Asynchronous Choice and Markov Equilibria: Theoretical Foundations and Applications

V. Bhaskar
Department of Economics
University of Essex
Colchester CO4 3SQ, UK

Fernando Vega-Redondo
Facultad de Económicas and
Instituto Valenciano de Investigaciones Económicas
Universidad de Alicante

March 1998
Revised: August, 1998

Abstract

This paper provides a theoretical foundation for Markov (perfect) equilibria in repeated games with asynchronous moves that is based on memory costs. We show that if players incur a "complexity cost" which depends on the memory length required by their strategies, then any strategy which is iteratively undominated is Markovian. Thus, every Nash or perfect equilibrium is Markovian as well. We also provide a dynamic learning rationale for this conclusion. Our result has interesting implications for repeated asynchronous choice games where the stage game is of common interest. If players are sufficiently patient, iterated dominance ensures repeated play of the efficient stage-game equilibrium if this equilibrium satisfies a risk-related condition that, in a 2×2 game, is equivalent to risk-dominance.

1 Introduction

Markov equilibria occupy an important position in the theory of dynamic and stochastic games. A large literature on a variety of topics ranging from industrial organization (e.g. Beggs and Klemperer (1992)) to renewable resources (e.g. Levhari and Mirman (1980)) restricts attention to Markov equilibria — players are restricted to strategies which condition only on the payoff relevant "state" variables, and are not allowed to condition upon payoff irrelevant events. There has however been little theoretical justification for such a restriction.[1] Furthermore, in the case of standard repeated games, where the "state" is constant over time (the simultaneous action stage game is fixed), Markov equilibria are trivial and uninteresting, since they consist of repeated play of the Nash equilibrium of the stage game. This contrasts sharply with the "folk theorems" for these games (e.g. Fudenberg and Maskin (1986)) which show that a wide range of alternative behavior (and payoffs) may arise in repeated game equilibria, under relatively mild conditions.

The first interesting example of Markov equilibria in a repeated game context is due to Maskin & Tirole (1988*a*, 1988*b*), who studied a variety of models of oligopolistic competition. They showed that if firms take their decisions in an *asynchronous* manner and alternate in changing actions, the Markov equilibria of the resulting game are non-trivial. For example, in the price-setting context, phenomena such as the kinked demand curve and the Edgeworth cycle may arise in Markov equilibrium, leading to firms' profits in equilibrium bounded well above Bertrand profits. This contrast between Markov equilibrium outcomes of synchronous and asynchronous games is striking. With asynchronous moves, a richer pattern of equilibrium (Markov) behavior becomes possible since a player may react to her opponents' currently fixed actions, that are payoff-relevant. Further developments of this approach include Eaton and Engers (1990) and De Fraja (1993).[2]

The present paper provides a theoretical foundation for Markov equilibria of repeated games with asynchronous moves that is based on (essentially arbitrary) memory costs. We consider two-player repeated games with discounting where players move in alternate periods. Players may condition their actions upon payoff irrelevant past events, but we assume that such

[1] A recent attempt at such a justification may be found in Maskin and Tirole (1997), to be discussed below.

[2] Jéhiel (1995) also analyzes alternating-move games, but from a "limited-forecast" perspective.

conditioning is costly. Specifically, we suppose that players' preferences over alternative configurations respond both to payoffs and memory requirements, i.e. they are increasing in payoffs for identical memory requirements, but decreasing in these requirements for equal induced payoffs. In principle, this allows for either an scenario where complexity costs are conmeasurable with payoffs or, alternatively, the oft-considered context where complexity costs are lexicographically less important than stage-game payoffs. We shall focus in the first scenario as our leading model, although we also analyze the latter formulation to test of robustness of our conclusions.

The setup described defines a repeated game where, given a certain strategy profile, a player's overall payoff depends upon two factors: the discounted sum of stage-game payoffs and memory costs. In this context, we show that any strategy of a player that survives a process of the iterated elimination of dominated strategies must be a Markov strategy. This result applies whether we focus on the strategic form of the repeated game or perform the elimination process at each separate information set in the extensive form game. This implies that every rationalizable strategy of the game (whether in its strategic or extensive-form variants – cf. Bernheim (1984) and Pearce (1984)) is Markovian. A fortiori, therefore, it follows that every Nash or perfect equilibrium of the repeated game with memory costs must be Markovian as well. Since iterated dominance is a weak requirement of rationality, this also allows us to provide a relatively strong justification for Markov strategies from a social-learning perspective. In particular, it follows from an adaptation of well-known results in the learning literature that any monotonic (evolutionary) system formalizing a general process of social learning will weed out all non-Markovian strategies in the long run.

It is worth noting at this point that we obtain similar results regardless of whether we require optimality only along the "equilibrium path" or after each information set (i.e. including those which are not reached at equilibrium). In either case, iteratively undominated strategies are Markovian. Hence our approach yields robust results. This may be contrasted with the pioneering work of Abreu and Rubinstein (1988), who introduce complexity considerations in a repeated game context with *simultaneous* moves. They focus on the Nash equilibrium concept, a certain strategy taken to be preferred to an alternative more complex one if both strategies yield the same payoff against the opponent's strategy on the *equilibrium* path. They showed that this is enough to reduce quite substantially the wide range of Nash equilibrium payoffs typically resulting from standard "folk theorems". However, Kalai and Neme (1992) subsequently showed that just "a little perfection" (i.e. the requirement of equilibrium behavior at histories reached after only one deviation) is sufficient to restore the usual folk-theorem conclusions in this context.

The conclusions described are warranted only in repeated games where players move asynchronously. With synchronous moves, it is easy to see that memory considerations are not sufficient to rule out non-Markov equilibria. In this sense, our analysis may be interpreted as providing theoretical justification for Markov equilibria precisely in those cases where these equilibria may produce interesting behavior (recall our former discussion).

To underscore the latter point, our second objective of this paper is to identify a wide enough family of games where the Markov condition displays significant implications. As a first step in this direction, we focus on contexts where the stage game is an arbitrary one of common interest (i.e. with an action profile that Pareto dominates all others). The sharpest conclusions occur when each player has only two actions (i.e. the 2×2 case). There, we find that if the efficient action is (weakly) risk dominant for at least one player (in the sense of Harsanyi–Selten (1988)) and players are sufficiently patient, every (Markov) strategy profile that survives iterated extensive-form dominance has players choosing the efficient action except possibly in the first two periods. Thus, of course, the same must happen in every Perfect equilibrium of the repeated game. In contrast, efficiency is not ensured at equilibrium if the efficient action is risky for both players. For arbitrary (i.e. larger) games of common interest, analogous conclusions are obtained but a generalization of risk dominance beyond the 2×2 case is required.

Our latter selection results are related to recent work by Lagunoff and Matsui (1997). These authors focus on *pure* coordination games (where players obtain *identical* payoffs at every action profile) and show that, if players are patient enough, the only equilibrium payoff of the repeated game is the efficient one. Their result depends critically on the fact that a pure coordination game has non-generic payoffs, as demonstrated by a recent paper by Dutta (1995). Specifically, Dutta shows that, in *any* general stochastic game where a full-dimensionality condition on the set of attainable payoffs is satisfied, any individually-rational payoff can be supported at some equilibrium. Of course, Dutta's folk theorem applies to any infinitely repeated game with asynchronous moves, and hence to the repeated common-interest games considered in the present paper if memory considerations were absent. Thus, as outlined above, our analysis demonstrates that memory considerations, arbitrarily large or small, can have powerful effects in reducing the multiplicity of equilibria in *generic* scenarios.

The rest of the paper is organized as follows. Section 2 sets out our basic model of asynchronous repeated games. Section 3 has our main result inducing Markov strategies and Markov equilibria. Section 4 explores a general model of social learning that leads to Markov behavior. Section 5 undertakes an explicit consideration of lexicographic memory costs. Section 6 turns to our selection results when the stage game has common-interests. Section 7 concludes.

2 The model

For expositional simplicity, we restrict our formal discussion to the case of two player interaction. Time (i.e. the stage of play) is indexed discretely, $t = 0, 1, 2, ...$ At every t, each player takes an action $s_i^t \in S_i$. Short-run (or instantaneous) payoffs are given by stationary functions $\pi_i : S_1 \times S_2 \to \mathbb{R}$, indicating the payoff $\pi_i^t \equiv \pi_i(s_1^t, s_2^t)$ attained by each player i at t.

