Markovian Dynamics in Asynchronous Stochastic Models

Subir. K. Chakrabarti[1]

April 24, 2019

Department of Economics
Indiana University Purdue University Indianapolis(IUPUI)
425 University Blvd.
Indianapolis, IN 46202

Abstract: We study Asynchronous Stochastic Dynamic Models in which a single agent chooses in each period. Examples of such dynamic models are encountered frequently. Markets in which a buyer and a seller make alternating offers, Auctions in which bidders bid sequentially, or a duopoly in which the firms react to the strategy of the other firms are all examples of such dynamic models. Many of these have a Markovian structure so it is natural to look for equilibrium in Markov strategies which are strategies that depend only on the realized state in each period. We use Asynchronous Stochastic Games as the general framework to study such equilibria and examine the conditions under which Markov perfect equilibrium exists. We show that under some regularity conditions there exist Markov-perfect equilibrium in pure strategies when the game is a finite horizon game. The infinite horizon game has a Markov perfect equilibrium in randomized strategies. If there are a finite number of players and the game has a fixed cycle, then the game has a stationary Markov-perfect equilbrium. If the state space is an atomless measure space, and satisfies the condition of coarser transition kernel, then the game has a Markov perfect equilibrium in pure strategies. We show that the results can be applied to some well-known dynamic models like the dynamic oligopoly models and the extraction of common property resources.

Key Words: Stochastic games, Markov-Perfect equilibrium, Stationary markov equilibrium, Subgame perfect equilibrium, R & D games, Dynamic market competition, Stationary equilibrium, Dynamic economic models.

JEL Classification: C7, C62, C72, C73, C61, D81.

[1] The paper has benefited from presentations at the Midwwest Economic Theory Conference, October 2018 and seminars at Vanderbilt University and University of Cincinnati, April 2019. The author thanks the participants at these presentations for comments and suggestions but would especially like to thank Andrew Daughety, Jennifer Reinganum, Santanu Roy, Rene Saran, Iryna Topolyan, John Weymark and Myrna Wooders. The usual disclaimer applies.

1 Introduction

Asynchronous dynamic models describe situations in which the agents choose actions in a sequence with a single agent choosing in any one period. The well known Stackelberg model is one example. Other examples of asynchronous dynamic models arise in negotiations and bargaining where offers and counter-offers are made in successive rounds, Auctions in which bidders bid one after another, as well as in dynamic oligopoly models like the ones discussed by Maskin and Tirole in [15] and [16]. Principal Agent models are also asynchronous dynamic models as the offer of a principal precedes the actions taken by an agent. Some networks also have this feature. Thus there is a fairly large class of dynamic models in economics that are Asynchronous.

Here we study asynchronous dynamic models within the general framework of Asynchronous Stochastic games as these have a structure that is rich enough to allow the modelling of uncertainty and noise. We address two interrelated questions; the first is about the nature of the solution for these games and the second is about the conditions under which a solution can be found and the characteristics of the solution. A resolution of the first question can be had by suggesting that given the wide use of a noncooperative solution like Nash equilibrium and/or subgame perfect equilibrium, a closely related concept which is appropriate for this class of games, could be used as the equilibrium concept. Therefore, we focus our attention on Markov perfect equilibrium strategies as this seems to fit most closely the structure of the games that we want to discuss. The second set of issues deal with the conditions under which such an equilibrium would exist and the characteristics of these equilibrium.

A stochastic game is one in which players choose actions in each period, and these actions then determine the distribution of the states next period. The single-period payoffs depend on both the realized state and the actions of the players. The payoff of a player in the entire game is the discounted sum of the single-period payoffs. The distribution of next period's states are given by a transition probability function which depends on both the current state and the current actions. Therefore, in each period a player has to choose an action that maximizes the sum of the single-period payoff and the discounted expected future payoff. In a general stochastic game, the single-period payoff depends on the actions taken by the player as well as the actions taken by the other players in that period. In an Asynchronous stochastic game, the single-period payoff depends on the action taken by the player and the realized state. The strategic interactions between the players therefore occur through the effect that a player has on the payoffs of the players who choose in the subsequent periods, through the effect of the player's action on the future states.

As these are stochastic games with the single-period payoff depending on the realized state and actions of that period, and in which the transition probability is a function of only the current state and actions, it is useful to ask whether an equilibrium exists in which the strategies of the players depend only on the current state. We therefore look for Markov perfect equilibrium in which the strategies of the players are functions of only the current state, but are nonetheless perfect equilibrium in the sense that the equilibrium strategy of the player is a best response to the strategies of the other players, and this is so for every realization of the state in each period. We find that if the single-period payoffs are measurable in the state variable and continuous in the actions, and the transition probability is measurable in the state variable is norm-continuous in actions, then a finite horizon game has a pure strategy Markov perfect equilibrium. If the game is an infinite horizon game then there ia a Markov perfect equilibrium in randomized strategies. We also find that if the game has a finite number of players and the players choose in a fixed cycle, then the game has a stationary Markov-perfect equilibrium. If the state space is a non-atomic measure space and the transition probability has a decomposable coarser transition kernel (a condition introduced in He and Sun [11]) then the game has a Markov-perfect equilibrium in pure strategies.

Quite a large literature has examined the question of the existence of a stationary Markov-perfect equilibrium for general stochastic games in which a number of players simultaneously choose actions in each period. This is a natural starting point, as a stationary Markov-perfect equilibrium is an attractive solution of a stochastic games as it reflects the stationary nature of such a game as well as its Markovian structure, as the transition probability is a function of only the current state and actions. However, the examples in Levy [13] and Levy and McLennan [14] show that under the condition of just norm-continuity and absolute continuity of the transition probability function, stationary Markov equilibrium do not exist. These examples thus indicate that additional conditions are needed to guarantee the existence of a stationary Markov equilibrium. He and Sun [11] shows that a stationary Markov equilibrium exists if the transition probability function satisfies the additional condition of *coarser transition kernel*. Duggan [7] also shows that a stationary Markov perfect equilibrium exists when the stochastic game is "noisy," that is the state space has a non-atomic component that is not directly affected by the previous period's state and actions. Although the condition in He and Sun [22] implies the condition of "noise" in Duggan [7], the condition in Duggan [7] is itself interesting as it provides a useful interpretation of the condition in He and Sun [11].

Given the results in the literature in stochastic games and given the fact that players choose actions simultaneously in each period, a natural question to raise is how the results would be affected if the players chose actions in a sequence rather than simultaneously, and the implications for the application of these results. As Asynchronous stochastic games have a different structure from general stochastic game since only a single player chooses in each period, the results for general stochastic games do not apply directly. This is true in particular about the example of Levy [13] and of Levy and McLennan [14]. The paper is laid out as follows. In section 2 we describe the general model. In section 3 we discuss payoffs and strategies. In section 4 we discuss the results on the existence of Markov-perfect equilibrium for the finite and infinite horizon games. In section 5 we discuss the purification result under the condition that the state space is a non-atomic measure space and the transition probability has a decomposable coarser transition kernel. In each section we present an application to which the result can be applied directly. In section 7 we conclude.

2 The Model

In an Asynchronous stochastic game in any period there is only one player who chooses from an action set. Thus an Asynchronous stochastic game is the following tuple:
$[(S, \Sigma, \nu_t), (u_{it})_{i=1}^n, A_t, q_t]_t$ where

(i) (S, Σ, ν_t) is the *State Space* in period t. S is a *complete, separable metric space*, Σ is the Borel sigma algebra and ν_t is a *probability measure.*

(ii) A_t is the action set in period t. It is a *compact metric space.*

(iii) $u_{it} : S \times A_t \to \mathbb{R}$, is the single period payoff function of player i in period t. It depends on the current state and the current action. We will assume that u_{it} is *measurable* on S, and *continuous* on A_t. Further, there is a $K > 0$ such that for all i and t, $u_{it}(s, a) \leq K$ for all (s, a).

We will use the notation $\mathcal{M}(S)$ to denote the space of bounded measurable functions and $\mathcal{P}(S)$ to denote the space of probability measures on S. We note that the set $\mathcal{M}(S)$ can be viewed as a set of linear functionals on $\mathcal{P}(S)$ [1]. The *weak* topology on $\mathcal{P}(S)$ denoted by $\sigma(\mathcal{P}(S), \mathcal{M}(S))$ is the weakest topology which makes these linear functionals continuous. We will denote this dual topology by σ_c^*. We will denote the resulting topological space by $(\mathcal{P}(S), \sigma_c^*)$. (Chapter 3 of Rudin [1973] has a detailed discussion of the *weak* and the *weak** topologies).

(iv) $q_t : S \times A \to (\mathcal{P}(S), \sigma_c^*)$ is the *transition probability* in period t. We will assume that the transition probability is product measurable on $S \times A$ and continuous on A.

