bounded flow payoff $u_i : S \times A \to \mathbb{R}$ and a discount factor $\delta_i \in [0,1)$. Total payoffs in the game are the discounted sum of flow payoffs. The dynamic game is given by $\Gamma = \left\{ S, \mathcal{N}, \iota, q, (u_i)_{i \in \mathcal{N}} \right\}$.

This formulation allows for both deterministic and stochastic finite horizons: one (or more) of the states may be absorbing, and gives all players a zero payoff.

The game starts in a state s_0 at period 0 determined by $q(\varnothing)$ and the history at period $t \geq 1$ is a sequence of states and actions, $H^t = (S \times A)^t$. Some histories may not be feasible: if after a history $h = (s_\tau, a_\tau)_{\tau=0}^t$, the state s has zero probability under $q(\cdot \mid s_t, a_t)$, then that state cannot arise after the history h. Since infeasible histories arise with zero probability and the set of all histories is countable, without loss of generality our notation often ignores the possibility of infeasible histories. Let $H^0 = \{\varnothing\}$ and $H = \cup_{t=0}^\infty H^t$; we write h for a typical element of H, $\tau(h)$ for the length of the history (i.e., $\tau(h)$ is the t for which $h \in H^t$), and $H^\infty = (S \times A)^\infty$ for the set of outcomes (infinite histories) with typical element h^∞. We sometimes write (h, s) for $(h, s_{\tau(h)}) = (s_0, a_0; s_1, a_1; \ldots, s_{\tau(h)-1}, a_{\tau(h)-1}; s_{\tau(h)})$, with the understanding that $s = s_{\tau(h)}$. Player i's payoff as a function of outcome, $U_i : H^\infty \to \mathbb{R}$, is

$$U_i \left(h^\infty \right) = U_i \left((s_t, a_t)_{t=0}^\infty \right) = (1 - \delta_i) \sum_{t=0}^\infty \delta_i^t u_i(s_t, a_t).$$

A (*behavioral*) *strategy* for player i is a mapping $b_i : H \times S(i) \to \Delta(A)$. Write B_i for the set of strategies of player i. A strategy profile $b = (b_i)_{i \in \mathcal{N}}$ can be understood as a mapping $b : H \times S \to \Delta(A)$, specifying a mixed action at every history. Write $V_i \left(b \mid h, s \right)$ for player i's expected continuation utility from the strategy profile b at the history (h, s). This value is given recursively by

$$V_i \left(b \mid h, s \right) = \sum_{a \in A} b_{\iota(s)} \left(a \mid h, s \right) \left\{ (1 - \delta_i) u_i \left(s, a \right) \right.$$

$$\left. + \delta_i \sum_{s' \in S} q \left(s' \mid s, a \right) V_i \left(b \mid (h, s, a), s' \right) \right\}.$$

We write $V_i \left(b \right) \equiv \sum q(s \mid \varnothing) V_i \left(b \mid (\varnothing, s) \right)$ for player i's ex ante utility under strategy profile b.

Definition 1 *A strategy b_i is* Markov *if for each $s \in S(i)$ and histories $h, h' \in H$ of the same length (i.e., $\tau(h) = \tau(h')$),*

$$b_i(h, s) = b_i(h', s).$$

A Markov strategy is stationary *if the two histories h and h' can be of different lengths.*

A Markov profile is eventually stationary *if there exists ℓ such that for histories h and h' where $\tau(h) \geq \ell$ and $\tau(h') \geq \ell$, for all $i \in \mathcal{N}$ and all $s \in S(i)$,*

$$b_i(h, s) = b_i(h', s).$$

Definition 2 *A strategy profile b is a subgame perfect Nash equilibrium (SPNE) if, for all $s \in S$, $h \in H$, and each $i \in \mathcal{N}$ and $b_i' \in B_i$,*

$$V_i((b_i, b_{-i}) \mid h, s) \geq V_i((b_i', b_{-i}) \mid h, s). \tag{1}$$

If b is both Markov and a SPNE, it is a Markov perfect equilibrium.

Many games fit into our general setting:

1. Repeated perfect information games. The state s tracks play in the perfect information stage game; $u_i(s, a)$ is zero whenever (s, a) results in a non-terminal node of the stage game, and is the payoff at the terminal node otherwise.

2. Perfect information games played between overlapping generations of players (Bhaskar (1998) and Muthoo and Shepsle (2010)).

3. Extensive form games between long-lived and short-lived players. Such games arise naturally in the reputation literature (e.g., Fudenberg and Levine (1989); Ahn (1997, Chapter 3)).

4. Infinitely repeated games with asynchronous moves, either with a deterministic order of moves (as in Maskin and Tirole (1987, 1988a,b), Jehiel (1995), Lagunoff and Matsui (1997) and Bhaskar and Vega-Redondo (2002)) or with a random order of moves (as in Matsui and Matsuyama (1995)).[3] In both cases, the state s is the profile of actions of players whose actions are fixed, and $u_i(s, a)$ is the stage game payoff.

---

[3]To incorporate the Poisson process of opportunities to change actions, as in Matsui and Matsuyama (1995), we would have to incorporate a richer timing structure into our model. Lagunoff and Matsui (1995) describe a class of "asynchronously repeated games" allowing this straightforward extension.

	c_1	c_2	d
c_1	$11, 11$	$6, 9$	$-20, 20$
c_2	$9, 6$	$10, 10$	$-20, 20$
d	$20, -20$	$20, -20$	$0, 0$

Figure 2: Payoffs for an augmented prisoners' dilemma.

5. Non-cooperative bargaining. In each period, a proposer makes an offer and other players decide sequentially whether to accept or reject the offer, with either deterministic order of moves (Rubinstein, 1982) or random order (Chatterjee, Dutta, Ray, and Sengupta, 1993).

Our next examples show that there are interesting Markov equilibria in perfect information games, the possibility for non-stationary Markov behavior, as well as the restrictive power of Markov.

Example 1 [An asynchronous move repeated game] Consider the augmented prisoners' dilemma illustrated in Figure 2. With asynchronous moves, player 1 moves in odd periods and player 2 in even periods (since time begins at $t = 0$, player 2 makes the first move). State and action sets are $S = A = \{c_1, c_2, d\}$ and the state encodes the action taken in the previous period (so $q(s' \mid s, a) = 1$ if $s' = a$ and 0 otherwise). Suppose the initial state is given by c_1.

There are two stationary pure strategy Markov equilibria: Let $b^* : S \to A$ be the Markov strategy given by $b^*(s) = s$. It is straightforward to verify that b^* is a perfect equilibrium for $\delta \in \left[\frac{1}{2}, \frac{20}{31}\right]$. Let $b^\dagger : S \to A$ be the Markov strategy given by $b^\dagger(c_1) = b^\dagger(c_2) = c_2$ and $b^\dagger(d) = d$. It is straightforward to verify that $b^\dagger$ is a perfect equilibrium for $\delta \in \left[\frac{1}{2}, \frac{2}{3}\right]$.