At $t = 0$, players choose their actions s_i^0 $(i = 1, 2)$ simultaneously. Thereafter, they revise their actions in alternation: player 1 at odd periods $t = 1, 3, 5, ...$ and player 2 at even periods $t = 2, 4, 6, ...$ Once the game reaches stage t, its prior history is given by the specification of past play $[(s_1^0, s_2^0), (s_1^1, s_2^1), ..., (s_1^{t-1}, s_2^{t-1})]$. This list of play contains redundancies derived from the fact that each player can revise her action only every two periods (consecutive strategy profiles can only differ in one of its two components). Therefore, it will be convenient to rely on a more compact description of histories based on the single specification, for each $t \geq 1$, of the action adopted by the player who is unable to revise it then. As explained above, this action defines the state of the system, since it is the only payoff-relevant information derived from prior history. Formally, it may be defined as follows:

$$\begin{aligned} \omega^t &= s_1^t \quad \text{if } t \text{ is even} \\ &= s_2^t \quad \text{if } t \text{ is odd.} \end{aligned}$$

If, for notational simplicity, we make $\omega^0 = s_1^0$, any t-long history $(t \geq 1)$ may be fully specified through the list $h^t = (\omega^\tau)_{\tau=0}^t$. The set of all such histories may be partitioned into two subsets, $\hat{H}_1$ and $\hat{H}_2$, corresponding to the set of histories h^t $(t \geq 1)$, where either it is the turn of player 1 (when t is odd) or 2 (when t is even) to revise her action. Denoting by h^0 the "empty history" prevailing at $t = 0$, the full set of histories after which either player 1 or 2 moves is respectively given by $H_1 = \hat{H}_1 \cup \{h^0\}$ and $H_2 = \hat{H}_2 \cup \{h^0\}$.

In the context described, a general strategy for player i is given by a mapping $f_i : H_i \to \Delta(S_i)$ specifying the probability $f_i(h^t)(s)$ with which player i chooses each action $s_i \in S_i$ after every possible history h^t such that it is her turn to move. Let F_i denote the set of strategies of player i. As explained, we are interested in evaluating the memory requirements of different strategies. To formalize this idea, we consider a sequence of progressively richer subset of strategies for each player as follows. First, we define the zero-memory strategies:

$$F_i^0 \equiv \left\{ \begin{array}{c} f_i \in F_i : \left[h^t = (\omega^\tau)_{\tau=0}^t,\ \tilde{h}^{t'} = (\tilde{\omega}^\tau)_{\tau=0}^{t'} \in H_i,\ \omega^t = \tilde{\omega}^{t'}\right] \\ \Rightarrow f_i(h^t) = f_i(\tilde{h}^{t'}) \end{array} \right\}.$$

Verbally, F_i^0 stands for the strategies of player i which are only responsive to the current state of the system, i.e. the fixed action to which her opponent is committed from the previous period or empty (initial) history. As customary, they are labelled *Markov strategies.*

As player i considers tailoring her action to (payoff-irrelevant) history of progressively more protracted length, the following strategy subsets obtain for each $k = 1, 2, ...$:

$$F_i^k \equiv \left\{f_i \in F_i : \left[h^t,\, \tilde{h}^{t'} \in H_i,\ \omega^{t-s} = \tilde{\omega}^{t'-s},\ s = 0, 1, ...k\right] \Rightarrow f_i(h^t) = f_i(\tilde{h}^{t'})\right\}. \tag{1}$$

For each $k = 1, 2, ...$, the respective strategy set F_i^k allows player i's action to depend on histories of length k. In particular, a strategy belonging to this set does not permit the player's action to depend on history when the last k observations are identical. If a certain strategy f_i needs exactly k-period memory to be carried out (i.e. $f_i \in F_i^k \backslash F_i^{k-1}$), we shall denote $m(f_i) \equiv k$. On the other hand, if $f_i \in F_i \backslash \bigcup_{k=0}^{\infty} F_i^k$, we make $m(f_i) = \infty$.

Given any strategy pair $f = (f_1, f_2)$, we can uniquely define a probability measure over resulting action paths. Denote by $\sigma^t(f) \in \Delta(S_1) \times \Delta(S_2)$ the probability vector over players' actions at t induced by strategy profile f. This allows us to define the expected discounted flow of stage payoffs for any player i as follows:

$$V_i(f) = (1 - \delta) \sum_{t=0}^{\infty} \delta^t\, \mathbf{E}\left[\pi_i(\sigma^t(f))\right],$$

where $\mathbf{E}[\cdot]$ stands for the expectation operator.

We are now ready to specify players' preferences over strategy profiles in $F = F_1 \times F_2$, a typical element of which is denoted by $f = (f_1, f_2)$. There are three dimensions which we must take into account in specifying these preferences. The first is the discounted sum of stage game payoffs which accrue to the player along the path induced by f. The second is the memory requirement of player i's strategy, $m(f_i)$. The final dimension is the expected discounted sum of stage game payoffs induced by f after arbitrary histories, i.e. histories that may be reached only if one or both players deviate from f. We shall define two different classes of preference relations, labelled Ξ^N and Ξ^P, incorporating these dimensions in different ways.

The class Ξ^N incorporates only the first two dimensions, and does not take into account payoffs after histories which are off the path induced by f. This approach is hence similar to that adopted by Abreu and Rubinstein (1988).

The second class of preference relations, Ξ^P, also takes into account repeated game payoffs which arise after deviations from f. In particular, preferences belonging to this class do not require that memory considerations are infinitely more important than considerations of subgame perfection. This is akin to the approach of Kalai and Neme (1992).

Player i's binary preference relation $\succcurlyeq_i$ on the set F is assumed reflexive, transitive and complete. In addition, it will be assumed to belong either to the class Ξ^N or to the class Ξ^P. Formally, if it belongs to Ξ^N, it must verify the following two axioms:

(N) Suppose that $f, \tilde{f} \in F$ satisfy *both*:
(i) $V_i(f) \geq V_i(\tilde{f})$;
(ii) $m(\tilde{f}_i) \geq m(f_i)$.
Then, $f \succcurlyeq_i \tilde{f}$. If, furthermore,(i) or/and (ii) apply strictly, $f \succ_i \tilde{f}$. [3]

(C) There exists some $v_i \in \mathbb{N}$ such that:

$$[f_i \in F_i\,,\, m(f_i) \geq v_i] \Rightarrow \left[\begin{array}{c} \exists \hat{f}_i \in F_i : \forall f_j \in F_j (j \neq i), \\ (\hat{f}_i, f_j) \succ_i (f_i, f_j) \end{array} \right].$$

Axiom (N) requires little explanation: if a strategy profile f is no worse than another profile $\tilde{f}$, given a history h^t on both of the following counts:

[3]As customary, $f \succ_i \tilde{f}$ means: $f \succcurlyeq_i \tilde{f}$ and $\neg(\tilde{f} \succcurlyeq_i f)$.

(i) discounted payoffs; (ii) memory costs, then f is weakly preferred to $\tilde{f}$; furthermore, if f is also strictly better than $\tilde{f}$ in one of those two respects, it is also strictly preferred.

On the other hand, Axiom (C) states that there exists some sufficiently high memory threshold v_i for each player, such that any strategy with a larger memory requirement is strictly dominated. This condition will be met if, for example, the marginal memory cost of an additional period of memory is strictly positive and bounded away from zero. However, it is incompatible with those lexicographic preferences for which memory costs are "infinitesimally less important" than stage payoffs. This case may be addressed by a slight variation of our approach, as described in Section 5.

Note that our axioms are quite weak and, in particular, do not completely specify the preference relation, i.e. they are consistent with a wide range of possible players' preferences. For a simple example, consider a memory cost function $c_i : F_i \to \mathbb{R}$, where $c_i(f_i) = \zeta(m(f_i))$ for some increasing unbounded real function $\zeta(\cdot)$. Then, for some $\lambda \in (0,1)$, we may define each preference $\succcurlyeq_i$ as follows:

$$\forall f,\, \tilde{f} \in F, \quad f \succcurlyeq_i \tilde{f} \Leftrightarrow \lambda V_i(f) + (1-\lambda)c_i(f_i) \geq \lambda V_i(\tilde{f}) + (1-\lambda)\, c_i(\tilde{f}_i).$$

The second class of preferences, Ξ^P, allow us incorporate considerations of perfection, i.e. the requirement of optimality after *every* possible history. Abusing previous notation, denote by $V_i(f, h^t)$ the continuation payoff earned by player i after history h^t when the strategy profile is f. Preferences in Ξ^P continue to satisfy (C) but, instead of (N), they are assumed to verify the following alternative axiom:

(P) Suppose that $f,\, \tilde{f} \in F$ satisfy *both*:

(i) $V_i(f, h^t) \geq V_i(\tilde{f}, h^t)$, $\forall h^t \in H$;

(ii) $m(\tilde{f}_i) \geq m(f_i)$.

Then, $f \succcurlyeq_i \tilde{f}$. If, furthermore, at least one of the following applies:

(a) $\exists\, h^t \in H : V_i(f, h^t) > V_i(\tilde{f}, h^t)$;

(b) $m(\tilde{f}_i) > m(f_i)$

then, $f \succ_i \tilde{f}$

The essence of axiom (P) is that a player can be ensured to prefer one strategy profile to another only if the former does better in payoff terms than the latter after all histories, including those which are not reached when the strategy profile is played. In other words, a player may prefer to use a strategy with a longer memory requirement even if there is lower memory strategy which does as well in the absence of any deviations, provided the former does better in the event of some deviations.