For every $B \in \Sigma$, the transition probability q_t describes a function $q_t(B|.,.) : S \times A \to [0,1]$ which gives the measure of B as a function of (s,a). The product measurability condition implies that for every $B \in \Sigma$ the function $q_t(B|.,.) : S \times A \to [0,1]$ is a jointly measurable function. Notice that we described the transition probability on the state space as depending only on the immediate past. Therefore, in the definition of a transition probability we have implicitly assumed that the transition probability is Markovian.

3 Payoffs and Strategies

A history up to period t is a sequence of actions $(a_1, \cdots, a_t)$ chosen by the players from periods 1 through t. We will denote such a history by h_t and denote the set of histories up to period t by H_t.

Definition 1 *A behavior strategy of a player i is a sequence $\{f_{it}\}_t$ such that for all t $f_{it} : H_{t-1} \times S \to \mathcal{P}(A_t)$ is a measurable function for each t at which player i's action set is A_t[2], where $\mathcal{P}(A_t)$ has the topology of weak convergence of probability measures[3] and the measurable subsets of $\mathcal{P}(A_t)$ are the Borel sets.*

[1] For any $f \in \mathcal{M}(S)$ one may define the functional $\Lambda_f : \mathcal{P}(S) \to \mathbb{R}$ as $\Lambda_f(\nu) = \int_S f(s)\nu(ds)$.

[2] We note that since the game is an Asynchronous stochastic game, at each period there is only one player for whom the action set is A_t.

[3] The topology of weak convergence of probability measures is the $\sigma(\mathcal{P}(A_t), \mathcal{C}(A_t))$ dual topology on $\mathcal{P}(A_t)$ where $\mathcal{C}(A_t)$ is the set of all continuous functions on A_t. For detailed discussions of this topology one may refer to Billingsley [?] and Parthasarathy [17]

We will focus on a special class of strategies, namely *Markov strategies.*

Definition 2 *A* **Markov strategy** b_i *is a sequence* $\{b_{it}\}$ *such that* $b_{it} : S \to \mathcal{P}(A_t)$ *is measurable for every* t. *A* **Markov strategy combination** $b = (b_i)_{i \in N}$ *is a combination of Markov strategies.*

Clearly, "Markov" strategies restrict players to make their choice of action in each period as a function of only on the current state. This can be a fairly severe restriction on the kind of strategies players can use. However, with the assumptions we have made about the Markovian nature of the transition probability, a player can do just as well by using a Markov strategy. This is so because the current and future payoff of a player is given by

$$u_{it}(s, a) + \delta \int_S v_{i(t+1)}(b)(s')q_{t+1}(ds'; s, a). \tag{1}$$

Given a Markov strategy combination $b = (b_1, \cdots, b_n)$, the future expected payoff of player i in period t is given by

$$\begin{aligned}
&v_i(b)(s) = \int_S \int_{A_{t+1}} u_{i(t+1)}(s_{t+1}, a)b_{i(t+1)}(s_{t+1})(da))q(ds_t, a_t) + \\
&\delta \int_S \int_{A_{t+1}} [\int_S \int_{A_{t+2}} u_{i(t+2)}(s_{t+2}, a)b_{i(t+2)}(s_{t+2})(da)q(ds_{t+1}, a_{t+1})]b_{i(t+1)}(s_{t+1}(da))q(ds_t, a_t) + \cdots \\
&+\delta^{T-1} \int_S \int_{A_{t+1}} \cdots [\int_S \int_{A_T} u_{iT}(s_T, a)b_{iT}(s_T)(da)q(ds_{T-1}, a_{T-1})] \cdots b_{i(t+1)}(s_{t+1})(da)q(ds_t, a_t) + \\
&\sum_{\ell=T+1}^{\infty} \delta^{\ell-1} \int_S \int_{A_{t+1}} \cdots [\int_S \int_{A_\ell} u_{i\ell}(s_\ell, a)b_{i\ell}(s_\ell)(da)q(ds_{\ell-1}, a_{\ell-1})] \cdots b_{i(t+1)}(s_{t+1})(da)q(ds_t, a_t) \quad (2)
\end{aligned}$$

We note that since $u_{it}(s, a) \leq K$ for all (s, a), for any $\epsilon > 0$, there is a T sufficiently large, such that, the sum of the terms from period T onwards satisfies

$$\begin{aligned}
&\sum_{\ell=T+1}^{\infty} \delta^{\ell} \int_S \int_{A_{t+1}} \cdots [\int_S \int_{A_\ell} u_{i\ell}(s_\ell, a)b_{i\ell}(s_\ell)(da)q(ds_{\ell-1}, a_{\ell-1})] \cdots b_{i(t+1)}(s_{t+1})(da)q(ds_t, a_t) \\
&\leq \frac{\delta^{T-1}K}{1-\delta} < \epsilon
\end{aligned}$$

for $0 < \delta < 1$. Therefore, the expected future payoff given in (2) converges to a limit for each $s \in S$. We further note that as b is a Markov strategy each $b_{it} : S \to \mathcal{P}(A_t)$ is measurable. As the transition probability function $q : S \times A_t \to \mathcal{P}(S)$ is measurable on S, each term in (2) is a measurable function of s_t. Therefore, the expected future payoff of player i from the behavior strategy combination $b = (b_1, \cdots, b_n)$, given by $v_i(b) : S \to \mathbb{R}$ is a measurable function on S.

The payoff of a player i in period t is then given by

$$u_{it}(s_t, a_t) + \delta \int_S v_i(b)(s')q(ds'; s_t, a_t) \tag{3}$$

where $u_{it}(s_t, a_t)$ is the single-period payoff and $\delta \int_S V_i(b)(s')q(ds'; s_t, a_t)$ is the expected future payoff from the Markov strategy combination $b = (b_1, \cdots, b_n)$. We now note that if the players use a Markov strategy combination, the player who makes the choice in period t can choose an action in A_t to maximize his payoff simply as a function of the realized state $s_t \in S$. That is, the player cannot do any better by using a general behavior strategy than the optimal Markov strategy. Therefore, if every other player uses a Markov strategy, then *an optimal* Markov strategy of player i in period t is an optimal strategy for period t. Hence an equilibrium in Markov strategies then is an equilibrium of the game.

Definition 3 *A Markov strategy combination b^* is a* **Markov perfect equilibrium** *if for any s_t in any period t and for every player i*

$$v_{it}(b^*)(s_t) \geq v_{it}(b^*_{-i}, b_i)(h_{t-1}, s_t)$$

for any behavior strategy b_i (not necessarily Markovian).

It is of interest to note that when one considers deviations from the Markov perfect equilibrium strategy one allows a deviation using a general behavior strategy.

4 Markov Perfect Equilibrium for Asynchronous Stochastic Games

We present two results for Asynchronous stochastic games. In the first we show that there exists a Markov-perfect equilibrium in pure strategies when the game has a finite horizon. In the second result we show that there exists a Markov-perfect equilibrium for an infinite horizon game.

4.1 Finite Horizon Game

The proof will use a backward induction argument of the kind that is frequently used in analyzing dynamic models. We have already observed that the expected payoff of a player i in period t is given by

$$u_{it}(s, a_t) + \delta \int_S v_{i(t+1)}(s')q(ds'; s, a_t)$$

where $v_{i(t+1)} : S \to \mathbb{R}$ is the expected future payoff of player i from period $t+1$ onwards.

The next result shows that a pure strategy Markov-perfect equilibrium exists for a finite horizon game. A finite horizon game is one which has a terminal period T.

Theorem 1 *A finite horizon Asynchronous stochastic game has a pure strategy Markov-perfect equilibrium.*

Proof: The proof uses the familiar backward induction argument. In period T, the player who chooses in period T, chooses $a_t \in A_T$ to maximize the single-period payoff $u_T : S \times A_T \to \mathbb{R}$. Let $m_t : S \to \mathbb{R}$ be given by

$$m_T(s) = \text{argmax}_{a \in A_T} u_T(s, a)$$

and the correspondence $M_T : S \to A_T$ be defined as

$$M_T(s) = \{\hat{a}_T \in A_T \mid u_T(s, \hat{a}_T) = m_T(s)\}.$$

As $u_T : S \times A_T \to \mathbb{R}$ is measurable on S and continuous on A_T it is jointly measurable on $S \times A_T$. Then, by the measurable maximum theorem (see for example Theorem 18.19 of Aliprantis and Border [1]), the correspondence $M_T : S \to A_T$ is measurable and has a measurable selection $\hat{f}_T : S \to A_T$ such that $\hat{f}_T(s) \in M_T(s)$ for all $s \in S$. That is, there is an optimal choice function $b^\star_T : S \to A_T$ given by $b_T(s) = f_T(s)$ for $s \in S$ in period T.