Finally, denote by $b^\alpha : S \to \Delta(A)$ the Markov strategy given by $b^\alpha(c_1) = \alpha \circ c_1 + (1 - \alpha) \circ c_2$, $b^\alpha(c_2) = c_2$, and $b^\alpha(d) = d$. Suppose it is player i's turn. At (h, c_2), the payoff from following b^α is

$$V_i(b^\alpha \mid h, c_2) = 10. \tag{2}$$

At (h, c_1), the payoff from choosing c_1, and then following b^α, is

$$(1 - \delta)11 + \delta\alpha\{(1 - \delta)11 + \delta V_i(b^\alpha \mid (h, c_1, c_1), c_1)\}$$
$$+ \delta(1 - \alpha)\{(1 - \delta)6 + \delta V_i(b^\alpha \mid (h, c_1, c_1), c_2)\}, \tag{3}$$

while the payoff from choosing c_2, and then following b^α, is

$$(1 - \delta)9 + \delta V_i(b^\alpha \mid (h, c_1), c_2)\}. \tag{4}$$

coalition	1's payoff	2's payoff	3's payoff
$\{1,2\}$	9	3	0
$\{2,3\}$	0	9	3
$\{1,3\}$	3	0	9

Figure 3: Payoffs to players in each pairwise coalition for Example 2. The excluded player receives a payoff of 0.

In order for player i to be willing to randomize, (3) must equal (4), with this common value being $V_i(b^\alpha \mid h, c_1)$. Since $V_i(b^\alpha \mid (h, c_1), c_2) = 10$, (4) implies $V_i(b^\alpha \mid h, c_1) = 9 + \delta$, and solving (3) for α yields

$$\alpha = \frac{(4\delta - 2)}{(5 - \delta)\delta}. \tag{5}$$

This is a well defined probability for $\delta \geq \frac{1}{2}$. Moreover, b^α, for α satisfying (5), is a Markov equilibrium for $\delta \in \left[\frac{1}{2}, \frac{2}{3}\right]$.

For any time t, the nonstationary Markov strategy specifying for periods before or at t, play according to b^*, and for periods after t, play according to b^α, for α satisfying (5), is a Markov perfect equilibrium for $\delta \in (\frac{1}{2}, \frac{2}{3})$.

An outcome path in which player 1 always plays c_1 and player 2 always plays c_2 is a subgame perfect equilibrium outcome path for sufficiently patient players – this is supported by permanent d after any deviation. However, this is not the outcome path of any Markov perfect equilibrium.
★

The nonstationary equilibrium in Example 1 is, in a sense, trivial, since it is eventually stationary. As we will see, this type of non-stationarity is not purifiable.

The next example illustrates nontrivial nonstationarity.

Example 2 [Pure Markov requires nonstationarity] This coalition formation game is a simplification of Livshits (2002). There are three players. In the initial period, a player i is selected randomly and uniformly to propose a coalition with one other player j, who can accept or reject. If j accepts, the game is over with payoffs given in Figure 3. If j rejects, play proceeds to the next period, with a new proposer randomly selected. If no coalition is formed, all players receive a payoff of 0. Note that there is a cycle in

preferences: 1 prefers the coalition with 2, who prefers the coalition with 3, who in turn prefers the coalition with 1.

For $\delta < 3/4$, there is a unique Markov perfect equilibrium and this equilibrium is stationary and in pure strategies with every proposal immediately accepted. For $\delta > 3/4$, there is no Markov perfect equilibrium in stationary pure strategies. There is a stationary Markov perfect equilibrium in mixed strategies, with the responder randomizing between acceptance and rejection, accepting with probability $3(1 - \delta)/\delta$.

If $3/4 < \delta < \sqrt{3/4}$, there are two nonstationary pure strategy Markov equilibria. In one, offers are accepted in odd periods and rejected in even periods, while in the other offers are accepted in even periods and rejected in odd. (For larger values of δ, pure strategy Markov equilibria display longer cycles.)

An example of a non-Markov perfect equilibrium when $3/4 < \delta < \sqrt{3/4}$ is the following: in the first period, if 1 is selected, then 1 chooses 3, who accepts (the specification of play after 3 rejects can follow any Markov perfect equilibrium). If 1 chooses 2, then 2 rejects, with play then following the Markov perfect equilibrium with acceptance in odd periods (recall that the first period is period 0, so that after 2's rejection, the next period is odd and so there is immediate acceptance). If 2 or 3 are selected, then play follows the Markov perfect equilibrium with acceptance in even periods.
★

3.2   Markov equilibria and payoff relevance

In Definition 1, we have taken the state space and the corresponding notion of a *Markov* strategy as a primitive. The restriction to Markov strategies is often motivated by the desire to restrict behavior to only depend on the payoff relevant aspects of history. By construction, since payoffs at any date only depend upon the state and upon the action taken, our states capture everything that is payoff-relevant in the history of the game (but may capture more). Maskin and Tirole (2001) describe a natural definition of the coarsest possible notion of payoff-relevant (or Markov) state.[4] While in general, two states s and s' reached in the same period in our state space S may be payoff equivalent, a sufficient condition ruling this out is that for

---

[4]Loosely, Maskin and Tirole (2001) use payoff equivalence to induce a partition $\mathcal{H}^*$ over histories of the same length. A *Markov* state is an element of this partition. $\mathcal{H}^*$ is the coarsest partition with the property that for every profile measurable with respect to that partition, each player has a best response measurable with respect to that partition.

every pair $s, s' \in S(i)$, $u_i(s, a)$ is not an affine transformation of $u_i(s', a)$ (Mailath and Samuelson, 2006, Proposition 5.6.2).[5]

Since small changes in payoffs can destroy payoff equivalence, the set of Markov states as defined by Maskin and Tirole (2001) does not behave continuously with respect to the payoffs u_i, i.e. the set of payoffs for which the set of Markov states is given by S is *not* closed).[6] This failure of continuity may be viewed as a criticism of the concept. In particular, suppose that $\hat{\Gamma}$ denotes a game with the same extensive form as Γ, where payoffs are allowed to depend nontrivially on histories, so that payoffs are given by $\hat{u}_i : H \times S \times A \to \mathbb{R}$. Apart from the specification of payoffs, Γ and $\hat{\Gamma}$ coincide. Then, for generic assignment of payoffs, every SPNE of $\hat{\Gamma}$ will be Markov, and Markov has no restrictive power in $\hat{\Gamma}$. Moreover, this will be true even if $|\hat{u}_i(h, s, a) - u_i(s, a)|$ is small.

However, as Maskin and Tirole (2001, p. 194) note, one of the philosophical considerations embodied by Markov perfect equilibrium is that "minors causes should have minor effects."[7] We interpret this as a requirement that if two histories are *almost* payoff irrelevant, then behavior at these two histories should be almost the same. This immediately yields upper hemi-continuity:[8]

Lemma 1 *Suppose $\{\hat{\Gamma}^m\}_m$ is a sequence of perturbations of Γ satisfying,*

---

[5]Maskin and Tirole (2001) use cardinal preferences to determine payoff equivalence, and hence the presence of affine transformations. If there are no strictly dominated strategies, this corresponds to saying that games are equivalent if the better-response relation is the same for all conjectures in both games. This is, in general, a stronger requirement than the equivalence of best responses in both games (Morris and Ui, 2004).

[6]Roughly speaking, two histories correspond to different Maskin and Tirole (2001) states if a collection of equalities (corresponding to payoff equivalence) fail. Clearly, equalities can fail along a convergent sequence and yet hold in the limit. Consequently, two histories will not be payoff equivalent (i.e., correspond to different states) along a sequence of convergent payoffs, and yet be payoff equivalent (correspond to the *same* state) in the limit .

[7]This is not simply a technical requirement. Markov in the absence of some kind of continuity requirement can be consistent with behavior that is clearly not in the spirit of Markov, see Proposition 2 and the following discussion in Mailath and Samuelson (2001).