Let $\Xi = \Xi^N \cup \Xi^N$. The generality of our main results derive from the fact that they apply to *any* preference relation $\succcurlyeq_i$ in the class Ξ and, moreover, only rely on dominance criteria.. Specifically, it will be enough to focus on those strategy profiles that survive an iterative elimination of dominated strategies in terms of some preference $\succcurlyeq_i$ in Ξ. This implies, in particular, that the resulting strategies are only required to be rationalizable within Ξ, in the sense of Bernheim (1984) and Pearce (1984).

Of course, since our selection results only require rationalizability, they also hold if we use the standard Nash equilibrium concept for any pair of players' preferences $\succcurlyeq_i$ in the class Ξ. However, observe that if these preferences belong to the class Ξ^P, the induced equilibrium embodies notions of perfection, whereas it does not if preferences belong to the class Ξ^N. To emphasize this distinction, the corresponding equilibrium will be labelled Perfect* Equilibrium or Nash* Equilibrium when the class of preferences is restricted accordingly.

Definition 1 *Given any pair of preferences* $(\succcurlyeq_1, \succcurlyeq_2) \in \Xi^2$, *a strategy profile* $f^* = (f_1^*, f_2^*) \in F$ *is an* Equilibrium *if, for each* $i = 1, 2$ *and any* $f_i \in F_i$, *we have* $f^* \succcurlyeq_i (f_i, f_j^*)$, $j \neq i$. *If each* $\succcurlyeq_i \in \Xi^N$ $(i = 1, 2)$, f^* *is called a* Nash* Equilibrium (N*E). *If each* $\succcurlyeq_i \in \Xi^P$ $(i = 1, 2)$,.f^* *is called a* Perfect* Equilibrium (P*E).

It is worth noting that the notion of a P^*E is neither stronger nor weaker than the notion of a N^*E.[4] It is obviously not weaker since the latter may prescribe sub-optimal behavior at unreached information sets. It is not stronger, since in the case of P*E, memory considerations are invoked only if the alternative strategy does as well after all histories (see Kalai and Neme (1992) for a further discussion, in a related context).Our preferred solution concept is P^*E. However, our results are quite similar irrespective of the solution concept we employ. As explained in the introduction, this stands in contrast with the existing literature which introduces complexity costs in repeated games (Abreu and Rubinstein (1988), Kalai and Neme (1992)).

[4]More precisely, it is possible that f^* be an equilibrium for some $\succcurlyeq_i \in \Xi^N$ but not an equilibrium for some $\succcurlyeq_i \in \Xi^P$. Similarly, it is possible that f^* be an equilibrium for some $\succcurlyeq_i \in \Xi^P$ but not an equilibrium for some $\succcurlyeq_i \in \Xi^N$.

3 Main results

First of all, we establish that only Markov strategies survive an iterative elimination of dominated strategies if players preferences belong to Ξ i.e. they are the only *iteratively undominated strategies* within Ξ.

Theorem 1 *Consider any alternating-move game as described above with players' preferences $\succcurlyeq_i$ belonging to the class Ξ. Let f_i be a rationalizable strategy. Then, $f_i \in F_i^0$.*

Proof. First, we introduce some notation. Given preferences $\succcurlyeq_i \in \Xi$, let $\beta_i(f_j)$ stand for the set of player i's strategies that are a best response (in terms of $\succcurlyeq_i$) to any given opponent's strategy $f_j \in F_j$. Correspondingly, associated to any subset $D_j \subset F_j$, let

$$\beta_i(D_j) \equiv \{f_i \in F_i : \exists f_j \in D_j \text{ s.t. } f_i \in \beta_i(f_j)\}.$$

Finally, for every $D_j \subset F_j$, we extend previous notation and define $m(D_j) \equiv \max\{m(f_j) : f_j \in D_j\}$, possibly equal to ∞ if the set in question is unbounded.

With this notation in hand, the main steps of the proof are a consequence of the following two claims.

Claim 1 For each $i, j = 1, 2\, (i \neq j)$, $m(\beta_i(F_j)) < v_i$.

Claim 2 For each $i, j = 1, 2\, (i \neq j)$, $[D_j \subset F_j,\ m(D_j) \geq 1] \Rightarrow m(\beta_i(D_j)) < m(D_j)$.

Claim 1 is a direct consequence of Axiom (C). To prove Claim 2, assume for the sake of contradiction that there are some $i, j = 1, 2$ $(i \neq j)$ with $D_j \subset F_j$ and $m(D_j) > 0$ but $m(\beta_i(D_j)) \geq m(D_j)$. First, we can rule out that $m(\beta_i(D_j)) > m(D_j)$ since, by a direct adaptation of a standard argument in dynamic programming, player j need not make her choice depend on longer histories than her opponent in order to maximize her discounted flow of stage payoffs. Thus, since memory is costly (by either (N) or (P)), she will not do so, i.e. $m(\beta_i(D_j)) \leq m(D_j)$.

Suppose then that $m(\beta_i(D_j)) = m(D_j) \equiv k$ and let $f_i^* \in \beta_i(f_j^*)$ with $f_j^* \in D_j$ and $m(f_i^*) = k$. We now argue that, in this case, there is an alternative strategy $\hat{f}_i \neq f_i^*$ such that $\hat{f}_i \in F_i^{k-1}$ (i.e. $m(\hat{f}_i) \leq k-1$) and

$$V_i(f^*, h^t) = V_i((\hat{f}_i, f_j^*), h^t) \text{ for every } h^t \in H_i. \tag{2}$$

In view of (N) or (P), this obviously contradicts that $f^* \in \beta_i(D_j)$.

To construct such a strategy $\hat{f}_i$, proceed as follows. First, identify all pair of histories $h^t = (\omega^\tau)_{\tau=0}^t$, $\tilde{h}^{t'} = (\tilde{\omega}^\tau)_{\tau=0}^{t'} \in H_i$ such that

$$\omega^{t-s} = \tilde{\omega}^{t'-s}, \quad \forall s = 0, 1, ..., k-1, \tag{3}$$

and then write $h^t \, R \, \tilde{h}^{t'}$, interpreting R as a binary relation on the set of histories. On the other hand, identify all those histories h^t such that there is no other $\tilde{h}^{t'} \neq h^t$ for which (3) applies and denote by $\bar{H}$ the set of all those histories. Then, for any pair $\hat{h}^t, \check{h}^{t'} \in \bar{H}$, write $\hat{h}^t \, P \, \check{h}^{t'}$, again interpreting P as a trivial binary relation. Define the binary relation $R^* = R \cup P$. This relation is obviously reflexive, symmetric, and transitive. Therefore, it partitions the set of all histories $H \equiv H_1 \cup H_2$ into equivalence classes. Clearly, one equivalence class is $\bar{H}$. On the other hand, we have the set of equivalence classes $\mathcal{Q}$ including all those $Q \subset H$ induced by R. Associated to every $Q \in \mathcal{Q}$, denote by $s(Q)$ any particular (arbitrarily chosen) action with the property:

$$\exists h^t \in Q \text{ s.t. } s(Q) = f_i^*(h^t) \text{ and} \tag{4}$$

Then, define the strategy $\hat{f}_i$ as follows:

$$\begin{aligned} \hat{f}_i(h^t) \; &= s(Q) \qquad \text{if } h^t \in Q, \; Q \in \mathcal{Q}; \\ &= f_i^*(h^t) \qquad \text{if } h^t \in \bar{H}. \end{aligned}$$

By construction, $\hat{f}_i \in F_i^{k-1}$. On the other hand, to see that we must also have (2), consider any $Q \in \mathcal{Q}$ and any $h^t \in Q$ for which $\hat{f}_i$ and f_i^* differ, i.e. $f_i^*(h^t) \neq s(Q) = \hat{f}_i(h^t)$. By (4), there exists some history $\tilde{h}^{t'}$ such that $f_i^*(\tilde{h}^{t'}) = s(Q)$. Since $f_j^* \in F_j^{k'}$ and $k' \leq k$ $(j \neq i)$, it follows that

$$f_j^*(h^t, s(Q)) = f_j^*(\tilde{h}^{t'}, s(Q))$$

and, therefore, if $s(Q)$ is an optimal action for player i after history $\tilde{h}^{t'}$, it must also be optimal after history h^t. (Note that h^t and $\tilde{h}^{t'}$ only differ in payoff-irrelevant details that, moreover, will be ignored by both players in subsequent periods since their strategies rely on at most k-length memories.) Hence (2) results, which completes the proof of Claim 2.