Now in period $T-1$, given the optimal choice function $b^\star_T : S \to A_T$ the payoff of the player who chooses in period $T-1$ is given by

$$u_{T-1}(s, a) + \delta \int_S v_{T-1,T}(b^\star_T)(s')q(ds'; s, a) \tag{4}$$

where $v_{T-1,T}(b^\star_T)(s') = u_{T-1,T}(s', b^\star_T(s'))$ and $u_{T-1,T} : S \times A_T \to \mathbb{R}$ is the single-period payoff in period $T-1$ of the player who chooses in period $T-1$. Since $u_{T-1,T} : S \times A_T \to \mathbb{R}$ is measurable on S and continuous on A_{T-1}, it is jointly measurable on $S \times A_{T-1}$. Hence,, as $b_T^\star : S \to A_T$ is measurable, the function $v_{T-1,T}(b_T^\star) : S \to \mathbb{R}$ is measurable. We now claim that

$$\int_S v_{T-1,T}(b_T^\star)(s')q(ds'; s, a)$$

is continuous on A_{T-1} and measurable on S. The first part of the claim follows from noting that the transition probability $q : S \times A_{T-1} \to \mathcal{P}(S)$ is continuous on A in the $\sigma_c^\star$ topology on $\mathcal{P}(S)$. The second part of the claim follows from the fact that the transition probability is measurable on S. Therefore, the payoff function of the player who chooses in period $T-1$ given by (4) is measurable on S and continuous on A_{T-1}. Let

$$m_{T-1}(s) = \text{argmax}_{a \in A_T}[u_{T-1}(s, a) + \delta \int_S v_{T-1,T}(b_T^\star)(s')q(ds'; s, a)]$$

and the correspondence $M_{T-1} : S \to A_{T-1}$ be defined as

$$M_{T-1}(s) = \{\hat{a}_{T-1} \in A_{T-1} \mid [u_{T-1}(s, \hat{a}_{t-1}) + \delta \int_S v_{T-1,T}(b_T^\star)(s')q(ds'; s, \hat{a}_{t-1})] = m_{T-1}(s)\}.$$

This correspondence is a measurable correspondence and has a measurable selection $\hat{f}_{T-1} : S \to A_{T-1}$ such that $\hat{f}_{T-1}(s) \in M_{T-1}(s)$ for all $s \in S$. Define the optimal choice function of the player who chooses in period $T-1$ as $b_{T-1}^\star(s) = \hat{f}_{T-1}(s)$ for all $s \in S$.

One can now find $b_{T-2}^\star : S \to A_{T-2}$ which is optimal for the player who chooses in period $T-2$, given that the players who choose in periods $T-1$ and T use the choice functions $b_{T-1}^\star$ and $b_T^\star$ respectively. Proceeding in this manner one can finally obtain the function $b_1^\star$, that gives the optimal choice function for period 1 given that the optimal choice functions in the following periods are given by $(b_2^\star, \cdots, b_{T-1}^\star, b_T^\star)$.

It should be clear from the construction of the sequence of functions $(b_1^\star, \cdots . b_T^\star)$ that in each period t, $b_t^\star$ is optimal given that in the following periods the choices are given by the functions $(b_{t+1}^\star, \cdots . b_T^\star)$. But this shows that $(b_1^\star, \cdots . b_T^\star)$ is a Markov-perfect equilibrium of the finite horizon Asynchronous stochastic game. We also note that the strategies of the players are pure strategies. ■

4.2 The Generalized Stackelberg Model

As in the Stackelberg Model there are two firms that produce an output that are close substitutes. The demand for firm 1 is given by $p_1 = f_1(q_1, q_2)$ and the demand for firm 2 is given by $p_2 = f_2(q_1, q_2)$. In period 1, firm 1 chooses an output q_1 which then determines the distribution of the prices $p_1(q_1, .)$ and $p_2(q_1, .)$ of the two firms. That is the demand functions in period 2 depend stochastically on the choice q_1 of firm 1 in period 1. Given the realizations of the demand functions in period 2, firm 2 chooses an output level q_2. The payoffs of the firms are then the profits of the firms in period 2. This leads to the following two-period Asynchronous stochastic game.

(i)Let C denote the set of non-increasing continuous functions from $[0, Q] \to [0, P]$ endowed with the sup-norm topology. Then the state S in period 2 is $C \times C$, the set of possible demand functions of the two firms. Then S is a complete separable metric space. Let μ be a probability measure on S which can be non-atomic.

(ii) Let $q(.|q_1)$ be the transition probability function. We will assume that it is norm-continuous in q_1 and absolutely continuous with respect to the measure μ on S.

(iii) $u_1(q_1) = 0$, $u_2(q_1) = 0$ as neither firm receives a payoff in period 1.

(iv) $u_1(s, q_2) = p_1(q_1, q_2)q_1 - c_1(q_1)$ and $u_2(s, q_2) = p_2(q_1, q_2)q_2 - c_2(q_2)$, are the profits of the two firms respectively, given the realized state $s = \{p_i(q_1, q_2)\}_i = 1^2$ of firm $i = 1, 2$. Here $c_i(q_i)$ is the cost function of the firm $i = 1, 2$ which are continuous and increasing in q_i.

Proposition 1 *The two-period Asynchronous game of the Generalized Stackelberg model has a Markov-perfect equilibrium in pure strategies.*

Proof: This follows from Theorem 1 as the two-period asynchronous stochastic game satisfies all the conditions of theorem 1. ∎

The Markov-perfect equilibrium here is a subgame perfect equilibrium although it appears to be different from the usual subgame perfect equilibrium in the Stackelberg model. The optimal choice of firm 2 does not depend on the output q_1 directly. But firm 1 does control the payoff of firm 2 through the transition probability that determines the realization of the demand function of firm 2 in period 2. In the usual Stackelberg model this is determined precisely as the model is deterministic; here the actual demand is determined stochastically. Because of that firm 2 does not respond directly to the output of firm 2 but only to the actual demand that is realized.

4.3 The Infinite Horizon Game

We start this section with some preliminaries. Let

$$M := \{f : S \to \mathbb{R} \mid f \quad \text{is measurable and} \quad |f| \leq K\}. \tag{5}$$

Since $L_\infty(S,\nu_t)$ can be identified with $L_1^*(S,\nu_t)$, the dual space of $L_1(S,\nu_t)$ (see Dunford and Schwartz [1988] theorem 5, page 289) we can consider M as a subset of $L_\infty(S,\nu_t)$ and topologise it with the relative $\sigma(L_\infty(S,\nu_t), L_1(S,\nu_t))$ topology that it inherits as a subset of $L_\infty(S,\nu_t)$. We will denote the topological space $(M, \sigma(L_\infty(S,\nu_t), L_1(S,\nu_t))$ by M_t and the dual topology $\sigma(L_\infty(S,\nu_t), L_1(S,\nu_t))$ by σ_t^∞.

Remark: The dual topology σ_t^∞ is the same as the *weak** topology of $L_\infty(S,\nu_t)$ and $f_n \to f$ in this topology if and only if $\int_S f_n(s)g(s)\nu_t(ds) \to \int_S f(s)g(s)\nu_t(ds)$ for every $g \in L_1(S,\Sigma,\nu_t)$.

The next result is important enough to state as a lemma.

Lemma 1 *M_t is a compact metric space.*

Proof: Since M_t is the K-ball in $L_1^*(S,\nu_t)$ by the Banach-Alaoglu theorem (see theorem 3.15, page 66 of Rudin [1973]) M_t is compact in the σ_t^∞ topology. Since S is a separable metric space, Σ is the Borel sigma algebra and ν_t is a probability measure, therefore, the metric space of measurable sets is a separable metric space (see theorem B page 168 of Halmos [1974]). But this implies that $L_1(S,\nu_t)$ is separable. Therefore, M_t is metrizable (see theorem 3.16 page 68 of Rudin [1973]). Hence, M_t is a compact metric space. ∎

In section 2 we had defined the σ_c^* topology on $\mathcal{P}(S)$. We now define a stronger topology on $\mathcal{P}(S)$, the *norm topology*[4] which we denote by $||.||$. $\nu_n \to \nu$ in $||.||$ if $\nu_n(E) \to \nu(E)$ uniformly for every $E \in \Sigma$ and we will write this as $||\nu_n - \nu|| \to 0$. The next assumption is a strengthening of the assumption on transition probabilities. Compare this with (iv) in section 2.