[8]The failure of the Markov perfect equilibrium correspondence to be upper hemi-continuous in Maskin and Tirole (2001) (see their footnote 11) is due to their notion of Markov state, and is related to our earlier observation that the set of payoffs for which the set of Markov states in a particular period is given by the states in S reachable in that period is *not* closed. Maskin and Tirole (2001, Section 4) shows that for finite games, generically (in the space of payoffs of Γ), Markov equilibria of Γ (in the Maskin-Tirole sense) can be approximated by Markov equilibria of near-by $\hat{\Gamma}$.

for all $h \in H$, for all $s \in S$, and all $a \in A$,

$$|\hat{u}_i^m(h, s, a) - u_i(s, a)| \to 0 \quad as \quad m \to \infty.$$

Suppose b^m is a SPNE of $\hat{\Gamma}^m$ and $b^m \to b^*$ as $m \to \infty$. If, for all $h, h' \in H$ of the same length, for all $s \in S$,

$$\left| b_{\iota(s)}^m(h, s) - b_{\iota(s)}^m(h', s) \right| \to 0 \quad as \quad m \to \infty, \tag{6}$$

then b^* is a Markov perfect equilibrium of Γ.

Proof. The profile b^* is trivially a SPNE of Γ (this is just upper-hemicontinuity of the SPNE correspondence). Fix a state s and two histories of the same length, h and h'. For all $\varepsilon > 0$, for sufficiently large m, the convergence of b^m to b^* implies $\left| b_{\iota(s)}^m(h, s) - b_{\iota(s)}^*(h, s) \right| < \varepsilon$ and $\left| b_{\iota(s)}^m(h', s) - b_{\iota(s)}^*(h', s) \right| < \varepsilon$, and (6) implies $\left| b_{\iota(s)}^m(h, s) - b_{\iota(s)}^m(h', s) \right| < \varepsilon$. The triangle inequality then yields $\left| b_{\iota(s)}^*(h, s) - b_{\iota(s)}^*(h', s) \right| < 3\varepsilon$. Because this inequality holds for all $\varepsilon > 0$, we have $\left| b_{\iota(s)}^*(h, s) - b_{\iota(s)}^*(h', s) \right| = 0$, and so b^* is Markov. ∎

Note that Lemma 1 shows that b^* is Markov in the sense of Definition 1, not in the sense of Maskin and Tirole (2001).

The observation that for generic payoff assignments, every subgame perfect equilibrium of $\hat{\Gamma}$ will be Markov does *not* imply that every non-Markov perfect equilibrium of Γ can be approximated by a (Markov) equilibrium in any near-by $\hat{\Gamma}$. We present an example in the appendix.

3.3   The Game with Payoff Shocks

We now allow for the payoffs in the underlying game to be perturbed, as in Harsanyi (1973). We require that the payoff shocks respect the recursive payoff structure of the infinite horizon game, i.e., to not depend upon history except via the state: Let Z be a full dimensional compact subset of $\mathbb{R}^{|A|}$ and write $\Delta^*(Z)$ for the set of measures with support Z generated by strictly positive densities.[9] At each history (h, s), a payoff shock $z_i \in Z$ is drawn according to $\mu_i^s \in \Delta^*(Z)$.[10] The payoff shocks are independently distributed

---

[9]Our analysis only requires that the support be in Z, but notation is considerably simplified by assuming Z is the support.

[10]Since both A and S are countable, there are only countably many histories, and so we do not require a continuum of independent random variables.

across players and histories. We write $\mu^s := \times_i \mu_i^s$ for the product measure on $Z^{\mathcal{N}}$. If the player moving at state s chooses action a, i's payoff is augmented by εz_i^a, where $\varepsilon > 0$. Thus, players' stage payoffs in the perturbed game depend only on the current state, action, and payoff shock (s, a, z), and are given by

$$\tilde{u}_i(s, a, z) = u_i(s, a) + \varepsilon z_i^a.$$

We denote the perturbed game by $\Gamma(\varepsilon, \mu)$.

To describe strategies, we first describe players' information more precisely. Write $\mathbf{z}_i(h, s)$ for the sequence of payoff shocks realized for player i along (h, s), and $\tilde{z}_i(h, s)$ for player i's current shock (thus $\tilde{z}_i(h, s)$ is the last element of the sequence $\mathbf{z}_i(h, s)$); and $\mathbf{z}(h, s)$ for the sequence of payoff shock profiles realized for all players up to (h, s). Hopefully without confusion, we suppress the arguments (h, s), leaving the context to clarify the dimensionality of the various vectors, so that for (h, s),

$$\tilde{z}_i \in Z, \ \mathbf{z}_i \in Z^{\tau(h)+1}, \text{ and } \mathbf{z} \in (Z^{\tau(h)+1})^{\mathcal{N}}.$$

A *behavior strategy* for player i in the perturbed game, $\tilde{b}_i$, specifies player i's mixed action $\tilde{b}_i(h, s, \mathbf{z}_i)$, at every history (h, s) with $s \in S(i)$ and for every realization of i's payoff shocks $\mathbf{z}_i$. The set of all behavior strategies for player i is denoted $\widetilde{B}_i$.

The definition of sequential rationality requires us to have notation to cover unreached information sets. A belief assessment for player i specifies, for every feasible history $h \in H$ and $s \in S(i)$, a belief

$$\pi_i^{h,s} \in \Delta\left(\prod_{j \neq i} Z^{\tau(h)+1}\right) \tag{7}$$

over the payoff shocks $\mathbf{z}_{-i}$ that have been observed by other players at history (h, s). Note that, as suggested by the structure of the perturbed game, we require that these beliefs are independent of player i's private payoff shocks, $\mathbf{z}_i$; beyond this requirement, we impose no further restrictions (such as independence of payoff shocks across players or periods)—see Remark 1.

Player i's "value" function is recursively given by, for a given strategy profile $\tilde{b}$,

$$\tilde{V}_i(\tilde{b} \mid h, s, \mathbf{z}) = \sum_{a \in A} \tilde{b}_{\iota(s)}(a \mid h, s, \mathbf{z}_{\iota(s)}) \left[(1 - \delta_i)\tilde{u}_i(s, a, \tilde{z}_i) \right.$$

$$\left. + \delta_i \sum_{s' \in S} q(s' \mid s, a) \int \tilde{V}_i(\tilde{b} \mid (h, s, a), s', (\mathbf{z}, z))\mu^{s'}(dz) \right].$$

13

Since player i does not know all the coordinates of $\mathbf{z}$, player i's expected payoff from the profile $\tilde{b}$ is given by

$$\int \tilde{V}_i(\tilde{b} \mid h, s, (\mathbf{z}_i, \mathbf{z}_{-i})) \ \pi_i^{h,s}(d\mathbf{z}_{-i}). \tag{8}$$

Definition 3 *Strategy $\tilde{b}_i$ is a* sequential best response *to $(\tilde{b}_{-i}, \pi_i)$, if for each $h \in H$, $s \in S(i)$, $\mathbf{z}_i \in Z^{\tau(h)+1}$, and $\tilde{b}_i' \in \widetilde{B}_i$,*

$$\int \tilde{V}_i((\tilde{b}_i, \tilde{b}_{-i}) \mid h, s, ((\mathbf{z}_i, \mathbf{z}_{-i})) \ \pi_i^{h,s}(d\mathbf{z}_{-i})$$

$$\geq \int \tilde{V}_i((\tilde{b}_i', \tilde{b}_{-i}) \mid h, s, (\mathbf{z}_i, \mathbf{z}_{-i})) \ \pi_i^{h,s}(d\mathbf{z}_{-i}).$$

Strategy $\tilde{b}_i$ is a sequential best response to $\tilde{b}_{-i}$ if strategy $\tilde{b}_i$ is a sequential best response to $(\tilde{b}_{-i}, \pi_i)$ for some π_i.