To complete the proof of the Theorem, we simply need to recall the concept of rationalizable strategies (Bernheim (1984) and Pearce (1984)). Construct a sequence of subsets $\{D_i^q\}_{q=0}^\infty$ for each $i = 1, 2$, as follows: First,

$D_i^0 = F_i$, $i = 1, 2$. Then, for each $q = 1, 2, ...$, we make $D_i^q = \beta(D_j^{q-1})$. The set of rationalizable strategies for each player i is given by $R_i \equiv \bigcap_{q=1}^{\infty} D_i^q$. It is immediate to see that, in view of Claims 1 and 2, $R_i \subseteq F_i^0$, which completes the proof. ■

Theorem 1 obviously ensures that every N*E of the repeated game is Markovian, as is every P*E. That is, we have the following Corollary.

Corollary 1 *Consider any alternating-move game as described above with players' preferences belonging to* Ξ^N *(resp.* Ξ^P*). If* $f^* = (f_1^*, f_2^*)$ *is a N***E (resp. P***E), then* $f_i^* \in F_i^0$ *for each* $i = 1, 2,$ *.*

In view of this Corollary, the question arises of whether a N*E, or a P*E, can be guaranteed to exist in our context. A positive answer to this question is established by the following result.

Theorem 2 *Consider any alternating-move game as described above with players' preferences* $\succcurlyeq_i$ *belonging to the class* Ξ*. An equilibrium* $f^* = (f_1^*, f_2^*)$ *always exists.*

Proof. First, notice that any Markov Perfect Equilibrium (MPE) of the game – a PE that involves strategies in F_i^0 for each player i (cf. Maskin & Tirole (1988*a*, *b*). – is both a P*E and a N*E. In view of Definition 1, this follows directly from Axioms (N), (P), and (C).

Thus, it is enough to prove existence of some MPE. Here, we rely on a slight adaptation of an argument put forward by Fudenberg & Tirole (1991, Theorem 13.1, p. 504). On the one hand, note that the (Markov) state space of the game is finite: it coincides with $\Omega \equiv S_1 \cup S_2 \cup \{h^0\}$, each state interpreted as either the action adopted by the player who is unable to revise in the current period or the empty (initial) history. Therefore, one may define an *instrumental* finite game in strategic form where players choose one of their respective Markov strategies. For each pair of these strategies, payoffs are identified with the discounted payoffs induced by the respective $\pi_i(\cdot)$, when the initial state is chosen through *any* given probability measure over Ω with full support. (Thus, unlike in the original game, the initial history need not be h^0.) Since the constructed games is finite, it has a Nash equilibrium, that is easily seen to define a MPE of the original dynamic game. To confirm it, two are the essential observations to be made. First, if every state is chosen with positive probability as the initial state, then the equilibrium strategies are evaluated after every ("Markov") history. Second, if the opponent chooses a Markov strategy, there is a best response to it that is itself Markov. This completes the proof. ■

Thus, memory considerations have a striking implication in asynchronous choice games: any equilibrium must be in Markov strategies. This result is robust to whether we use the Nash equilibrium criterion (N^*E) or require perfection (P^*E). As explained, this robustness sharply contrasts with the existing literature on strategic complexity with simultaneous moves (recall our former discussion of Abreu and Rubinstein (1988) as compared to Kalai and Neme (1992)).

It is important to understand that asynchronicity of choice plays a critical role in our result. For, in simultaneous move repeated games, memory considerations do *not* imply Markov strategies. This is a subject of our ongoing research, and beyond the scope of the present paper. However, it may be useful to sketch an illustration, in the context of the repeated Prisoners' Dilemma with simultaneous moves. Assume that players preferences are given by an element of Ξ^P. More specifically, assume that the cost of one period memory is small relative to the payoff benefit from responding optimally at histories which arise from a single deviation from the equilibrium path. Consider the following strategy which requires one period memory. Play C at $t = 1$ or if the action profile at $t - 1$ was (C, C). Play D otherwise. It is clear that if both players play this strategy, this constitutes a P*E which supports the play of (C, C) in every period, provided that δ is sufficiently large. However, the only Markov Perfect Equilibrium of this repeated simultaneous-move game is one where the players choose D every period irrespectively of history.[5]

It is well known that Markov Perfect Equilibria in asynchronous-move games span a relatively narrow range of outcomes. (Recall, for example, the analysis by Maskin and Tirole (1988*a*, 1988*b*).of dynamic oligopoly, summarized in the Introduction.) Thus, in this sense, Corollary 1 may be viewed as having potentially strong "anti-folk" implications that contrast with the general folk theorem proven by Dutta (1995) for general stochastic games in the absence of complexity considerations. (Note that any repeated game with asynchronous moves can be formulated as a stochastic game). As advanced,

[5] Sabourian (1991, 1997) provides variants of the folk theorem for repeated games with bounded memory.

the anti-folk theorem implications of our result will be further explored in Section 6, within the context of general games of common interest.

4 Learning to play Markov strategies

Under preference axioms set out above, the strong conclusion of Theorem 1 may be used to provide a sharp social-learning rationale for Markov strategies. For concreteness, we shall pursue here an evolutionary approach that reflects a process of social-learning in a large-population and random-matching context. Alternatively, we could approach the issue in what Fudenberg & Levine (1998) call a fixed-player framework, where a variety of alternative learning models (e.g. fictitious play or some of its generalizations[6]) would lead to similar conclusions.

Suppose that we have two large (continuum) populations of equal measure, 1 and 2, whose respective members are randomly matched in pairs every "learning period" to play a repeated game with asynchronous moves as described. At the end of every such period, any pair of matched players with strategy profile (f_1, f_2) receive respective von Neuman-Morgensten payoffs $\psi_1(f_1, f_2)$ and $\psi_2(f_1, f_2)$. Here, the functions $\psi_i(\cdot)$ are supposed to reflect "inclusive" payoffs, i.e. they embody both discounted payoffs and memory costs. In particular, the preferences $\succcurlyeq_i$ induced by each $\psi_i(\cdot)$:

$$(f_1, f_2) \succcurlyeq_i (f_1', f_2') \Leftrightarrow \psi_i(f_1, f_2) \geq \psi_i(f_1', f_2') \tag{5}$$

are assumed to satisfy Axioms (C) and (N) (or (P)) above.

For analytical tractability, suppose that the set of admissible strategies in each population is finite (although sufficiently "rich", as detailed below). For each population i, denote by Y_i the set of such admissible strategies. Furthermore, let $\Theta_i \equiv \Delta(Y_i)$ stand for the set of possible population configurations, a typical $\theta_i \in \Theta_i$ specifying the frequency $\theta_i(f_i)$ of individuals in population i who adopt any given strategy $f_i \in Y_i$.

The state space of the learning process is $\Theta = \Theta_1 \times \Theta_2$, a typical element denoted by $\theta \equiv (\theta_1, \theta_2)$. Social learning is modelled as a dynamic process on the space Θ with a law of motion of the following form:

$$\theta_i^{q+1}(f_i) = \theta_i^q(f_i)\, \Phi_i(f_i; \theta^q) \qquad (i = 1, 2;\ f_i \in Y_i), \tag{6}$$

[6]See, for example, Milgrom & Roberts (1991), Fudenberg & Kreps (1993), or Kaniovski & Young (1995).

for some Lipschitz-continuous functions $\Phi_i(\cdot)$, where $q \in \{0, 1, 2, ...\}$ indexes learning periods. Since the state space Θ must be invariant under this dynamics, one must have that, for all $\theta \in \Theta$, $i = 1, 2$, $\sum_{f_i \in Y_i} \theta_i(f_i)\, \Phi_i(f_i; \theta) \equiv 1$.

Naturally, we want to impose some conditions on the functions $\Phi_i(f_i; \theta^q)$ suitably formalizing the idea that (6) represents a social learning process. In recent evolutionary literature, a common such requirement is labelled *payoff monotonicity* (cf. Nachbar (1990)). As adapted to the present case, it postulates that for all $i, j = 1, 2\, (i \neq j)$, f_i, $f_i' \in Y_i$, $\theta \in \Theta$, the following condition should hold:

$$\Phi_i(f_i; \theta) \geq \Phi_i(f_i'; \theta) \Leftrightarrow \sum_{f_j \in Y_j} \psi_i(f_i, f_j)\, \theta_j(f_j) \geq \sum_{f_j \in Y_j} \psi_i(f_i', f_j)\, \theta_j(f_j). \tag{7}$$

Of course, in view of (6), the key implication of (7) is that any strategy whose average payoff is higher than some other alternative one will grow at a faster rate (or shrink more slowly) than the latter. This is a rather weak requirement, consistent with a wide variety of specific formulations of social learning (e.g. inter-agent imitation).[7] A particular well-known instance of a payoff-monotonic system is given by Replicator Dynamics (RD), which is formulated as follows:

$$\Phi_i^R(f_i; \theta) \equiv 1 + \alpha \left\{ \sum_{f_j \in Y_j} \psi_i(f_i, f_j)\, \theta_j(f_j) - \sum_{f_i' \in Y_i} \theta_i(f_i) \left[\sum_{f_j \in Y_j} \psi_i(f_i', f_j)\, \theta_j(f_j) \right] \right\}, \tag{8}$$

where $\alpha > 0$ is a constant determining the speed of adjustment.