Assumption 1 *The transition probability $q_t : S \times A \to (\mathcal{P}(S), ||.||)$ is absolutely continuous with respect to the probability measure ν_t and continuous on A.*

4.4 Properties of the Payoff function in a period t and a Recursion Result

In each period t, given the Markov strategy combination that is to be used in period $t+1$ onwards, there is an expected future payoff $v_{i,t+1} : S \to \mathbb{R}$ which is measurable on S, (see for example (2)). The payoff function of a player i in period t, is then given by the function $V_{i,t} : S \times A \times M_t \to \mathbb{R}$ defined as

[4]For $\mu \in \mathcal{P}(S)$, $||\mu|| = \sup_{|f|\leq 1} |\int_S f d\mu|$ is a norm that yields the norm topology.

$$V_{i,t}(s,a,v_{i,t+1}) := u_i(s,a) + \delta_i \int_S v_{i,t+1}(s')q_t(ds';s,a). \tag{6}$$

where $v_{i,t+1} : S \to \mathbb{R}$ is a measurable function that gives the future expected payoff of player i.

Lemma 2 *The function $V_{i,t} : S \times A \times M_t \to \mathbb{R}$ satisfies the following properties:*

(i) *For every $f \in M_t$, $V_{i,t}(.,.,f) : S \times A \to \mathbb{R}$ is measurable on S and continuous on A_t.*

(ii) *For every $s \in S$, $V_{i,t}(s,.,.) : A \times M_t \to \mathbb{R}$ is continuous.*

Proof: The first conclusion follows directly from assumption 1 and the fact that $f : S \to \mathbb{R}$ is measurable. To check that $V_{i,t}$ is jointly continuous on $A \times M_t$ it is enough to show that if we take a sequence $(a_k, f_k) \to (a,f)$ in the metric space $A \times M_t$ then $V_{i,t}(s,a_k,f_k) \to V_{i,t}(s,a,f)$. Since u_i is continuous on A_t, we need only show that $\int_S f_k(s')q_t(ds';s,a_k) \to \int_S f(s')q_t(ds';s,a)$ as $(a_k,f_k) \to (a,f)$. We have,

$$\begin{aligned} |\int_S f_k(s')q_t(ds';s,a_k) - \int_S f(s')q_t(ds';s,a)| &\leq |\int_S f_k(s')q_t(ds';s,a_k) - \int_S f_k(s')q_t(ds';s,a)| \\ &+ |\int_S f_k(s')q_t(ds';s,a) - \int_S f(s')q_t(ds';s,a)|. \end{aligned} \tag{7}$$

Now the norm continuity of the transition probability q_t (see assumption 1) implies the following

$$|\int_S f_k(s')q_t(ds';s,a_K) - \int_S f_k(s')q_t(ds';s,a)| \leq \int |f_k|||q_t(.|s,a_k) - q_t(.|s,a)|| \leq K\epsilon \tag{8}$$

for every $f_k \in M_t$ and all $k \geq k_o$.

Also, if $g \in L_1(S,\nu_t)$ is the Radon-Nikodym derivative of $q_t(.|s,a)$ with respect to ν_t, then as $f_k \to f$ in the σ_t^∞ topology, we have

$$\begin{aligned} |\int_S f_k(s')q_t(ds';s,a) - \int_S f(s')q_t(ds';s,a)| &= |\int_S f_k(s')g(s')\nu_t(ds') - \int_S f(s')g(s')\nu_t(ds')| \\ &\leq \epsilon \end{aligned} \tag{9}$$

for all $k \geq k_1$. Therefore, for all $k \geq \max\{k_o, k_1\}$,

$$|\int_S f_k(s')q_t(ds';s,a_k) - \int_S f(s')q_t(ds';s,a)| \leq (K+1)\epsilon.$$

This shows that $|V_t(s,a,f) - V_t(s,a_k,f_k)| \to 0$ as $k \to \infty$ and completes the proof.[5] ■

Let M_t^n denote the n-product space $M_t \times \cdots \times M_t$ endowed with the product topology[6]. For $f \in M_{t+1}^n$, where $f = (f_1, \cdots, f_n)$ the payoff of the player who chooses in period t maximizes the payoff function

$$V_t(s, a_t, f_t) = u_t(s_t, a_t) + \delta \int_S f_t(s') q(ds'(s_t, a_t).$$

As V_t is measurable on S and continuous on A_t, the correspondence B_t from S into A_t that gives the optimal choice of the player has a measurable graph which is compact-valued and admits a measurable selection (see for example Aliprantis and Border [1], Theorem 18.19).

When mixed strategies are used, the action set in any period is given by the set of probability distributions on A_t which we denote by $\mathcal{P}(A_t)$. The topology on this set is the topology of weak convergence of probability measures so that $\mathcal{P}(A_t)$[7]. The correspondence $B_t^C : S \to \mathcal{P}(A_t)$ given by

$$B_t^C(s) = Co.B_t(s)$$

for $s \in S$, where $Co.B_t(s)$ is the convex hull of $B_t(s)$. As the correspondence B_t is a measurable correspondence which is compact-valued and nonempty-valued, the correspondence B_t^C is a measurable correspondence that is nonempty-valued and compact-valued. We now note that for each $\mu_t \in Co.B_t(s)$, the payoff of a player i (this would be any player i not just the player who chooses in period t) is given by

$$\int_{A_t} V_{i,t}(s, a_t, f_i) \mu_t(da_t).$$

As $V_{i,t}(s, a_t, f_i)$ is continuous on A_t, the function $\int_S V_{i,t}(s, a_t, f_i)\mu_t(da_t)$ is continuous in $\mu_t \in \mathcal{P}(A_t)$. Therefore for any measurable selection $b_t : S \to \mathcal{P}(A_t)$, the payoff function

[5] The second part of the result here that the function $V_t : S \times A_t \times M_t$ is jointly continuous on $A_t \times M_t$ is where the stronger condition that the transition probability is norm-continuous (assumption 1) is used.

[6] Actually for every player $i = 1, \cdots, n$ the set M_t is different as the $K_i's$ which are used as the uniform bounds could be different. Since the rest of the discussion is unaffected by whether we take note of this or not, for notational simplicity we will assume that the $K_i's$ are all equal.

[7] In the definition of a behavior strategy (see definition 1) we topologised $\mathcal{P}(A_t)$ with the topology of weak convergence of probability measures. Since A_t is a compact metric space $\mathcal{P}(A_t)$ is also a compact metric space in this topology see e.g. theorem 6.4 on page 45 of Parthasarathy [1967].

$\int_S V_{i,t}(s, a_t, f_i)b_t(s)(da_t)$ is a measurable function of $s \in S$. For any $f \in M^n_{t+1}$ define the correspondence $P_t : S \to \mathbb{R}^n$ as

$$P_t(f)(s) = \{[\int_{A_t} V_{i,t}(s, a_t, f_i)b_t(s)(da_t)]^n_{i=1} \mid b_t : S \to \mathcal{P}(A_t) \text{ is measurable and } b_t(s) \in B^C_t(s)\}.$$

Thus, $P_t(f)(s)$ therefore is the set of payoff vectors of the players in period t when player t maximizes the payoff in period t, given that the future expected payoff vector is $f = (f_1, \cdots, f_n)$.

We now define the correspondence $\Psi_t : M^n_{t+1} \to M^n_t$ as

$$\Psi_t(f) := \{g : S \to \mathbb{R}^n \mid g_t(s) \text{ is in } P_t(f)(s)\, a.e.\nu_t\}. \tag{10}$$

This correspondence gives the set of payoff functions of the players in period t, when the player choosing in period t chooses optimally, given the expected payoff from period $t+1$ onwards. The argument we made above shows that the correspondence is nonempty-valued. We show here that this correspondence is well behaved. Before we go into the result we introduce some notation. For a correspondence $P : S \to \mathbb{R}^n$ we will denote the set of integrable a.e. selections of P by $\tilde{S}(P)$. Formally, if $L_1(S,\mathbb{R}^n)$ is the set of ν integrable functions from S to $\mathbb{R}^n$, then

$$\tilde{S}(P) = \{\sigma \in L_1(S,\mathbb{R}^n) : \sigma(s) \in P(s)\ a.e.\ \nu\}.$$

The integral of a correspondence P will be written as $\int_S P(s)\nu(ds)$ and is defined as

$$\int_S P(s)\nu(ds) = \{x \in \mathbb{R}^n : x = \int_S \sigma(s)\nu(ds), \sigma \in \tilde{S}(P)\}.$$

This is the Aumann integral of the correspondence (see Klein and Thompson [1984], page 185).

Lemma 3 $\Psi_t : M^n_{t+1} \to M^n_t$ *is nonempty-valued and upper semi-continuous.*

Proof: Let $f_k \to f$ in M^n_{t+1}. Let the sequence $\{g_k\}$ in M^n_t be such that $g_k \in \Psi_t(f_k)$ for every $k \in \mathbb{N}$ and $g_k \to g$ in M^n_t. We need to show that $g \in \Psi_t(f)$.