Definition 4 *A strategy $\tilde{b}_i$ is* shock history independent *if for all $h \in H$, $s \in S(i)$, and shock histories $\mathbf{z}_i, \mathbf{z}_i' \in Z^{\tau(h)+1}$,*

$$\tilde{b}_i(h, s, \mathbf{z}_i) = \tilde{b}_i(h, s, \mathbf{z}_i')$$

when $\tilde{z}_i = \tilde{z}_i' = z$, for almost all $z \in Z$.

Lemma 2 *If $\tilde{b}_i$ is a sequential best response to any $\tilde{b}_{-i}$, then $\tilde{b}_i$ is a shock history independent strategy.*

Proof. Fix a player i, $h \in H$, $s \in S(i)$, and payoff shock history $\mathbf{z}$. Player i's *next period* expected continuation payoff under $\tilde{b}$ from choosing action a this period, $\mathbf{V}_i(a, \tilde{b}_{-i}, \pi_i \mid h, s)$, is given by

$$\sum_{s'} q(s' \mid s, a) \iint \max_{\tilde{b}_i} \tilde{V}_i(\tilde{b}_i, \tilde{b}_{-i} \mid (h, s, a), s', \mathbf{z}, z') \ \mu^{s'}(dz')\pi_i^{h,s,a,s'}(d\mathbf{z}_{-i}).$$

Since $\tilde{b}_{-i}$ and $\pi_i^{h,s,a,s'}$ do not depend on player i's shocks, the maximization implies that $\mathbf{V}_i(a, \tilde{b}_{-i}, \pi_i \mid h, s)$ also does not depend on those shocks. Thus, his total utility is

$$(1 - \delta_i)[u_i(s, a) + \varepsilon \tilde{z}_i^a] + \delta_i \mathbf{V}_i(a, \tilde{b}_{-i}, \pi_i \mid h, s).$$

Since Z has full dimension and μ^s is absolutely continuous, player i can only be indifferent between two actions a and a' for a zero measure set of

14

$z \in Z$. For other z, there is a unique best response, and so it is shock history independent. ∎

A shock history independent strategy (ignoring realization of z of measure 0) can be written as

$$\tilde{b}_i : H \times S(i) \times Z \rightarrow \Delta\left(A\right).$$

If *all* players are following shock history independent strategies, we can recursively define value functions for a given strategy profile $\tilde{b}$ that do not depend on *any* payoff shock realizations:

$$V_i^*(\tilde{b} \mid h, s) = \int \sum_{a \in A} \tilde{b}_{\iota(s)}(a \mid h, s, \tilde{z}_{\iota(s)}) \left[(1 - \delta_i)\tilde{u}_i(s, a, z) \right.$$

$$\left. + \delta_i \sum_{s' \in S} q(s' \mid s, a)V_i^*(\tilde{b} \mid (h, s, a), s') \right] \mu^s(dz). \quad (9)$$

It is now immediate from Lemma 2 that beliefs over unreached information sets are essentially irrelevant in the notion of sequential best responses, because, while behavior can in principle depend upon prior payoff shocks, optimal behavior does not.

Lemma 3 *A profile $\tilde{b}$ is a profile of mutual sequential best responses if, and only if, for all i, $\tilde{b}_i$ is shock history independent, and for each $h \in H$, $s \in S(i)$ and $\tilde{b}_i' \in \tilde{B}_i$,*

$$V_i^*((\tilde{b}_i, \tilde{b}_{-i}) \mid h, s) \geq V_i^*((\tilde{b}_i', \tilde{b}_{-i}) \mid h, s). \quad (10)$$

Remark 1 Because the perturbed game has a continuum of possible payoff shocks in each period, and players may have sequences of unreached information sets, there is no standard solution concept that we may appeal to. Our notion of sequential best response is very weak (not even requiring that the beliefs respect Bayes' rule *on the path of play*). The only requirement is that each player's beliefs over other players' payoff shocks be independent of his own shocks. For information sets on the path of play, this requirement is implied by Bayes' rule. Tremble-based refinements imply such a requirement at all information sets, though they may imply additional restrictions across information sets. This requirement is not implied by the notion of "weak perfect Bayesian equilibrium" from Mas-Colell, Whinston, and Green (1995), where no restrictions are placed on beliefs off the equilibrium path:

this would allow players to have different beliefs about past payoff shocks depending on their realized current payoff realization.

However, Lemma 3 implies that once we impose mutuality of sequential best responses, any additional restrictions have no restrictive power. It is worth noting why no belief assessment $\pi_i^{h,s}$ appears either in the description of V_i^*, (9), or in Lemma 3: Player i's expected payoff from the profile $\tilde{b}$, given in (8), is the expectation over past payoff shocks of other players, $\mathbf{z}_{-i}(h, s)$, as well as all future payoff shocks. Critically, in this expectation, as implied by the structure of the perturbed game, it is assumed that all future shocks are distributed according to μ, independent of all past shocks. ♦

Given Lemma 3 and the discussion in Remark 1, the following definition is natural:

Definition 5 *A* perfect Bayesian equilibrium *is a profile of mutual sequential best responses.*

Even though the private payoff shocks are drawn from a continuum, each period's decision can be viewed as a finite dimensional one, and so existence of PBE is guaranteed for any μ and ε for standard reasons (a sketch of the proof is in the appendix).

Lemma 4 *The perturbed game* $\Gamma(\varepsilon, \mu)$ *has a PBE for all* $\varepsilon > 0$ *and* μ *with* $\mu^s \in \Delta^*(Z^{\mathcal{N}})$ *for all* $s \in S$.

The definition of *Markov* shock history independent strategies naturally generalizes that for the unperturbed game: a strategy $\tilde{b}_i$ is *Markov* if for each $s \in S(i)$, for almost all $z \in Z$, and histories $h, h' \in H$ with $\tau(h) = \tau(h')$,

$$\tilde{b}_i(h, s, z) = \tilde{b}_i(h', s, z).$$

Definition 6 *A shock history independent strategy* $\tilde{b}_i$ *has* K-recall *if for each* $s \in S(i)$, *histories* $h, h' \in H$ *satisfying* $\tau(h) = \tau(h') = t$, *and almost all* $z \in Z$,
$$\tilde{b}_i(h, s, z) = \tilde{b}_i(h', s, z)$$

whenever $(s_k, a_k)_{k=t-K}^{t-1} = (s_k, a_k)_{k=t-K}^{t-1}$. *A strategy* b_i *has* infinite recall *if it does not have* K-recall *for any* K. *A* Markov *strategy is a* 0-recall *strategy (there being no restriction on* h *and* h'*).*

A K-recall *strategy is* stationary *if the two histories can be of different lengths.*

The following is the key result of the paper.

Lemma 5 *If $\tilde{b}_i$ is a sequential best response to $\tilde{b}_{-i}$ and does not have K-recall, then for some $j \neq i$, $\tilde{b}_j$ does not have $(K+1)$-recall.*

Proof. If $\tilde{b}_i$ does not have K-recall, then there exist h and h' with $\tau(h) = \tau(h') = t \geq K$ and $s \in S(i)$

$$(s_k, a_k)_{k=t-K}^{t-1} = (s'_k, a'_k)_{k=t-K}^{t-1}$$

and

$$\tilde{b}_i(h, s, z) \neq \tilde{b}_i(h', s, z) \tag{11}$$

for a positive measure of z.