As advanced, to obtain interesting long-run conclusions we must require that the set of admissible strategies be sufficiently rich. Depending on the generality demanded from the evolutionary system, two alternative possibilities will be considered in this respect. The first one, labelled (PS) below, simply states that the set of admissible strategies for each population must *include* all those *pure* strategies with memory requirements up to some large enough level (in particular, this set must include pure strategies whose memory demands are so large that, in view of (C), they will never be used).

(PS) Let $\hat{F}_i \equiv \{f_i : H_i \to S_i\}$ denote the set of pure strategies of population i. For each $i = 1, 2$, $Y_i \supset \{f_i \in \hat{F}_i : m(f_i) \geq v_i\}$, where v_i is as in (C).

[7] We refer the reader to some of the recent monographs on Evolutionary Game Theory (e.g. Weibull (1995) or Samuelson (1997)) for an elaboration on these matters

The above condition only requires that each Y_i should include all *pure* strategies that are below a certain high enough level of complexity. This is enough for our purposes if we restrict our consideration to some relatively narrow family of evolutionary systems (e.g. the RD). However, if we enlarge the range of the analysis to other payoff-monotonic systems, we may need to include some mixed strategies in the admissible set. To address this issue precisely, we now introduce some additional notation.

Assume player i's preferences satisfy (5), and let $f_i \in F_i$ be any non-Markovian strategy for population i. Then, $m(f_i) = k$ for some $k \in \mathbb{N}$ and it follows from previous analysis (cf. the proof of Theorem 1) that $f_i \notin \beta_i(F_j^k)$, i.e. f_i is *not* a best response to opponent's strategies that involve memory requirements smaller or equal than those of f_i. By a standard argument establishing the coincidence of rationalizability and iterated dominance in bilateral games (cf. Pearce (1984)), $f_i \notin \beta_i(F_j^k)$ implies that there exists some $f_i' \in F_i$ (in fact, we must have $f_i' \in F_i^{k'}$ for some $k' < k$) such that

$$\psi_i(f_i', \tilde{f}_j) > \psi_i(f_i, \tilde{f}_j), \quad \forall \tilde{f}_j \in F_j^k.$$

That is, f_i is (strictly) dominated by f_i' if player j's strategies are restricted to lie in F_j^k.

Given any $f_i \in F_i$, denote by $\mathcal{D}(f_i)$ the set of strategies in F_i which dominate f_i when player j's strategies are restricted to lie in $F_j^{m(f_i)}$. As explained, $\mathcal{D}(f_i) \neq \emptyset$ whenever $f_i \notin F_i^0$, i.e. as long as f_i is not Markovian. We now formulate an alternative richness condition on the set of admissible strategies as follows:

(DS) For each population $i = 1, 2$,

$$[f_i \in Y_i,\ \mathcal{D}(f_i) \neq \emptyset] \Rightarrow \mathcal{D}(f_i) \cap Y_i \neq \emptyset.$$

Building upon each of the two previous richness conditions, (PS) and (DS), we have the following two results.

Theorem 3 *Suppose that the evolutionary system (6) satisfies (8) – i.e. it is the Replicator Dynamics – with the set of admissible strategies verifying (PS) and payoffs satisfying (5) for preferences* $\succcurlyeq_i \in \Xi$. *Then, if the initial conditions of the system* $\theta^0 \in int(\Theta)$, *there exists some* $\bar{\alpha} > 0$ *such that, if* $\alpha \leq \bar{\alpha}$, $\lim_{q\to\infty} \theta_i^q(f_i) = 0$ *for all* $f_i \notin F_i^0$, $i = 1, 2$.

Theorem 4 *Suppose that the evolutionary system (6) satisfies (7) – i.e. it is payoff-monotonic – with the set of admissible strategies verifying (DS) and payoffs satisfying (5) for preferences* $\succcurlyeq_i \in \Xi$. *Then, if the initial conditions of the system* $\theta^0 \in int(\Theta)$, $\lim_{q\to\infty} \theta_i^q(f_i) = 0$ *for all* $f_i \notin F_i^0$, $i = 1, 2$.

The proof of the previous two results follows from an application of standard arguments in Evolutionary Game Theory (see e.g. Weibull (1995)). By way of illustration, we only outline a proof of the latter one.

Proof of Theorem 4 (sketch). Suppose, for the sake of contradiction, that $\theta^0 \in int(\Theta)$ but there exists a strategy $f_i \in Y_i$ with $m(f_i) > 0$ whose long-run frequency satisfies $\limsup_{q\to\infty} \theta_i^q(f_i) > 0$. Let $\tilde{f}_i$ be the strategy displaying this feature whose memory requirements $m(\tilde{f}_i)$ is maximum in the two populations. By Condition (DS) and the continuity of each $\Phi_i(\cdot)$, there is some $f_i' \in Y_i$ and $\bar{q} \in \mathbb{N}$ such that, for all $q \geq \bar{q}$,

$$\sum_{f_j \in Y_j} \psi_i(f_i', f_j)\, \theta_j^q(f_j) - \sum_{f_j \in Y_j} \psi_i(\tilde{f}_i, f_j)\, \theta_j^q(f_j) > 0. \tag{9}$$

We now argue that this implies $\lim_{q\to\infty} \theta_i^q(\tilde{f}_i) = 0$. For suppose otherwise. Then, using (6), we may write:

$$\frac{\theta^{q+1}(\tilde{f}_i)}{\theta^q(\tilde{f}_i)} - \frac{\theta^{q+1}(f_i')}{\theta^q(f_i')} = \Phi_i(\tilde{f}_i; \theta^q) - \Phi_i(f_i'; \theta^q)$$

or, rearranging terms:

$$\frac{\theta^{q+1}(\tilde{f}_i)/\theta^q(\tilde{f}_i)}{\theta^{q+1}(f_i')/\theta^q(f_i')} = \frac{\theta^{q+1}(\tilde{f}_i)/\theta^{q+1}(f_i')}{\theta^q(\tilde{f}_i)/\theta^q(f_i')} = 1 + \frac{\Phi_i(\tilde{f}_i; \theta^q) - \Phi_i(f_i'; \theta^q)}{\theta^{q+1}(f_i')/\theta^q(f_i')}$$

that, using (6) again, as well (7) and (9), implies:

$$0 < \frac{\theta^{q+1}(\tilde{f}_i)/\theta^q(\tilde{f}_i)}{\theta^{q+1}(f_i')/\theta^q(f_i')} = 1 + \frac{\Phi_i(\tilde{f}_i; \theta^q)}{\Phi_i(f_i'; \theta^q)} - 1 \leq \zeta < 1.$$

Hence we conclude that

$$\lim_{q\to\infty} \frac{\theta^q(\tilde{f}_i)}{\theta^q(f_i')} \leq \lim_{q\to\infty} \zeta^q = 0,$$

a contradiction. ■

Theorem 4 indicates that, under condition (DS), every payoff-monotonic process of social learning will have the fraction of players relying on non-Markovian strategies in each population i, $\sum_{f_i \notin F_i^0} \theta^q(f_i)$, vanish in the long run. Of course, this does not imply that they must learn to play some (Markovian) equilibrium. However, this will happen if the repeated game is, for example, dominance solvable (i.e. has a unique strategy profile that survives iterated elimination of dominated strategies and, therefore, is the unique rationalizable profile). In fact, a context of this type will be encountered in Section 6 where the analysis focuses on asynchronous repeated games of common interest.

5 Lexicographic preferences

In the game-theoretic literature with complexity costs, it is often postulated that these costs are of only subsidiary importance when compared to discounted stage payoffs. This heuristic idea is typically formalized by means of the so-called *lexicographic preferences* (cf. Abreu & Rubinstein (1988)). In our context, this kind of preferences may be characterized by the following axiom:

(L) Given any $h^t \in H_i$, let f, $\tilde{f} \in F$ satisfy one of the following:
(i) $V_i(f) > V_i(\tilde{f})$;
(ii) $V_i(f) = V_i(\tilde{f})$ and $m(f_i) < m(\tilde{f})$.
Then, $f_i \succ_i \tilde{f}$. Otherwise, $\tilde{f} \not\succ_i f_i$.

As explained above, (L) is inconsistent with Axiom (C) – cf. Section 2. Therefore, we cannot apply Theorem 1 to conclude that every iteratively undominated strategy is Markovian. However, if we focus on N*E, the weaker conclusion of Corollary 1 obtains for lexicographic preferences if we also postulate the following condition:

(B) There exist some u_i $(i = 1, 2)$ such that the set of admissible strategies is given by $\hat{F}_i \equiv \{f_i \in F_i : m(f_i) < u_i\}$ where, possibly, $u_i = \infty$.