From lemma 2, for every $s \in S$, the payoff function $V_t(s, ., .) : A \times M^n_{t+1} \to \mathbb{R}$ is jointly continuous on $A \times M^n_{t+1}$. Therefore, for every $s \in S$ the correspondence $P_t(.)(s)$, which takes M^n_{t+1} to the set of optimal payoffs of the player who chooses in period t, is an upper semi-continuous correspondence. The upper semi-continuity of $P_t(.)(s)$ implies that as $f_k \to f$ in M_{t+1}^n, for each $s \in S$, for any open set G where $P_t(s) \subset G$, there is a k_1 such that for all $k \geq k_1$, $P_t(f_k)(s) \subset G$. As M_t^n has the *weak** topology $g_k \to g$ in the *weak** topology. Therefore, for almost every $s \in S$ there is a convex combination $co.g_k(s)$ of $g_k(s)$ such that $co.g_k(s) \to g(s)$. Hence, for almost every $s \in S$,

$$g(s) \in Co.P_t(s) = P_t(s).$$

This shows that $g \in \Psi_t(f)$. ■

We now go on to construct the main body of the proof. But before we do that we make the following observation.

Lemma 4 *If $C \subseteq L_\infty(\nu_{t+1})$ is compact, then $\Psi_t(C)$ is a compact subset of $L_\infty(\nu_t)$.*

Proof: We will show that a sequence $\{c_k\}$ in $\Psi_t(C)$ will have a convergent subsequence. Since $\{c_k\}$ is a sequence in $\Psi_t(C)$, therefore, there is a sequence $\{a_k\}$ in C such that $c_k \in \Psi_t(a_k)$ for every k. Because of the compactness of C and M_t^n, without loss of generality we may assume that $a_k \to a$ and $c_k \to c$ via a subsequence. Then from lemma 3 it follows that $c \in \Psi_t(a)$. But this implies that $c \in \Psi_t(C)$ and completes the proof. ■

The idea of the proof of the main result, which establishes the existence of a Markov perfect equilibrium, is to essentially ask what would happen if the infinite horizon stochastic game was terminated at some period T. The set of possible payoffs at the end of period T would then be the set M_T^n. Given this set of expected future payoffs, the possible set of equilibrium payoffs in period $T-1$ will be in the set $\Psi_{T-1}(M_T^n)$ and we know that this set is nonempty and a compact subset of $L_\infty(S, \nu_{T-1})$. In period $T-2$, the players knowing that they will play an equilibrium in period $T-1$, given that the expected future payoffs are in M_T^n will play assuming that the expected future payoffs are in $\Psi_{T-1}(M_T^n)$. This will result in expected payoffs in the set $\Psi_{T-2}\Psi_{T-1}(M_T^n))$. In period $T-3$ players will play on the assumption that the future plays are in $\Psi_{T-2}\Psi_{T-1}(M_T^n)$. This process continues till the initial period is reached when the players play on the assumption that the future payoffs lie in the set $\Psi_1\Psi_2\cdots\Psi_{T-1}(M_T^n)$. Now this is the set of equilibrium expected payoffs that would result if the game was terminated in period T. If, however, the game was terminated in period $T+1$, then the set of possible expected future payoffs when playing in period T will be in the set $\Psi_T(M_{T+1}^n)$. Similarly, if the game is terminated in period $T+2$, then the set of possible future payoffs in period T is the set $\Psi_T\Psi_{T+1}(M_{T+2}^n) \subset \Psi_T(M_{T+1}^n)$. Therefore, if the game is never terminated then the set of possible equilibrium payoffs in period T is the set $\cap_{k=1}^{\infty}\Psi_{T+k-1}(M_{T+k}^n)$. Since $\Psi_T\Psi_{T+1}\cdots\Psi_{T+k}(M_{T+k+1}^n)$ is a nonempty compact subset of $\Psi_T\Psi_{T+1}\cdots\Psi_{T+k}(M_{T+k-1}^n)$ the set $\cap_{k=1}^{\infty}\Psi_{T+k-1}(M_{T+k}^n)$ is nonempty. A strategy combination that lead to these payoffs can be shown to be a Markov perfect equilibrium.

The following fact about the correspondences Ψ_t is of some importance

$$\Psi_t(M_{t+1}^n) \subset M_t^n \text{ for every } t \geq 1. \tag{11}$$

We will use the notation $\Psi_t^k(M_{t+k+1})$ to denote the set

$$\Psi_t\Psi_{t+1}\cdots\Psi_{t+k}(M_{t+k+1}).$$

Therefore the set $\Psi_t^{\infty} = \cap_{k=1}^{\infty}\Psi_t^k(M_{t+k})$ describes the set of expected future payoffs at time t that is generated by playing an equilibrium at every period in the future on the assumption that the players would continue to do the same.

Lemma 5 *The set Ψ_t^{∞} is nonempty for every t.*

Proof: Notice that $\Psi_t^{\infty} = \cap_{k=1}^{\infty}\Psi_t^k(M_{t+k+1}^n)$, where by lemma 6 $\Psi_t^k(M_{t+k+1})$ is a nonempty compact subset of M_{t+k}^n which is a compact metric space. From (22) and the obvious induction result, $\cap_{k=1}^{\infty}\Psi_t^k(M_{t+k+1}^n)$ is the intersection of a nested sequence of nonempty compact sets. Hence, by the finite intersection property for a collection of compact sets, the set $\cap_{k=1}^{\infty}\Psi_t^k(M_{t+k+1}^n)$ is nonempty. Since the argument can be made for any $t \geq 1$, we have the result. ∎

Lemma 5 gives us the following important result. There is a sequence of functions $\{\hat{x}_t\}_{t\in\mathbb{N}}$ such that $\hat{x}_t : S \to \mathbb{R}^n$ satisfying the condition that

$$\hat{x}_t \in \Psi_t(\hat{x}_{t+1}). \tag{12}$$

This is quickly seen by noting that since Ψ_1^{∞} is nonempty there is an $\hat{x}_1 \in \Psi_1^{\infty}$. But this means that there is an $\hat{x}_2 \in \Psi_2^{\infty}$ such that $\hat{x}_1 \in \Psi_1(\hat{x}_2)$, and so on. We now go on to show that there is a Markov perfect equilibrium strategy combination which will have payoffs in the set Ψ_t^{∞}. We construct the equilibrium strategy combination in the result that follows.

Theorem 2 *An infinite horizon Asynchronous stochastic game that satisfies assumption 1 has a Markov perfect equilibrium strategy combination.*

Proof: Consider the correspondence $\gamma_1 : S \to \mathcal{P}(A_1)$ given by

$$\gamma_1(s) = \{\hat{\mu}_1(s) \in \mathcal{P}(A_1) \mid \int_{A_1} V_1(s, a_1, \hat{x}_2^1)\mu_1(s) = \hat{x}_1^1(s)\}$$

where $\hat{x}_1^1(s)$ is the payoff of player 1, the player who chooses in period 1. As $V_1(s, a_1, \hat{x}_2^1)$ is continuous in a_1 and $\mathcal{P}(A_1)$ has the topology of weak convergence, $V_1(s, a_1, \hat{x}_2^1)$ is a Caratheodory function. Therefore, since $\hat{x}_1^1 : S \to \mathbb{R}$ is measurable, by Filippov's Implicit Function Theorem (see theorem 18.17 of Aliprantis and Border [1]), the correspondence γ_1 is measurable and measurable selection $g_1^* : S \to \mathcal{P}(A)$ such that $\int_A V_i(s, a, \hat{x}_2^1) g_1^*(da_1, s) = \hat{x}_1^1(s)$ for every $s \in S$. Thus $g_1^* : S \to \mathcal{P}(A_1)$ is an equilibrium strategy in period 1.

Similarly, for period 2, there is a strategy $g_2^* : S \to \mathcal{P}(A_2)$ such that $g_2^*(s)$ is an equilibrium strategy for period 2. We construct the Markov strategy combination $g^* = (g_1^*, g_2^*, \cdots)$ by the forward induction method outlined above. We claim that this is a Markov perfect equilibrium strategy combination of the infinite horizon Asynchronous stochastic game.