Suppose that $\tilde{b}_j$ has $(K+1)$-recall for each $j \neq i$. Since histories (h, s, a) and (h', s, a) agree in the last $K+1$ periods, player i's continuation value from playing action a at (h, s) and at (h', s) is the same, for all s':

$$V_i^*((\tilde{b}_i, \tilde{b}_{-i}) \mid (h, s, a), s') = V_i^*((\tilde{b}_i, \tilde{b}_{-i}) \mid (h', s, a), s').$$

Hence, player i's total expected utility from choosing action a at *either* (h, s) or (h', s) is

$$(1 - \delta_i)\tilde{u}_i(s, a, z) + \delta_i \sum_{s' \in S} q(s' \mid s, a) V_i^*((\tilde{b}_i, \tilde{b}_{-i}) \mid (h, s, a), s').$$

For almost all z, there will be a unique a maximizing this expression, contradicting our premise (11). ∎

Corollary 1 *If $\tilde{b}$ is a perfect Bayesian equilibrium of the perturbed game, then either $\tilde{b}$ is Markov or at least two players have infinite recall.*

Example 3 [The Chain Store] It is an implication of Lemma 5 that the one-period recall mixed strategy equilibria from the chain store game of Section 2 cannot be approximated by any equilibrium of $\Gamma(\varepsilon, \mu)$ when the short-lived player is restricted to a bounded recall strategy. On the other hand, any such equilibrium can be approximated if we allow both players to play infinite recall strategies.

We now show this explicitly for the mixed equilibrium where the entrant chooses Out with probability one after Out or F in the previous period and randomizes with probability $c/[\delta(1+c)]$ after In, while the incumbent

randomizes with equal probability on A and F after any history. The construction is similar to one described in Bhaskar, Mailath, and Morris (2008, Section 5). For any history $h \in H \equiv \cup_t \{\mathsf{Out}, \mathsf{A}, \mathsf{F}\}^t$, let $k(h)$ be the number of periods of A at the end of h (so that if $k(h) = 0$, A was not played in the last period). Define $H(\kappa) \equiv \{h \in H : k(h) = \kappa\}$; the collection $\{H(\kappa) : \kappa \in \mathbb{Z}_+\}$ is a partition of H. Our strategy profile is measurable with respect to this partition. We denote by $\rho_\kappa : [0,1] \rightarrow \{\mathsf{Out}, \mathsf{In}\}$ the strategy of the entrant, and by $\sigma_\kappa : [0,1] \rightarrow \{\mathsf{A}, \mathsf{F}\}$ the strategy of the incumbent at any history $h \in H(\kappa)$. The payoff shocks for the entrant (incumbent, respectively) are drawn from $[0,1]$ according to the distribution function F_1 (F_2, resp.).

We begin by setting $\rho_0(z_1) = \mathsf{Out}$ for all $z \in [0,1)$, which requires the entrant face a probability of A of no more than $(1 + \varepsilon)/2$. We accordingly set $\sigma_0(z_2) = \mathsf{A}$ for all $z_2 \leq z_2^0$, where z_2^0 satisfies $F_2(z_2^0) = (1 + \varepsilon)/2$, implying a probability of A of exactly $(1 + \varepsilon)/2$ (which can be made arbitrarily close to $1/2$ by choosing ε small). The construction proceeds recursively. Let z_i^κ be the marginal type for player i at a history $h \in H(\kappa)$, so that the probability of In is $F_1(z_1^\kappa)$ and the probability of A is $F_2(z_2^\kappa)$ at such a history. Indifference for the entrant at h requires

$$1 + \varepsilon z_1^\kappa = F_2(z_2^\kappa)2,$$

yielding an equation determining the incumbent's marginal type as a function of the entrant's:

$$z_2^\kappa = F_2^{-1}[\tfrac{(1+\varepsilon z_1^\kappa)}{2}]. \tag{12}$$

Turning to optimality for the incumbent, let $V^\kappa(\mathsf{A})$ be the value to the incumbent from playing A at a history $h \in H(\kappa)$ after the entrant chooses In (which is current payoff shock independent), and $V^\kappa(\mathsf{F}, z_2)$ be the value to the incumbent from playing F at a history $h \in H(\kappa)$ after the entrant chooses In and payoff shock z_2. Then, for all $\kappa \geq 0$,

$$V^\kappa(\mathsf{F}, z_2) = (1 - \delta)(-c + \varepsilon z_2) + \delta W^0$$

and

$$V^\kappa(\mathsf{A}) = \delta F_1(z_1^{\kappa+1})V^{\kappa+1} + \delta(1 - F_1(z_1^{\kappa+1}))W^0,$$

where $W^0 = 1$ is the continuation value to the incumbent of the entrant choosing Out, and V^κ is the continuation value to the incumbent at a history $h \in H(\kappa)$ after the entrant has chosen In, but before the incumbent's payoff

shock is realized:

$$
\begin{aligned}
V^\kappa &= F_2(z_2^\kappa)V^\kappa(\mathsf{A}) + (1 - F_2(z_2^\kappa))E[V^\kappa(\mathsf{F}, z_2) \mid z_2 > z_2^\kappa] \\
&= F_2(z_2^\kappa)V^\kappa(\mathsf{F}, z_2^\kappa) + (1 - F_2(z_2^\kappa))E[V^\kappa(\mathsf{F}, z_2) \mid z_2 > z_2^\kappa] \\
&= V^\kappa(\mathsf{F}, z_2^\kappa) + (1 - F_2(z_2^\kappa))E[V^\kappa(\mathsf{F}, z_2) - V^\kappa(\mathsf{F}, z_2^\kappa) \mid z_2 > z_2^\kappa] \\
&= (1 - \delta)(-c + \varepsilon z_2^\kappa) + \delta + \varepsilon(1 - F_2(z_2^\kappa))(1 - \delta)E[z_2 - z_2^\kappa \mid z_2 > z_2^\kappa].
\end{aligned}
$$

Observe that we now have an expression for $V^{\kappa+1}$ in terms of $z_2^{\kappa+1}$, which using (12), gives $V^{\kappa+1}$ in terms of $z_1^{\kappa+1}$. We thus have an implicit difference equation, since given z_1^κ (and so z_2^κ from (12)), the requirement that z_2^κ be indifferent, $V^\kappa(\mathsf{F}, z_2^\kappa) = V^\kappa(\mathsf{A})$, implicitly determines $z_1^{\kappa+1}$, and so on. Observe that for $\varepsilon = 0$, for all $\kappa \geq 1$, $F_1(z_1^\kappa) = c/[\delta(c+1)]$ and $F_2(z_2^\kappa) = 1/2$, and so for ε small, the implied behavior in $\Gamma(\varepsilon, \mu)$ is close to the one-period recall mixed strategy equilibrium.

<div align="right">★</div>

3.4 Purification in the Games of Perfect Information

We now consider the purifiability of equilibria in the unperturbed game, while maintaining our restriction to bounded recall strategies. *Purification* has several meanings in the literature (see Morris (2008)). One question asked in the literature is when can we guarantee that every equilibrium is essentially pure by adding noise to payoffs (e.g., Radner and Rosenthal (1982))? It is trivially true that our shocks ensure that there is an essentially pure equilibrium (we build in enough independence to guarantee that this is the case) and that there are no equilibria with nontrivial mixing.

We follow Harsanyi (1973) in being interested in the relation between equilibria of the unperturbed game and equilibria of the perturbed game. But our definition of *purifiability* is very weak: we require only that there exists a sequence of equilibria of a sequence of perturbed games that converge to the desired behavior.