The above condition postulates that all admissible strategies involve finite (possibly unbounded) memory requirements. To see that this condition has the claimed implications, assume that both (B) and (L) hold and let $f^* =$

(f_1^*, f_2^*) be a N*E of the game. Suppose w.l.o.g. that, say, $m(f_1^*) \geq m(f_2^*)$. If $m(f_1^*) > 0$, a direct application of the argument used in the proof of Theorem 1 implies that player 1 can match the payoff achieved by f_1^* against f_2^* with an alternative strategy $\hat{f}_1$ such that $m(\hat{f}_1) < m(f_1^*)$. Therefore, f_1^* cannot be a best response to f_2^*, which is obviously a contradiction. We may conclude therefore that if $f^* = (f_1^*, f_2^*)$ is a N*E, it must involve $f_i^* \in F_i^0$ for each $i = 1, 2$.

Now suppose that Condition (B) is strengthened as follows:

(B)′ There exist some *finite* u_i $(i = 1, 2)$ such that the set of admissible strategies is given by $\hat{F}_i \equiv \{f_i \in F_i : m(f_i) \leq u_i\}$.

The above axiom imposes *directly* on the set of admissible strategies the boundedness condition induced by (C) on the (unrestricted) set of undominated strategies. Thus, it is immediate to verify that, relying again on the logic underlying Theorem 1, it follows that, under both (L) and (B)′, all (admissible) strategies that survive an iterative elimination of dominated strategies are Markovian. Moreover, since (L) obviously implies (C), (B) and (L) also lead to the long-run conclusion concerning payoff-monotonic learning process contained in Theorem 4. In this respect, the role of (B) is merely to justify the contemplated restriction to a finitely dimensional dynamical system (cf. Footnote **??**).

For limitations of space, the discussion in this section has abstracted from considerations of perfection (i.e. preferences have been assumed to not take into account payoffs which accrue after unreached histories). However, it is clear that the arguments presented here also apply if we take into account such considerations.

6 An application: common interest games

There is a large literature which considers dynamic games and focuses on Markov equilibria. Examples include Maskin and Tirole (1988a, 1988b) and Eaton and Engers (1990). Theorem 1 provides a justification for this focus, in the context where players move asynchronously. We now consider an application to the question of equilibrium selection. The recent game theoretic literature has focused on the question of equilibrium selection in such games, using an evolutionary approach. In a repeated game context, the folk theorem implies that there are no such selection results. However, we show that asynchronous moves in conjunction with our assumption that players seek to economize on the memory required to implement a strategy gives us strong selection results.

Our first result shows the possibility of efficient outcomes and applies to arbitrary two-player common interest games. Consider a two-player game G with action sets $S_1 = \{s_{11}, s_{12}, ..., s_{1m}\}$, $S_2 = \{s_{21}, s_{22}, ..., s_{2n}\}$.

Definition 2 *G is of common interest if there exists an action profile, say (s_{11}, s_{21}), which strictly Pareto-dominates all other profiles, i.e.* $\forall i = 1, 2,\ \forall (s_1, s_2) \in S \equiv S_1 \times S_2$,

$$\forall (s_1, s_2) \neq (s_{11}, s_{21}), \quad \pi_i(s_{11}, s_{21}) > \pi_i(s_1, s_2).$$

Let $s^1 = (s_{11}, s_{21})$ denote the efficient action profile and let π_i^* denote the payoff of player i at this profile. Our focus in upon the conditions under which the efficient outcome is uniquely selected if players preferences belong to Ξ^P, use iteratively undominated strategies, and are patient. As a preliminary to this, we show first that there always exists a MPE (and therefore a P$^*E$) where this outcome is played in every period.

Proposition 1 *Consider an alternating move game where the stage game has common interests, and any $\delta < 1$. There exists a P^*E, $\hat{f}$ which induces the paths of play, $(\hat{s}_1^t, \hat{s}_2^t)_{t=0}^{\infty}$ with $\hat{s}_i^t = s_{i1}$ for each $i = 1, 2$, and all $t \geq 0$.*

Proof. We show the existence of a MPE where the action profile $s^1 = (s_{11}, s_{21})$ is played in every period. Let $\Omega_1 = S_2$ be the state space for player 1, and let $\Omega_2 = S_1$ be the state space for player 2. Let $\tilde{F}_i^0$ be the set of all Markov strategies which play s_{i1} in the initial period and whenever the state is s_{j1} i.e.:

$$\tilde{F}_i^0 = \{f_i \in F_i^0 : f_i(h^0) = f_i(s_{j1}) = s_{i1}\}$$

We show first that if $f_j \in \tilde{F}_j^0$, then player i has a Markov best response to this which belongs to $\tilde{F}_i^0$. To see this, observe that if the state is either the initial state or s_{j1}, player i obtains a payoff of π_i^* in every period by playing a strategy in $\tilde{F}_i^0$. Since π_i^* is the maximal payoff in the game, this ensures that some strategy in $\tilde{F}_i^0$ is a best response to any strategy in $\tilde{F}_j^0$.

To show existence of a MPE, restrict each player i to strategies in $\tilde{F}_i^0$. By the same argument as in theorem 2, this restricted game has a MPE.

This is also a MPE of the unrestricted game, since for any strategy f_i' in the complement of $\tilde{F}_i^0$, there exists a strategy in $\tilde{F}_i^0$ and agrees with f_i' at all states except s_{j1} and the initial state, and is hence weakly better than f_i' . ■

Remark 1 *Observe that this proof also applies if s^1 is only weakly Pareto-efficient, i.e. the inequality in definition of common interest is weak.*

We now consider the conditions under which the efficient equilibrium is selected. We dub the condition which ensures efficient equilibrium selection *Generalized Bilateral Risk Dominance (GBRD)* which is defined as follows:

$$\exists i \in \{1,2\} : \forall s_{iq}, s_{ir} \in S_i,\ \forall s_{ju}, s_{jv} \in S_j\ (j = 3 - i),$$

$$\pi_i(s_{i1}, s_{j1}) + \pi_i(s_{i1}, s_{ju}) + \pi_i(s_{iq}, s_{ju}) \geq \pi_i(s_{ir}, s_{j1}) + \pi_i(s_{ir}, s_{jv}) + \pi_i(s_{iq}, s_{jv}).$$

Note that if we choose $s_{ju} = s_{jv}$, the above condition becomes:

$$\exists i \in \{1,2\} : \forall s_{ir} \in S_i,\ \forall s_{ju} \in S_j\ (j = 3 - i),$$

$$\pi_i(s_{i1}, s_{j1}) + \pi_i(s_{i1}, s_{ju}) \geq \pi_i(s_{ir}, s_{j1}) + \pi_i(s_{ir}, s_{ju}),$$

that is labelled *Bilateral Risk Dominance (BRD)*. Obviously, GBRD is a generalization of BRD.

Lemma 1 *Let GBRD be satisfied and let players players preferences belong to Ξ^P. Then, at the state $\omega_i = s_{j1}$,* any *Markov strategy for player i which is undominated requires this player to choose action s_{i1} at this state, for any belief about the strategy (pure or mixed) that i believes that player j is playing.*

Proof. Let σ_i be any Markov strategy of player i, which plays an action different from s_{i1} at the state $\omega_i^t = s_{j1}$. Let σ_j be an arbitrary pure strategy of player j. We shall show that for any σ_j, there exists a strategy of player i, σ_i' which strictly dominates σ_i. We construct σ_i' as follows: σ_i' chooses the action s_{i1} at the state $\omega_i^t = s_{j1}$. In period $t+2, \sigma_i'(\omega_i^{t+2}) = \sigma_i(s_{j1})\, \forall \omega_i^{t+2}$, i.e. σ_i' chooses the same actions at $t+2$ (independently of the state ω_i^{t+2}) that σ_i plays at the state s_{j1}, but two periods later. In subsequent periods, σ_i' chooses the same actions as σ_i. Notice that σ_i' involves a "two-step" deviation from σ_i.