From the construction of g^*, for any T there is an $\hat{x}_{T+1} \in M_{T+1}^n$ such that

$$V_T(b_T, g_{-T}^*)(\hat{x}_{T+1}, s) \leq V_T(g^*)(\hat{x}_{T+1}, s) \tag{13}$$

where $V_T(b_T, g_{-T}^*)(\hat{x}_{T,T+1}, s)$ is as defined in equation (6) with $\hat{x}_{T,T+1}$ denoting the expected payoff of the player who chooses in period T from period $T+1$ onwards. b_T is any behavior strategy used by the player who chooses in period T. Therefore, from (13) it follows that, given an $\epsilon > 0$ for all T sufficiently large

$$\begin{aligned} v_T(b_T, g_{-T}^*)(s) &\leq \epsilon + V_T(b_T, g_{-T}^*)(\hat{x}_{T+1}, s) \\ &\leq \epsilon + V_T(g^*)(\hat{x}_{T+1}, s). \end{aligned} \tag{14}$$

But then taking the limit as $T \to \infty$ we get

$$\begin{aligned} v_T(b_T, g_{-T}^*)(s) &\leq \epsilon + \lim_{T\to\infty} V_T(g^*)(\hat{x}_{T+1}, s) \\ &= \epsilon + v_T(g^*)(s). \end{aligned} \tag{15}$$

Since $\epsilon > 0$ was arbitrary, equation (15) must hold for every $\epsilon > 0$. Hence,

$$v_T(b_T, g_{-T}^*)(s) \leq v_i(g^*)(s). \tag{16}$$

Since (16) holds for every behavior strategy b_T of the player who chooses in period T, and for every period T, g^* is a Markov perfect equilibrium. ■

4.5 Stationary Markov-Perfect Equilibrium in Finite Player Games

Theorem 2 shows that a Markov-perfect equilibrium exists for an asynchronous stochastic game but such an equilibrium need not be stationary. What we show here is that if there are only a finite number of players then there is a Markov-perfect equilibrium that is stationary in the sense that a player uses the same choice function in each period in which the player chooses the action. Recall that the choice function of a player in a period is the function that designates the choice of action as a function of the state.

Definition 4 *A strategy of a player i is a stationary Markov strategy if the choice function $f_{i,t} : S \to A_i$ is the same in every period t in which player i chooses the action.*

We will say that an asynchronous stochastic game has a **fixed cycle** if players $1, \cdots, n$ choose actions in a sequence $1, 2, \cdots, n$ followed again by the same sequence $1, 2, \cdots, n$.

Theorem 3 *An asynchronous stochastic game with a finite number of players and a fixed cycle has a stationary Markov-perfect equilibrium.*

Proof: We will first prove this for the case when there are only two players and the players alternate in choosing actions. Since the players alternate in choosing actions, the game has a fixed cycle. As the game has two players, the set of possible payoff functions is given by M^2. As it is either player 1 or player 2 who choose in any period, we denote by $\Psi_1 : M^2 \to M^2$, the correspondence $\Psi_t : M^2 \to M^2$ defined in (10) when player 1 chooses, and by $\Psi_2 : M^2 \to M^2$ when player 2 chooses. From lemma 3 we know that the correspondences Ψ_1 and Ψ_2 are non-empty valued and upper semicontinuous. By definition these correspondences are convex-valued.

Now define the correspondence $\Psi_{12} : M^2 \to M^2$ as follows.

$$\Psi_{12}(u) = \Psi_2(\Psi_1)(u).$$

It should be clear that this correspondence is nonempty-valued, convex-valued and upper semicontinuous. Therefore, by the Fan-Glicksberg theorem it has a fixed point $u^\star : S \to \mathbb{R}^2$. Since $u^\star$ is a fixed point of Ψ_{12}, $u^\star \in \Psi_2(\Psi_1)(u^\star)$. Therefore, there must exist a $v^\star \in \Psi_1(u^\star)$ such that $u^\star \in \Psi_2(v^\star)$.

Let $\gamma_1 : S \to \mathcal{P}(A_1)$ be given by

$$\begin{aligned}\gamma_1(s) = \{\mu_1^\star \in \mathcal{P}(A_1) \quad | \quad & \int_{A_1} [u_1(s,a) + \delta \int_S u_1^\star(s') q(ds' : s,a)] \mu_1^\star = v_1^\star, \\ & \int_{A_1} [u_2(s,a) + \delta \int_S u_2^\star(s') q(ds' : s,a)] \mu_1^\star = v_2^\star\}.\end{aligned}$$

From Filippov's implicit function theorem and using arguments similar to those used in the proof of theorem 2 it follows that there is a measurable selection $b_1^\star : S \to \mathcal{P}(A_1)$ of the correspondence $\gamma_1 : S \to \mathcal{P}(A_1)$.

Now let $\gamma_2 : S \to \mathcal{P}(A_2)$ be given by

$$\begin{aligned}\gamma_2(s) = \{\mu_2^\star \in \mathcal{P}(A_2) \quad | \quad & \int_{A_2} [u_2(s,a) + \delta \int_S v_1^\star(s')q(ds' : s,a)]\mu_1^\star = u_1^\star, \\ & \int_{A_2} [u_1(s,a) + \delta \int_S v_1^\star(s')q(ds' : s,a)]\mu_1^\star = u_1^\star\}.\end{aligned}$$

Again from Filippov's implicit function theorem there is a measurable selection $b_2^\star : S \to \mathcal{P}(A_2)$ of the measurable correspondence γ_2.

It now follows that the Markov strategy combination $(b_1^\star, b_2^\star)$ is a stationary Markov-perfect equilibrium of the asynchronous game in which players 1 and 2 alternate in choosing actions.

The arguments made for the case of two players can now be extended to n players with a fixed cycle by using the correspondence $\Psi_{12\cdots n} : M^n \to M^n$ defined as

$$\Psi_{12\cdots n}(u) = \Psi_n(\Psi_{n-1}(\cdots(\Psi_1)(u).$$

One can then show, using the same arguments as in the case of two players, that there is a stationary Markov-perfect strategy combination $(b_1^\star, \cdots, b_n^\star)$. ■

4.6 Dynamic Oligopoly Games

The dynamic oligopoly models discussed in Maskin and Tirole [15] and [16] are examples of asynchronous dynamic models. These analyze a class of alternating-move infinite-horizon models of duopoly and analyze the properties of a Markov-perfect equilibrium which require strategies to depend only on the actions to which the rival firm is currently committed. In [15] a natural monopoly is analyzed in which fixed costs are so large that only one firm can make a profit. In equilibrium only one firm is active and practices the quantity analogue of limit pricing. In [16] the firms take turns choosing prices and the equilibrium studied is a Markov-perfect equilibrium in which a firm's choice of price in any period depends only on the other firm's current price. The Markov-perfect equilibria indicate the presence of kinked demand curve equilibria as well as Edgeworth cycles.

The interesting phenomenon described by the Markov-perfect equilibria in the models studied by Maskin and Tirole indicate that these kind of equilibria explain how firms behave in many instances, and especially when firms react mostly to each others short-term behavior. In fact in many instances that is what firms would want to focus on if the other firm is making decisions based on short-term reactions of its rival. If a firm changes its price then it would require the firm to do that as soon as it can. The dynamic models in Maskin and Tirole are deterministic and a firm needs to commit to its decision over two periods.

Here we study the stochastic version of such models in which the action of a firm in any period affects the payoff of the firm which chooses in the following period, through the impact of the action of the firm on the realized demand functions in the next period. The stochastic version has some advantages over a deterministic model in some respects. First, a firm which chooses in a period reacts to the realized demand function and not directly to the choice of the other firm so one does not need a firm to commit to its action over two periods. Second, a firm using only the realized demand in the current period to decide its optimal choice fits in more naturally as that is more relevant to a firm's payoff than the past history. In the stochastic version of the dynamic oligopoly model the idea of a Markov-perfect equilibrium is seems very natural.

The action set of a firm i is a set A_i which is a compact subset of a metric space. This could be the set of prices, quantities and investments from which the firm chooses. The realized state variable s which is an element of the state space S is the state of demand for the two firms and is an element of a complete, separable metric space. There is a transition probability function $q : S \times A_i \to \mathcal{P}(S)$ that gives the probability distribution on the next period's states as a function of the realized state and actions of the firm that chooses actions in the current period. The single-period payoff function of firm i is given by the profit of the firm $\pi_i(s, a)$ in that period, where s is the realized state variable and a is the action chosen in that period. It is important to observe here that the realized state variable s contains all the relevant information about past actions that affects the current payoff. Thus if the demand of a firm is a function of the output levels chosen by the other firms as well as the output of the firm that chooses in the current period, then that information about the output levels of chosen by the other firms in the preceding period is embedded in the realized state variable s.[8]

[8]For example, in the simple setting where firms are involved in choosing the output in each period, if firm 1 is the firm that set the output in the previous period, then the realized demand equation is $p = (A(\lambda) - q_1) - q_2$ where $A(\lambda)$ indicates the random realization of the intercept term of realized demand equation. Note that the output chosen by firm 1 in the preceding period is part of the realized

Proposition 2 *Let ν be the fixed probability measure on the state space S¿ If the transition probability is absolutely continuous with respect to ν, and norm continuous in the actions a and measurable in the state variable s, then the dynamic oligopoly model has a stationary Markov perfect equilibrium.*

Proof: The result follows from Theorem 3. ∎

5 Pure Strategy Markov Equilibrium

The preceding result shows that a Markov perfect equilibrium exists for asynchronous stochastic games under the condition of norm continuity of the transition probability. Here we show that, if the measure ν is nonatomic and the transition probability satisfies the condition of *decomposable coarser transition probability* of He and Sun [11], then there is a Markov perfect equilibrium in pure strategies. The condition of coarser transition probability allows one to purify the equilibrium strategies by replacing the payoffs in the convex hull of the equilibrium payoffs by payoffs that are in the set of equilibrium payoffs. The analyses here draws substantially on the insights from the results in He and Sun [11].