Fix a strategy profile b of the unperturbed game. We say that a sequence of current shock strategies $\tilde{b}_i^k$ in the perturbed game converges to a strategy b_i in the unperturbed game if expected behavior (taking expectations over shocks) converges, i.e., for each $h \in H$, $s \in S_i$ and $a \in A$,

$$
\int \tilde{b}_i^k(a \mid h, s, z)\mu^s(dz) \to b_i(a \mid h, s) \tag{13}
$$

Definition 7 *The strategy profile b is K-recall purifiable if there exist $\mu_i^k :$ $S \to \Delta^*(Z)$ and $\varepsilon^k \to 0$, such that there is a sequence of profiles $\{\tilde{b}^k\}_{k=1}^{\infty}$ converging to b, with $\tilde{b}^k$ a perfect Bayesian equilibrium of the perturbed game $\Gamma(\mu^k, \varepsilon^k)$ for each k, and $\tilde{b}_i^k$ having K-recall for all (but perhaps one) i.*

Since the supporting sequence of private payoff shocks is allowed to depend on the strategy profile b, and the distribution μ^k is itself indexed by k, this notion of purifiability is almost the weakest possible.[11] Our notion crucially maintains the recursive payoff structure of the infinite horizon game (in particular, we require that the payoff shocks are intertemporally independent). Allowing for intertemporally dependent payoff shocks violates the spirit of our analysis.

The strongest notion of purification, closer to the spirit of Harsanyi (1973), would require that *for all* fixed private shock distributions $\mu : S \to \Delta^*(Z)$, and for all sequences $\varepsilon^k \to 0$, there is an equilibrium $\tilde{b}^k$ of the perturbed game $\Gamma(\mu, \varepsilon^k)$ with $\tilde{b}^k$ converging to b. We refer to this notion as *Harsanyi* purification. Clearly, if a profile is Harsanyi purifiable, then it is purifiable.

Our main result is the following:

Proposition 1 *If b is a K-recall purifiable SPNE, then b is Markov. Moreover, if b is also eventually stationary, then it is stationary.*

Proof. The first assertion is an immediate implication of Corollary 1.

To prove the second assertion, suppose that b is eventually stationary. Then there exists a date t, a state s, and two histories, h and h', of length t such that the profile is stationary after t, and $b_{\iota(s)}(h, s) \neq b_{\iota(s)}(h', s)$. But then in near-by games with payoff shocks, there is a positive set of payoff shocks for which the optimal behavior at (h, s) differs from that at (h', s), which is impossible because the continuation play is identical after (h, s, a) and (h', s, a). ∎

We immediately have the following corollary. (An argument similar to that given in the proof of the second assertion of Proposition 1 above proves the second claim in this corollary.)

Corollary 2 *If Γ is the infinite repetition of a finite perfect information game with a unique backward induction equilibrium, then the only K-recall*

---

[11] It is also worth noting that we only require pointwise convergence in (13). For infinite horizon games, we may ask for uniform (in h) convergence, as is done in the positive result (Theorem 3) in Bhaskar, Mailath, and Morris (2008). Negative results are clearly stronger with pointwise convergence.

purifiable equilibrium is the infinite repetition of the backward induction equilibrium of Γ. If Γ has multiple backward induction equilibria, the only K-recall purifiable equilibria has infinite recall is an infinite sequence of history-independent specifications of a backward induction equilibrium.

4 Existence of Markov, purifiable Markov, and Harsanyi-purifiable Markov equilibria.

The unperturbed game is a particularly simple stochastic game of perfect information, since the set of players, set of states, and set of actions are all countable. The existence of stationary Markov equilibria in behavior strategies follows from Escobar (2008, Corollary 14).

Turning to purifiability, for finite players and states, a direct argument shows that every Markov perfect equilibrium is purifiable, even in the absence of any regularity assumptions or arguments. The proof is complicated by two features: future payoffs (including the contributions from the payoff shocks) affect current values and so the returns from different state transitions, and perturbing actions results in both perturbed flow payoffs and perturbed transitions in the dynamic.

Lemma 6 *Suppose there is a finite number of players, N, and S is finite. Every stationary Markov perfect equilibrium in the unperturbed game is purifiable.*

Proof. Let $b : S \to \Delta(A)$ be a stationary Markov equilibrium of the unperturbed game. Write $A^*(s)$ for the set of actions that are (possibly weak) best responses at state s. Fix a sequence of Markov strategy profiles $\{b^k\}$ converging to b (i.e., $b^k(a \mid s) \to b(a \mid s)$ for each a and s) with support of $b^k(\cdot \mid s)$ equal to $A^*(s)$. Recall that Z, a full dimensional compact subset of $\mathbb{R}^A$, is the support of each player's payoff shocks. Without loss of generality, we may assume the $\mathbf{0}$-vector is in the interior of Z. For each action $a \in A^*(s)$, write $Z^*(a, s)$ for the collection of payoff shock profiles favoring action a among those played with positive probability, i.e.,

$$Z^*(a, s) = \left\{ z \in Z^N \mid z^a_{\iota(s)} > z^{a'}_{\iota(s)} \text{ for all } a' \in A^*(s), \, a' \neq a \right\}.$$

Note that the union of the closures of the sets $(Z^*(a, s))_{a \in A^*(s)}$ is Z^N, i.e.,

$$\bigcup_{a \in A^*(s)} \text{cl}(Z^*(a, s)) = Z^N.$$

Write $a^*(z, s)$ for the action satisfying $z \in Z^*(a, s)$. (Such an action is unique for almost all z, choose arbitrarily when it is not unique.) For a Markov strategy b', we can write $V_i(b' \mid s)$ for the expected payoff to player i from b' in the unperturbed game starting in state s.

Two matrices, $\mathbb{Q}$ and $\mathbb{Q}^k$, describe the state transitions under b and b^k respectively. Their ss'-th elements are given by

$$[\mathbb{Q}]_{ss'} = \sum_a b_{\iota(s)}(a \mid s)q(s' \mid s, a)$$

and

$$[\mathbb{Q}^k]_{ss'} = \sum_a b^k_{\iota(s)}(a \mid s)q(s' \mid s, a).$$

For player i, the flow payoffs under b and b^k are described by the two $|S|$-dimensional vectors, $\mathbf{u}_i$ and $\mathbf{u}_i^k$, with s-th elements

$$[\mathbf{u}_i]_s = (1 - \delta_i) \sum_a b_{\iota(s)}(a \mid s)u_i(s, a)$$

and

$$[\mathbf{u}_i^k]_s = (1 - \delta_i) \sum_a b^k_{\iota(s)}(a \mid s)u_i(s, a).$$

It is immediate that the vector of values $\mathbf{V}_i(b) = [V_i(b \mid s)]_{s \in S}$ can be calculated as $(\mathbb{I} - \delta_i\mathbb{Q})^{-1}\mathbf{u}_i$, where $\mathbb{I}$ is the $|S|$-dimensional identity matrix. For each k, set

$$\Delta_i^k \equiv (\mathbb{I} - \delta_i\mathbb{Q}^k)(\mathbb{I} - \delta_i\mathbb{Q})^{-1}\mathbf{u}_i - \mathbf{u}_i^k.$$

Since $b^k \to b$, $\mathbb{Q}^k \to \mathbb{Q}$, $\mathbf{u}_i^k \to \mathbf{u}_i$, and so $\lim_k \Delta_i^k(s) = 0$ for all s. Set $\varepsilon^k = \max\left\{k^{-1}, \sqrt{\max_{i,s}|\Delta_i^k(s)|}\right\}$, so that $\varepsilon^k \to 0$. We first determine μ^k on $Z^*(a, s)$ be setting

$$\mu^k(Z^*(a, s)) = b^k_{\iota(s)}(a \mid s).$$

For k sufficiently large, we can complete the specification of $\mu^k(s) \in \Delta(Z^N)$ by choosing a strictly positive density on each $Z^*(a, s)$ so that[12]

$$\sum_{a \in A^*(s)} b^k_{\iota(s)}(a \mid s)E_{\mu^k(s)}[z_i^a \mid Z^*(a, s)] = \frac{\Delta_i^k(s)}{(1 - \delta_i)\varepsilon^k}. \tag{14}$$

---

[12]Since $\mathbf{0}$ is in the interior of Z, there are both positive and negative values of z_i^a in $Z^*(a, s)$, and since for large k, the right side of (14) is close to 0, such densities will exist.