Let σ_i choose $s_{ir} \neq s_{i1}$ (with positive probability) at $\omega_i = s_{j1}$, and suppose that the pure strategy σ_j chooses action s_{ju} at $\omega_j = s_{i1}$ and s_{jv} at $\omega_j = s_{ir}$. The continuation payoff from playing σ_i at $\omega_i = s_{j1}$ is given by

$$V_i(\sigma_i, \sigma_j; s_{j1}) = \pi_i(s_{ir}, s_{j1}) + \delta\pi_i(s_{ir}, s_{jv}) + \delta^2\pi_i(s_{iq}, s_{jv}) + \delta^3\tilde{V}_i((\sigma_i, \sigma_j; s_{iq})$$

where $\tilde{V}_i((\sigma_i, \sigma_j; s_{iq})$ denotes the continuation payoff to i when it is $j's$ turn to move and the state is $\omega_j = s_{iq}$. The continuation payoff from playing σ_i' at $\omega_i = s_{j1}$ is given by

$$V_i(\sigma_i', \sigma_j; s_{j1}) = \pi_i(s_{i1}, s_{j1}) + \delta\pi_i(s_{i1}, s_{ju}) + \delta^2\pi_i(s_{iq}, s_{ju}) + \delta^3\tilde{V}_i((\sigma_i, \sigma_j; s_{iq})$$

Clearly, GBRD implies that $V_i(\sigma_i', \sigma_j; s_{j1}) > V_i(\sigma_i, \sigma_j; s_{j1})$ if δ is sufficiently close to 1. Observe further that the strategy σ_i' which we have constructed does not depend upon σ_j, i.e. σ_i' dominates σ_i independently of the Markov pure strategy played by j. Hence σ_i' dominates σ_i even when j plays any mixed Markov strategy. ■

We now show that GBRD ensures play of the efficient equilibrium by an iteration of strict dominance arguments. To this end, consider the following claims:

1. If player i plays a Markov strategy which plays s_{i1} at $\omega_i = s_{j1}$, then it is strictly optimal for player j to play s_{j1} at $\omega_j = s_{i1}$. By doing so, j earns a payoff of $\pi_j(s_{i1}, s_{j1})$, which is his maximal payoff in the game, whereas by choosing any different action, j earns strictly less.

2. If player i plays a Markov strategy which plays s_{i1} at $\omega_i = s_{j1}$, then it is strictly optimal for player j to play s_{j1} at every state except possibly the initial null state. Let $\omega_i = s_{ju}$, so that the payoff to player i from the candidate strategy is given by $(1-\delta)\pi_i(s_{i1}, s_{ju}) + \delta\pi_i(s_{i1}, s_{j1})$. Let $s_{ir} \neq s_{i1}$ be any other action for i, and let s_{jv} be the action taken by j when i plays s_{ir}. Hence a one-step deviation from the candidate strategy yields a payoff of $(1-\delta)\pi_i(s_{ir}, s_{ju}) + \delta(1-\delta)\pi_i(s_{ir}, s_{jv}) + \delta^2\pi_i(s_{i1}, s_{jv}) + \delta^3\pi_i(s_{i1}, s_{j1})$. GBRD implies that the one-step deviation payoff is strictly less if δ is sufficiently close to 1.

These claims establish that player i plays s_{i1} in the first period t where he moves alone, i.e. either $t = 1$ (if $i = 1$) or at $t = 2$ (if $i = 2$). Furthermore, once i does this, j also plays s_{j1} thereafter. We have therefore proved the following proposition.

Proposition 2 *Consider any alternating-move game where the stage game is a common interest game which satisfies GBRD, and players preferences belong to* Ξ^P*. Then, there exists some* $\delta_0 < 1$ *such that, if* $\delta \geq \delta_0$*, every* f^* *which is iteratively undominated induces a path of play* $(s_1^{t*}, s_2^{t*})_{t=0}^{\infty}$ *such that* $s_i^{t*} = s_{i1}$ *for each* $i = 1, 2$*, and all* $t \geq 3$*.*

To illustrate this selection result, consider the following 2×2 common interest game, with payoffs given by the following table:

	s_{21}	s_{22}
s_{11}	a_1, a_2	d_1, c_2
s_{12}	c_1, d_2	b_1, b_2

Table 1

For $i \in \{1, 2\}$,we assume $a_i > c_i$, $b_i > d_i$, and $a_i > b_i$. In other words, $s^1 = (s_{11}, s_{21})$ and $s^2 = (s_{12}, s_{22})$ are strict Nash equilibria, and s^1 is the Pareto-efficient one.

In a 2×2 game, it is easily verified that GBRD is equivalent to BRD. Hence the above game satisfies GBRD if there is at least one player, i, such that this player's first action, s_{i1}, is weakly preferred to her other action s_{i2}, given that player i expects player j to play both his actions with equal probability, i.e. if $a_i + d_i < c_i + b_i$. Hence the above proposition shows at if the efficient action is a "safe action" *for at least one player*, and players are patient and have preferences belonging to Ξ^P, then it must be played if players use iteratively undominated strategies. Our selection criterion is closely related to the notion of risk-dominance in the sense of Harsanyi and Selten (1988). In particular, if s^1 is *not* risk-dominated, this is sufficient to ensure that it is uniquely selected (this is also a necessary condition if the stage game is symmetric).

The intuition for this result is as follows. Suppose that the game satisfies BRD so that s_{i1} is a safe action for player i. Then player i will play s_{i1} at the state $\omega_i = s_{j1}$, no matter what pure Markov strategy j plays. Indeed, this is so even if j plays the action s_{j2} at every state. In this case, i cannot avoid a transition to the inferior equilibrium s^2. However, if s_{i1} is a safe strategy, i prefers that j initiate the transition, and hence will not initiate the transition himself. If i initiates such a transition, his payoff over the next two periods is $c_i + \delta b_i$, whereas he lets j initiate the transition, his payoff is $s_{i1} + \delta d_i$, and the latter is strictly greater for any δ if s_{i1} is a safe strategy.[8] Given that i will not initiate the transition, neither will j, and this suffices to ensure that the action profile s^1 is absorbing. This in turn implies that if players are playing the inefficient equilibrium, at least one of them will be willing to initiate a transition to the efficient one, secure in the knowledge that the other player will follow.

If BRD is not satisfied, i.e. if s_{i1} is the more risky action for both players, this argument does not follow. In this case, there is a Markov Perfect Equilibrium where each player $i, i \in \{1, 2\}$, plays the inefficient action, s_{i2}, at all states. In this case, if the state is $\omega_i = s_{j1}$, player i prefers to initiate the transition to s^2 himself; the difference in payoff from this, as compared to letting j initiate the transition is given by

$$\delta(b_i - d_i) - (s_{i1} - c_i)$$

which is strictly positive for δ sufficiently close to 1 if s_{i1} is more risky. Given this, it is also optimal for both players to play the action profile s^2 in the first period.

Proposition 3 *Consider any alternating-move coordination game as described above, and assume that $s_{i1} + d_i < b_i + c_i$ for every $i = 1, 2$. If δ is sufficiently large, there exists a P*E, $\tilde{f}$, which induces the path of play, $(\tilde{s}_1^t, \tilde{s}_2^t)_{t=0}^{\infty}$, with $\tilde{s}_i^t = B$ for each $i = 1, 2$, and all $t \geq 0$.*

Proof. We show that the pure strategy profile given by an initial choice of s^2 and the Markov strategies (s_{12}, s_{12}) and (s_{22}, s_{22}) are a MPE if δ is sufficiently large. The optimality of s_{i2} when the state is s_{j1} has been verified in the text. If the state is s_{j2}, the equilibrium strategy yields b_i in every period, whereas a one-step deviation to s_{i1} yields d_i for two periods followed by b_i thereafter, which is strictly less. Finally, the initial choice of s_{i2} is also optimal given that player j is playing s_{j2}. ■

We also provide an example to show that bilateral risk dominance is not sufficient to ensure efficiency. I.e. if we have game with only two pure strategy Nash equilibria, $s^1 = (s_{11}, s_{21})$ and $s^2 = (s_{12}, s_{22})$ where the Pareto-efficient equilibrium s^1 risk dominates s^2, this is not sufficient to rule out a P*E where

[8] In either event his payoffs after 2 periods are b_i in every period. Note that we have factored out the normalization factor $(1 - \delta)$ in the discussion in the text.

s^2 is played in every period. Although the third action s_{i3} may be strictly dominated for both players, s_{i2} may be a better response to s_{j3} than s_{i1} is, and this may ensure that a player does not want to always play s_{i1} when the state is s_{j1}. Consider the following symmetric game with payoffs given by the matrix below

	s_{j1}	s_{j2}	s_{j3}
s_{i1}	a_{11}	a_{12}	a_{13}
s_{i2}	a_{21}	a_{22}	a_{23}
s_{i3}	a_{31}	a_{32}	a_{33}

Assume that $s^1 = (s_{11}, s_{21})$ and $s^2 = (s_{21}, s_{22})$ are strict Nash equilibria, where s^1 Pareto-dominates and risk-dominates s^2,i.e. $a_{11} > a_{22}$ and $a_{11} + a_{12} > a_{21} + a_{22}$. Assume that:

1. $a_{31} + a_{32} + a_{22} > a_{11} + a_{13} + a_{23}$
2. $a_{31} + a_{32} > a_{21} + a_{22}$
3. $a_{22} + a_{23} > a_{32} + a_{33}$

Note that assumption (1) above ensures that GBRD is not satisfied.