In order to get to the result here we recall some concepts form the literature. Let ν be a probability measure on (S, Σ). Let $\mathcal{G}$ be a sub-σ-algebra of Σ. A set $D \in \Sigma$ of positive measure is said to be a $\mathcal{G}$ **atom** if the restricted σ-algebras $\mathcal{G}^D$ and Σ^D satisfy the condition that the strong completion of $\mathcal{G}^D$ is Σ^D. That is, the restricted σ-algebras $\mathcal{G}^D$ are essentially the same. Since σ algebras can be viewed as the information that one has about the realizations of the states, if a set $D \in \Sigma$ is a $\mathcal{G}$ atom, then the occurrence of the event D gives the same information whether one has the σ-algebra Σ or the σ-algebra $\mathcal{G}$. Then

Definition 5 *Let $\mathcal{G}$ be a sub-σ-algebra of Σ. A discounted stochastic game is said to have a* **coarser transition kernel** *if Σ has no $\mathcal{G}$-atom under ν and $q(.; s, a)$ is $\mathcal{G}$-measurable on S for each $s \in S$ and $a \in A_t$.*

The condition implies that as the transition probability does not provide full information

demand equation. Here of course we are assuming that the stochastic element of the demand equation is the intercept term. Typically the stochastic shock λ would depend on the output level chosen in the preceding period.

about the events in σ-algebra it provides less information about the realizations of the states than is provided by the events in Σ.

The purification result here uses a result of Dynkin and Evstigneev [5] about regular conditional expectations of correspondences. Let $C : S \to \mathbb{R}^n$ be a closed-valued, measurable correspondence which is integrably bounded (that is, there is measurable function $\phi : S \to \mathbb{R}^n$ such that $\|C(s)\| \leq \phi(s)$ for ν-almost all s). For a Σ-measurable selection f of C let $\mathbb{E}^\nu(f|\mathcal{G})$ denote the conditional expectation of g with respect to the sub-σ-algebra $\mathcal{G}$ of Σ.[9] We note that $\mathbb{E}^\nu(f|\mathcal{G})$ is $\mathcal{G}$ measurable, and is a measurable selection of the correspondence $C : S \to \mathbb{R}^n$. Let

$$\tilde{S}^{(\mathcal{G},\nu)}(C) = \{\mathbb{E}^\nu(f|\mathcal{G}) \ : \ f \text{ is a } \Sigma\text{-measurable selection of the correspondence } C\}.$$

Theorem 4 *(***Dynkin and Evstigneev***) If Σ has no $\mathcal{G}$-atom, then for any Σ-measurable ν-integrably bounded, closed-valued correspondence $C : S \to \mathbb{R}^n$,*

$$\tilde{S}^{(\mathcal{G},\nu)}(C) = \tilde{S}^{(\mathcal{G},\nu)}(Co.C).$$

If the transition probability satisfies the condition of coarser transition kernel, then the set of conditional expectations of integrable selections of C are the same as set of conditional expectations of integrable selections of the convex hull of C. As in He and Sun [11] this allows one to replace future expected payoff functions that have values in the convex hull of the set of equilibrium payoffs, with payoff functions that have values in the actual set of equilibrium payoffs.

Theorem 5 *An Asynchronous Stochastic Game with a non-atomic fixed measure ν has a Markov equilibrium in* **pure strategies** *if it has a decomposable coarser transition kernel. If the game has a finite number of players and a fixed cycle then there is a stationary Markov-perfect equilibrium in pure strategies.*

Proof: From lemma 5 and Theorem 2 it follows that the equilibrium payoffs are given by the payoff vectors $(\hat{x}_1, \cdots, \hat{x}_t, \cdots)$, where $\hat{x}_t : S \to \mathbb{R}^n$ satisfies the condition that

$$\hat{x}_t \in P_t(\hat{x}_{t+1}).$$

That is,

$$\hat{x}_t(s) = \max_{a \in A_t}\{u_t(s,a) + \int_S \hat{x}_{t,t+1}(s')q(ds'; s, a)\nu(ds'))\}. \tag{17}$$

[9]For a discussion of conditional expectation one may refer to [4].

Here, $\hat{x}_{t,t+1}(s')$ is the expected payoff of player t (the player who makes the choice in period t) in period $t+1$, and $q(ds'; s, a$ is the Radon-Nikodym derivative of the transition probability.

Recall that the correspondence $B_t(x'_{t,t+1}) : S \to A_t$ gives the optimal choices of the player who chooses in period t when the expected future payoff is given by $x'_{t,t+1} : S \to \mathbb{R}^n$. Let $\tilde{P}(x'_{t,t+1})(s)$ denote the payoff set of the players when the choices are in $B_t(x'_{t,t+1})(s)$. $\tilde{P}(x'_{t,t+1})(s)$ is thus set of payoffs when the player uses pure strategies. From the theorem of Dynkin and Evstigneev (Theorem 4), we have $\tilde{S}^{(\mathcal{G}.\nu)}(B_t) = \tilde{S}^{(\mathcal{G}.\nu)}(Co.B_t)$, so that there exists a Σ-measurable selection $x^\star_{t+1}$ of $\tilde{P}_{t+1}(\hat{x}_{t+2})$ such that $\mathbb{E}^\nu(x^\star_{t+1}|\mathcal{G}) = \mathbb{E}^\nu(\hat{x}_{t+1}|\mathcal{G})$. Therefore, as $q(. : s, a)$ is $\mathcal{G}$-measurable on S for each $s \in S$

$$\begin{aligned} \int_S \hat{x}_{t+1}(s')q(ds'; s, a)\nu(ds') &= \int_S \mathbb{E}^\nu(\hat{x}_{t+1}q(.; s, a)|\mathcal{G})(s')\nu(ds') \\ &= \int_S \mathbb{E}^\nu(\hat{x}_{t+1}|\mathcal{G})(s')q(ds'; s, a)\nu(ds') \\ &= \int_S \hat{x}^\star_{t+1}(s')q(ds'; s, a)\nu(ds'). \end{aligned} \tag{18}$$

Therefore, we have

$$\hat{x}_t(s) = \max_{a \in A_t}\{u_t(s, a) + \int_S x^\star_{t,t+1}(s')q(ds'; s, a)\nu(ds')\}. \tag{19}$$

As $x^\star_{t+1}(s) \in \tilde{P}_{t+1}(s)$ for all $s \in S$, there is measurable function $f^\star_{t+1} : S \to A_{t+1}$ such that $f^\star_{t+1}(s) \in B_{t+1}(s)$ for all $s \in S$ and

$$\int_S \hat{x}_{t,t+1}(s')q(ds'; s, a) = \int_S [u_{t+1}(s, f^\star_{t+1}(s)) + \delta \int_S \hat{x}_{t+1,t+2}(s')q(ds'; s, f^\star_{t+1}(s))]q(ds; s_t, a_t). \tag{20}$$

Thus, there is a selection of nonrandomized actions in period $t+1$ that gives the same expected payoff to the players as the expected payoff from the payoff function (possibly randomized) of $\hat{x}_{t+1}$. Since this can be done in each period $t+1$, for $t \geq 1$, it follows that the sequence of functions $(f^\star_1, f^\star_2, \cdots, f^\star_{t+1}. \cdots)$ is a Markov equilibrium in pure strategies.

If there are a finite number of players and the game has a fixed cycle, then again the arguments for the existence of a pure strategy stationary equilibrium can be made first for the two-player as the extension to the general n-player case follows immediately. The argument uses theorem 4 to show that the randomized strategies can be purified.