Now consider the strategy profile $\tilde{b}^k$ in the perturbed game $\Gamma\left(\mu^k, \varepsilon^k\right)$ given by

$$\tilde{b}^k_{\iota(s)}(a|s, z) = \begin{cases} 1, & \text{if } z \in Z^*(a, s), \\ 0, & \text{if } z \in Z^*(a', s), \text{ for any } a' \in A^*(s), a' \neq a. \end{cases}$$

The specification of $\tilde{b}$ on the (zero measure) boundaries of the sets $Z^*(a, s)$ is irrelevant. By construction, expected behavior under $\tilde{b}^k$ equals b^k.

It remains to verify that, for sufficiently large k, $\tilde{b}^k$ is a PBE of the perturbed game (i.e., each $\tilde{b}^k_i$ is a sequential best response to $\tilde{b}^k_{-i}$). The expected payoff from $\tilde{b}^k$ evaluated before the realization of the state s payoff shock is denoted $\tilde{V}^k_i(\tilde{b}^k \mid s)$. Expressed as the vector $\tilde{\mathbf{V}}_i(\tilde{b}) = [\tilde{V}^k_i(\tilde{b}^k \mid s)]_s$, it satisfies the equation

$$\tilde{\mathbf{V}}^k_i(\tilde{b}^k) = \mathbf{u}^k_i + \Delta^k_i + \delta_i \mathbb{Q}^k \tilde{\mathbf{V}}^k_i(\tilde{b}^k).$$

This implies

$$\begin{aligned} (\mathbb{I} - \delta_i \mathbb{Q}^k)\tilde{\mathbf{V}}^k_i(\tilde{b}^k) &= \mathbf{u}^k_i + \Delta^k_i \\ &= (\mathbb{I} - \delta_i \mathbb{Q}^k)(\mathbb{I} - \delta_i \mathbb{Q})^{-1}\mathbf{u}_i, \end{aligned}$$

and so

$$\tilde{\mathbf{V}}^k_i(\tilde{b}^k) = (\mathbb{I} - \delta_i \mathbb{Q})^{-1}\mathbf{u}_i = \mathbf{V}_i(b).$$

It is now immediate that each $\tilde{b}^k_i$ is a sequential best response to $\tilde{b}^k_{-i}$ for sufficiently large k, since the payoff from an action $a \in A$, conditional on the realization z_i in the perturbed game is

$$(1 - \delta_i)u_i(s, a) + \varepsilon^k(1 - \delta_i)z_i + \delta_i \sum_{s'} q(s' \mid s, a)\tilde{V}_i(\tilde{b}^k \mid s')$$

$$= (1 - \delta_i)u_i(s, a) + \varepsilon^k(1 - \delta_i)z_i + \delta_i \sum_{s'} q(s' \mid s, a)V_i(b \mid s').$$

(Since Z is bounded, even large realizations of z_i cannot reverse the strict suboptimality of actions not in $A^*(s)$ for sufficiently small ε^k). ∎

We conjecture that with additional regularity assumptions, Markov equilibria will be Harsanyi purifiable. Doraszelski and Escobar (2008) provided conditions for the Harsanyi purifiability of Markov equilibria. While the class of games they study do not encompass those in the present paper, it seems likely that appropriately modifying their method to account for the perfect information structure of the unperturbed game will yield Harsanyi purification of all Markov equilibria for generic games (where the genericity notion will need to reflect our game structure).

5 Exogenous Memory Bounds

The analysis of Section 3 does *not* require players to recall the entire history (as required by the definition of perfect information). In particular, suppose each player i has restricted memory in the sense that he or she only knows the previous T_i periods (because he or she was not in the game, or forgets earlier periods). If the player knows calendar time, then

$$b_i : \mathbb{Z}_+ \times H^{T_i} \times S(i) \to \Delta(A). \tag{15}$$

If the player does not know calendar time, then a behavior strategy is a mapping

$$b_i : H^{T_i} \times S(i) \to \Delta(A). \tag{16}$$

(In both (15) and (16), if player i could be asked to move in a period $t < T_i$, the domain of b_i should include shorter histories). The analysis in Section 3 applies essentially without change to either notion of behavior strategy. (While agents will now have beliefs over actions and states in periods before their memory, these beliefs play no role in the analysis.)

This observation immediately yields the following result.

Proposition 2 *Suppose there is at most one long-lived player and all the other players are finitely-lived, with a uniform bound on the length of their lives. Suppose there is a uniform bound on the length of history any short-lived player learns at birth. Then, the only purifiable SPNE are Markov. Moreover, if the short-lived players do not know calendar time (so that the strategies satisfy (16)), then the only purifiable SPNE are stationary Markov.*

6 Discussion

Our results do not extend to games where more than one player moves at a time, e.g. repeated synchronous move games. Mailath and Morris (2002) and Mailath and Olszewski (2008) give examples of strict, and hence purifiable, finite recall strategy profiles. In this context, one might conjecture the weaker result that purifiability would rule out the "belief-free" strategies recently introduced by Piccione (2002) and Ely and Välimäki (2002). Bhaskar, Mailath, and Morris (2008) show that the one period recall strategies of Ely and Välimäki (2002) are not purifiable via one period recall strategies in the perturbed game; however, they are purifiable via infinite recall strategies. The purifiability of such belief free strategies via finite recall strategies remains an open question.

A Appendix

A.1 Some non-Markov equilibria cannot be approximated.

We return to the chain store game of Section 2, denoted Γ, and modify it as follows. Assume, as before, that the short run player only observes the outcome of the stage game in the previous period, but now allow the payoffs in the stage game to depend upon this outcome.[13] Specifically, we assume that in $\hat{\Gamma}$, the entrant's payoff at each terminal node of the stage game is augmented by $\eta_{a'a}^{E}$, and that of the incumbent by $\eta_{a'a}^{I}$, where $a' \in \{\mathsf{Out}, \mathsf{A}, \mathsf{F}\}$ is the stage game terminal node reached in the previous period and a is the terminal node in the current period. Assume that all the values $\eta_{a'a}^{E}$ and $\eta_{a'a}^{I}$ are close to 0. Then, by arguments similar to those in Section 2, there cannot be a pure strategy equilibrium where entry is deterred in each period. Consider a stationary mixed strategy equilibrium where the incumbent fights with probability $\sigma(\mathsf{A}) \in (0, 1)$ when $a_{t-1} = \mathsf{A}$. This implies the indifference condition

$$\delta\tilde{V}(\mathsf{A}) = -(1-\delta)(c + \eta_{AA} - \eta_{AF}) + \delta\tilde{V}(\mathsf{F}),$$

where $\tilde{V}(a)$ denotes his value function when the outcome in the previous period was a.