Consider the following Markov strategy for player i :

$$f_i(\omega_i) = \begin{cases} s_{i2} \text{ if } \omega_i \neq s_{j1} \\ s_{i3} \text{ if } \omega_i = s_{j1} \end{cases}$$

We now show that $f = (f_1, f_2)$ is a MPE. To verify this, observe that assumption (1) above ensures that playing s_{i3} is better than playing s_{i1} at the state $\omega_i = s_{j1}$, while assumption (2) ensures that playing s_{i3} is better than playing s_{i2} at this state. Assumption (3) ensures that playing s_{i2} is better than playing s_{i3} when $\omega_i = s_{j3}$, while assumption (1) also ensures that playing s_{i1} is not optimal at this state. Finally, these assumptions in conjunction with the fact that s^2 is a strict Nash equilibrium also ensure that it is optimal to play s_{i2} at $\omega_i = s_{j2}$.[9]

[9] After the first version of this paper was completed, we came across the work of Haller and Lagunoff (1997), who analyze Markov perfect equilibria in asynchronous choice games. The focus of this paper is quite different from ours – whilst we are concerned with providing a foundation for MPE, they show that the set of MPE is generically finite. However, they

7 Conclusion

The main objective of this paper has been to provide a theoretical rationale for Markov equilibria in repeated games with asynchronous moves (Theorem 1). Although the restriction to Markov equilibria has been ubiquitous in applications, there has hitherto been relatively little in terms of theoretical justification for this assumption. An important exception is Maskin and Tirole (1997). Their main purpose is to develop and refine the notion of Markov perfect equilibrium in a large class of dynamic games, and to show that concept is robust to small perturbations in payoffs. They also provide a learning framework to justify MPE. Players are drawn from a large population and randomly matched to play a dynamic game. In between the plays of the dynamic game, each player observes the information partition of the other players. Players also have the option to increase (but not to decrease) the complexity of their strategies, as measured by the extent of history dependence. Each increase in complexity incurs a cost, which is sunk. Maskin and Tirole show that if players are sufficiently patient, they will prefer to start with simple strategies, since this allows them the option to upgrade subsequently. Since all players start with Markov strategies, this option is never used.

In comparison with Maskin and Tirole's results, the rationale for Markov equilibrium offered in this paper is more modest, and applies only to asynchronous move games. For such games however, our rationale is a robust one — if players economize on memory, any equilibrium, whether Nash or subgame perfect, is necessarily in Markov strategies. Our formal argument requires the iterative elimination of strictly dominated strategies, and hence one may also construct a fairly general learning process,[10] which ensures that only Markov strategies are played.

Our approach to measuring strategic complexity bears comparision with the literature on finite automata, as exemplified by Kalai and Stanford

also have an efficient equilibrium selection result for 2×2 common interest games. Their selection result is weaker than Proposition 2 in several respects. First, it relies on an equilibrium notion (MPE) rather than iterated dominance. Second, it applies only to 2×2 games. Third, the payoff condition contemplated is also stronger – they require that $d_1 > c_1$ and $d_2 > c_2$ (cf. Table 1), whereas we only demand that $a_i + d_i \geq b_i + c_i$ for some i (recall that $a_i > b_i$, since s^1 is taken to be the efficient equilibrium).

[10]In particular, players can be allowed to both upgrade or downgrade the memory requirements of their strategies.

(1988), Abreu and Rubinstien (1988), Lipman and Srivastava (1990) and Kalai and Neme (1992). While complexity considerations are acknowledged to be important, it seems to us that there is no universally accepted measure of strategic complexity. In particular, one obtains radically different results depending upon whether one uses the number of states measure (as in Abreu-Rubinstein) or if one also takes into account the transitions between states (as in Lipman-Srivastava). One also obtains radically different results depending upon whether considerations of subgame perfection are incorporated or not. In defense of our approach, we would like to point out that any finite memory strategy is also finitely complex, in terms of the number of states measure. On the other hand, the converse is not true — a finitely complex strategy may require infinite memory. Roughly speaking, given a set of possible histories, H, the strategic complexity of a strategy f_i is given by size of the coarsest partition of H such that f_i is measurable with respect to this partition. In making this partition, one is unrestricted — in particular, two histories which differ only in terms of actions taken very far back in the past, could be in different elements of the partition. Hence a strategy with a limited number of states can require large memory. To illustrate, let the stage game be the prisoners' dilemma. Consider the behavior induced by the grim-trigger strategy on two equal length histories h and h'. In both these histories, players have played C in every period except period one. In period one, both players have played C under history h, whereas one player has played D under h'. The grim-trigger strategy outputs different actions at these two histories, irrespective of the length of these histories, and hence requires unbounded memory. In our view, memory is also costly and it is interesting to investigate the consequences of players economizing on memory. Our results are robust — we obtain similar results regardless of whether we consider Nash or subgame perfect equilibria.We refer the reader to the interesting paper of Dow (1991) which investigates the implications of memory constraints in the context of search.

The second contribution of this paper has been to show that incorporating memory considerations leads to efficient equilibrium selection in a class of common interest games (Proposition 2). The risk dominance criterion and generalizations of it are found to play an important role, as in the evolutionary literature. For simplicity, our discussion has been concerned with a simple setup involving only two players who move in alternation. However, it is straightforward to verify that the argument underlying our main theorem (theorem 1) extends to any context involving a finite number of players who (almost surely) never receive a simultaneous opportunity to revise their actions. On the other hand, the efficiency result established by Proposition 2 is bound to apply under much more restrictive circumstances.

References

[1] Abreu, D. & A. Rubinstein (1988), "The structure of Nash equilibria in repeated games played by finite automata", *Econometrica* **56**, 1259-1281.

[2] Beggs, A., and P. Klemperer (1992), "Multi-period competition with switching costs", *Econometrica* **60**, 544-565.

[3] Bernheim, D. (1984): "Rationalizable strategic behavior", *Econometrica* **52**, 1007-28

[4] De Fraja, G. (1993), "Staggered vs synchronized wage setting in oligopoly", *European Economic Review* **37**, 1507-1522.

[5] Dow, J. (1991), "Search decisions with limited memory", *Review of Economic Studies* **58**, 1-14.

[6] Dutta, P. (1995): "A folk theorem for stochastic games",*Journal of Economic Theory* **66**, 1-32.

[7] Eaton, J. , and M. Engers, (1990), "Intertemporal price competition", *Econometrica* **58**, 637-659.

[8] Fudenberg, D. & D. Kreps (1993): "Learning mixed equilibria", *Games and Economic Behavior* **5**, 320-67.

[9] Fudenberg, D. & D. Levine (1998): *The Theory of Learning in Games,* Cambidge: MIT Press.

[10] Fudenberg, D. & J. Tirole (1991): *Game Theory,* Cambridge (Mass.): MIT Press.

[11] Haller, H., and R. Lagunoff (1997), "Markov equilibrium in asynchronous choice repeated games", mimeo.

[12] Jéhiel, P. (1995), "Limited horizon forecast in repeated alternate games", *Journal of Economic Theory* **67**,497-519.

[13] Kalai, E., and A. Neme.(1992), "The strength of a little perfection", *International Journal of Game Theory* **20**, 335-355.

[14] Kalai, E., and W. Stanford (1988), "Finite rationality and interpersonal complexity in repeated games", *Econometrica* **56**, 397-410.

[15] Kaniovski, Y. and P. Young (1995): "Learning dynamics in games with stochastic perturbations", *Games and Economic Behavior* **11**, 330-63.

[16] Lagunoff, R. & A. Matsui (1997): "Asynchronous choice in repeated coordination games", *Econometrica* **65**, 1467-77.

[17] Levhari, D. and L. Mirman (1980), "The great fish war", *Bell Journal of Economics* 12, 322-344.

[18] Lipman, B., and S. Srivastava (1990), "Informational requirements and strategic complexity in repeated games", *Games and Economic Behavior* **2**, 273-290.

[19] Maskin, E. & J. Tirole (1988*a*): "A theory of dynamic oligopoly, I: overview and quantity competition with large fixed costs", *Econometrica* **56**, 549-70.

[20] Maskin, E. & J. Tirole (1988*b*): "A theory of dynamic oligopoly, II: Price competition, kinked demand curves, and Edgeworth cycles", *Econometrica* **56**, 571-600.

[21] Maskin, E. & J. Tirole (1997), "Markov equilibrium", mimeo.

[22] Milgrom, P. & J. Roberts (1991). "Adaptative and sophisticated learning in repeated normal-form games", *Games and Economic Behavior* **3**, 82-100.

[23] Nachbar, J. (1990): "Evolutionary selection in dynamic games", *International Journal of Game Theory* **19**, 59-90.

[24] Pearce, D. (1984): "Rationalizable strategic behavior and the problem of perfection", *Econometrica* **52**, 1029-1050

[25] Sabourian, H. (1991), "The folk-theorem of repeated games with bounded (one-period) memory", mimeo.

[26] Sabourian, H., (1997), "Repeated games with M-period bounded memory (pure strategies)", *Journal of Mathematical Economics*, forthcoming.

[27] Samuelson, L. (1997): *Evolutionary Games and Equilibrium Selection*, Cambridge: MIT Press.

[28] Weibull, J. (1995): *Evolutionary Game Theory*, Cambridge: MIT Press.