The payoff difference between accommodation and fighting when $a_{t-1} = \mathsf{Out}$ is given by

$$\delta\tilde{V}(\mathsf{A}) + (1-\delta)(c + \eta_{OUTA} - \eta_{OUTF}) - \delta\tilde{V}(\mathsf{F}).$$

Restricting attention to generic values of η, we see that

$$\sigma(\mathsf{Out}) = \begin{cases} 0 \text{ if } \eta_{OUTA} - \eta_{OUTF} > \eta_{AA} - \eta_{AF} \\ 1 \text{ if } \eta_{OUTA} - \eta_{OUTF} < \eta_{AA} - \eta_{AF}. \end{cases}$$

For Out to be an absorbing state, a necessary condition is that $\eta_{OUTA} - \eta_{OUTF} < \eta_{AA} - \eta_{AF}$. In other words, the *outcome* of the mixed equilibrium of Γ where entry is deterred in every period can be approximated in $\hat{\Gamma}$ only under this condition, and not otherwise. Furthermore, the outcomes of *none* of the continuum of equilibria of Γ where entry occurs with positive probability after all one period histories can be approximated in $\hat{\Gamma}$, for generic values of η. This follows from $\sigma(\mathsf{Out})$ and $\sigma(\mathsf{A})$ can both be interior only if $\eta_{OUTA} - \eta_{OUTF} = \eta_{AA} - \eta_{AF}$.

---

[13]We assume that outcomes in previous periods do not affect payoffs in order to ensure that the short run player is playing a complete information game.

A.2 Sketch of proof of Lemma 4.

Consider first perturbed games with a finite horizon and state space (and so finite number of players). From Lemma 2, all sequential best responses must be shock history independent. Moreover, the action specified by sequential best response in a period has a finite-dimensional characterization: at each history-state, (h, s), for any two actions a and a', there is a half-space $\mathbb{H}$ such that player $\iota(s)$ prefer a to a' if the payoff shock is in $\mathbb{H} \cap Z$, and a' to a if not. In other words, player $\iota(s)$'s ranking of actions at history (h, s) has a finite dimensional description (since $\mathbb{H}$ does). Let $\Gamma^{\eta}(\varepsilon, \mu)$ denote the finite "doubly" perturbed game in which a choice of action a leads to a with probability $1 - \eta$ and a uniform distribution over A with probability η. This can be viewed as a finite game, and so it has a Nash equilibrium. Moreover, extracting a convergent subsequence from the sequence of equilibria as $\eta \to 0$ yields a PBE of the (finite) perturbed game $\Gamma(\varepsilon, \mu)$.

Finally, consider the general perturbed game $\Gamma(\varepsilon, \mu)$ with infinite horizon and/or countably infinite state spaces. Since such a game is continuous at infinity and a similar property is satisfied with respect to the state space, the game can be approximated by a sequence of perturbed games with finite horizons and finite state spaces. A PBE of $\Gamma(\varepsilon, \mu)$ can then obtained as the limit of a convergent sequence of PBE of the finite games.

References

AHN, I. (1997): "Three Essays on Repeated Games Without Perfect Information," Ph.D. thesis, University of Pennsylvania.

BHASKAR, V. (1998): "Informational Constraints and the Overlapping Generations Model: Folk and Anti-Folk Theorems," *Review of Economic Studies*, 65(1), 135–149.

BHASKAR, V., G. J. MAILATH, AND S. MORRIS (2008): "Purification in the Infinitely-Repeated Prisoners' Dilemma," *Review of Economic Dynamics*, 11(3), 515–528.

BHASKAR, V., AND F. VEGA-REDONDO (2002): "Asynchronous Choice and Markov Equilibria," *Journal of Economic Theory*, 103(2), 334–350.

CHATTERJEE, K., B. DUTTA, D. RAY, AND K. SENGUPTA (1993): "A Non-Cooperative Theory of Coalitional Bargaining," *Review of Economic Studies*, 60(2), 463–477.

DORASZELSKI, U., AND J. ESCOBAR (2008): "A Theory of Regular Markov Perfect Equilibria in Dynamic Stochastic Games: Genericity, Stability, and Purification," Discussion paper, Harvard University and Stanford University.

ELY, J. C., AND J. VÄLIMÄKI (2002): "A Robust Folk Theorem for the Prisoner's Dilemma," *Journal of Economic Theory*, 102(1), 84–105.

ESCOBAR, J. (2008): "Existence of Pure and Behavior Strategy Stationary Markov Equilibrium in Dynamic Stochastic Games," Discussion paper, Stanford University.

FUDENBERG, D., AND D. K. LEVINE (1989): "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57(4), 759–778.

HARSANYI, J. C. (1973): "Games with Randomly Disturbed Payoffs: A New Rationale for Mixed-Strategy Equilibrium Points," *International Journal of Game Theory*, 2(1), 1–23.

JEHIEL, P. (1995): "Limited Horizon Forecast in Repeated Alternate Games," *Journal of Economic Theory*, 67, 497–519.

LAGUNOFF, R., AND A. MATSUI (1995): "An "Anti-Folk Theorem" for a Class of Asynchronously Repeated Games," Georgetown University.

——— (1997): "Asynchronous Choice in Repeated Coordination Games," *Econometrica*, 6(5), 1467–1477.

LIVSHITS, I. (2002): "On Non-Existence of Pure Strategy Markov Perfect Equilibrium," *Economics Letters*, 76(3), 393–396.

MAILATH, G. J., AND S. MORRIS (2002): "Repeated Games with Almost-Public Monitoring," *Journal of Economic Theory*, 102(1), 189–228.

——— (2006): "Coordination Failure in Repeated Games with Almost-Public Monitoring," *Theoretical Economics*, 1, 311–340.

MAILATH, G. J., AND W. OLSZEWSKI (2008): "Folk Theorems with Bounded Recall under (Almost) Perfect Monitoring," PIER Working Paper 08-027, University of Pennsylvania.

MAILATH, G. J., AND L. SAMUELSON (2001): "Who Wants a Good Reputation?," *Review of Economic Studies*, 68(2), 415–441.

———— (2006): *Repeated Games and Reputations: Long-Run Relationships.* Oxford University Press, New York, NY.

Mas-Colell, A., M. D. Whinston, and J. Green (1995): *Microeconomic Theory.* Oxford University Press, New York, NY.

Maskin, E., and J. Tirole (1987): "A Theory of Dynamic Oligopoly, III: Cournot Competition," *European Economic Review*, 31(4), 947–968.

———— (1988a): "A Theory of Dynamic Oligopoly I: Overview with Quantity Competition and Large Fixed Costs," *Econometrica*, 56(3), 549–569.

———— (1988b): "A Theory of Dynamic Oligopoly, II: Price Competition, Kinked Demand Curves, and Edgeworth Cycles," *Econometrica*, 56(3), 571–599.

———— (2001): "Markov Perfect Equilibrium I. Observable Actions," *Journal of Economic Theory*, 100(2), 191–219.

Matsui, A., and K. Matsuyama (1995): "An Approach to Equilibrium Selection," *Journal of Economic Theory*, 65(2), 415–434.

Morris, S. (2008): "Purification," in *The New Palgrave Dictionary of Economics Second Edition*, ed. by S. Durlauf, and L. Blume, pp. 779–782. Macmillan Palgrave.

Morris, S., and T. Ui (2004): "Best Response Equivalence," *Games and Economic Behavior*, 49(2), 260–287.

Muthoo, A., and K. Shepsle (2010): "Information, Institutions and Constitutional Arrangements," *Public Choice*, 144(1–2), 1–36.

Piccione, M. (2002): "The Repeated Prisoner's Dilemma with Imperfect Private Monitoring," *Journal of Economic Theory*, 102(1), 70–83.

Radner, R., and R. Rosenthal (1982): "Private Information and Pure-Strategy Equilibria," *Mathematics of Operations Research*, 7(3), 401–409.

Rubinstein, A. (1982): "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50(1), 97–109.