For the two player case we observe that the payoff function $v^\star : S \to \mathbb{R}^2$ can be replaced by $\hat{v} : S \to \mathbb{R}^2$ so that

$$u_2(s, a) + \delta \int_S v^\star_2(s')q(ds' : s, a) = u_2(s, a) + \delta \int_S \hat{v}_2(s')q(ds' : s, a)$$

$$u_1(s,a) + \delta \int_S v_1^\star(s')q(ds' : s,a) = u_2(s,a) + \delta \int_S \hat{v}_1(s')q(ds' : s,a)$$

of players 2 and 1 where $\hat{v} \in \tilde{P}_1(u^\star)$. Also the payoff function $u^\star : S \to \mathbb{R}^2$ can be replaced by $\hat{u} : S \to \mathbb{R}^2$ so that

$$u_1(s,a) + \delta \int_S u_1^\star(s')q(ds' : s,a) = u_1(s,a) + \delta \int_S \hat{u}_1(s')q(ds' : s,a)$$
$$u_2(s,a) + \delta \int_S u_2^\star(s')q(ds' : s,a) = u_2(s,a) + \delta \int_S \hat{u}_2(s')q(ds' : s,a)$$

where $\hat{u} \in \tilde{P}_2(v^\star) = \tilde{P}_2(\hat{u})$. Then if $f_1^\star : S \to B_1(\hat{u})$ and $f_2^\star : S \to B_2(\hat{v})$ are measurable selections, then $(f_1^\star, f_2^\star)$ is a pure strategy stationary Markov-perfect equilibrium. A similar argument will work for the case of n players in a game with a fixed cycle. ∎

5.1 A Model of Resource Extraction with Noise

Here we provide an application of the result in Theorem 5 and show that a fairly general model of resource extraction has a stationary pure strategy Markov perfect equilibrium if the agents extract the resource in a given sequence.

Let the number of agents who extract the resource be some finite number n. Without loss of generality assume that the agents extract the resource in the sequence $1, 2, \cdots, n$ after which the agents extract the resource in the same sequence again. The state at the beginning of a period is given by the stock of the resource s which is an element of $[0, M]$. The amount of the resource that an agent can extract if s is the stock at the beginning of the period is $[0, s - \underline{c}]$. The idea here is that agents can extract the resource only until a small amount $\underline{c}$ of the resource is left because it is technologically not feasible to extract the resource once it falls below $\underline{c}$. The state space S can be decomposed into two components so that $S = X \times Z$, where $X = Y = [0, M]$ where M is the maximum amount of the resource available even if there is no extraction of the resource. The stock of the resource available for extraction at the beginning of a period is given by a non-atomic distribution $\mu(.|s,a)$. Let $\mu_x(.|s,a)$ denote the marginal of $\mu(.|s,a)$ on $x \in X$ and that this is non-atomic. Here $x \in X$ denotes the amount of the resource available for regeneration after a amount of it has been extracted for consumption from the stock s available at the beginning of the period.[10] The amount of

[10]The idea here is that once a amount of the stock has been extracted from a beginning of period stock of s, there is some random amount of the resource that is available for regeneration in the next period. This could be due to the fact that some amount of the resource is wasted during the extraction of the resource and is not consumed.

the resource available for extraction at the beginning of the next period is then drawn from the non-atomic distribution $\mu_Z(.|x)$ on Z which depends only on the realized value x of the resource available for regeneration and is independent of the state and action pair (s, a) of the previous period.[11]

Therefore, as in Duggan [7], the state space has two components X and Z, where $x \in X$ is the stock of the resource available for regeneration at the end of a period, and $z \in Z$ is the amount of the resource available at the beginning of the period. Each component of the state space $S = X \times Z$ is given by $[0, M]$ with the Borel σ-algebra and the Lebesgue measure. We note that here the state space is a compact metric space. We assume that the measure $\mu_x(.|s, a)$ has a density function $g(x|s, a)$ that is continuous in x and a and measurable in s. Furthermore, assume that conditional on $x \in X$, the distribution
$mu_Z(.|x)$ of $z \in Z$ is independent of the state and actions (s, a) of the preceding period and that the mapping $x \to \nu(.|x)$ is a regular conditional probability for z. The measure $\mu_Z(.|x)$ is absolutely continuous with respect to the Lebesgue measure on $[0, M]$ and is thus an atomless measure; and by the Radon-Nikodym theorem has a density $p(.|x)$ and this can be chosen so that $p(z|x)$ is jointly measurable in (z, x).

The model of resource extraction is then an Asynchronous Stochastic game with
(i) State space $S = X \times Z$ where $X = Z = [0, M]$ is a non-atomic measure space with the Borel σ-algebra and the non-atomic fixed measure λ.
(ii) A transition probability $\mu : S \times A \times S \to [0, 1]$ which satisfies the conditions described above
(iii) The action set $A = [\underline{c}, M]$ which describes the amount of the resource that an agent can feasibly extract.
(iv) The single period utility function $u_i : S \times A \to \mathbb{R}$ which describes the single-period utility of agent i from extracting and consuming the resource. It is measurable in the state variable s and continuous on the action set A.

Proposition 3 *The Asynchronous Resource Extraction Model with Noise has a Markov-perfect equilibrium in pure strategies.*

Proof: We first observe the state space S is a compact metric space and the measure

[11] This describes the fact that the stock available for extraction at the beginning of a period is determined only by the stock available for regeneration and is independent of (s, a), and that there is some “noise” involved in the process of regeneration.

$\mu_X(.|s,a)$ has a density function $g(x|s,a)$ that is continuous in x and a and measurable in s. Further, the measure $\mu_Z(.|x)$ is independent of (s,a). Therefore the transition probability $\mu : S \times A \times S \to [0,1]$ is norm continuous in (s,a) and absolutely continuous with respect to the fixed nonatomic measure λ. Further, we note that the conditions on the transition probability μ is the same as that of conditions (iv) and (v) in Duggan [7]. But this implies that it satisfies the conditions for a stochastic game with endogenous shocks in He and Sun [11]. Therefore, by corollary 1 in He and Sun [11], it has a decomposable coarser transition kernel. Therefore, the resource extraction model satisfies the conditions of Theorem 5 and has a pure strategy Markov-perfect equilibrium. ■

6 Conclusion

Here we have shown that Markov-perfect equilibrium exists for Asynchronous Stochastic games in which the single-period payoff is a measurable function of the state variable. Theorem 1 shows that the a finite horizon game has a Markov-perfect equilibrium in pure strategies. Theorem 2 shows a Markov-perfect equilibrium exists for infinite horizon games but in randomized strategies. Given these results, it would be natural to ask whether the result of Theorem 2 can be sharpened to one that showed the existence of equilibrium in pure strategies. In an infinite horizon game the optimal choice of a player who chooses in time period t, can change from one element to another in the action set, as the terminal period from which the backward induction process starts, changes from T to $T+1$, and then to $T+2$ and so on. Thus as the time period in the future from which the players start to optimize, is pushed further into the future, the optimal choice of the player who chooses in time period t, changes from one element to another in the action set. These optimal choices may never converge to any element of the action set, or if the sequence of optimal choices converge, the sequence converges to a convex combination of the elements of the action set. The result of Theorem 2 should be seen in this context, as should the result in Theorem 5. Viewed within this context the result in Theorem 5 is of particular interest. One interpretation of the result of Theorem 5 shows that when there is noise in the realization of the state variable (that is the measure μ is non-atomic, and the transition probability does not give an exact indication of this noise (the condition of decomposable coarser transition kernel), then there is a Markov-perfect equilibrium in pure strategies.

One may also ask how important is the condition of norm continuity for the transition probability and whether one may be able to relax this condition. The answer, as in the case of general stochastic games, is perhaps not much can be done to relax this condition. If the game is a finite horizon game, a condition like weak* continuity may work, which is the weakest topology that makes the future expected payoff continuous in actions. In the infinite horizon game, however, something like the norm continuity condition would be needed. This is most clearly seen in the result in lemma 3, where the condition of norm continuity is needed for the value function to be jointly continuous in actions and future expected payoff functions.

A final point perhaps should be addressed about Asynchronous stochastic games. These games have a structure that, in many respects is similar to that of stochastic dynamic programming, so one is quite naturally led to wonder if one may be able to get existence results similar to those for stochastic dynamic programming. There is, however, a big and very important difference between stochastic dynamic programming and Asynchronous stochastic games, namely, that the single-period payoff functions in the game differ from period to period, as these are the single-period payoff functions of different players. Thus there is nothing like the optimal value function of stochastic dynamic programming for asynchronous stochastic games. But this does not rule out interesting possibilities when the asynchronous stochastic games are of a special kind, like for example zero sum games.

While the main results use the framework of stochastic games it should be evident that the results apply to a wide variety of asynchronous stochastic models. Some of this is seen in the applications that have been discussed like the dynamic oligopoly models and the model of resource extraction, but there are obviously many others. Principal-agent models in which the principal makes an offer that is followed by an action of the agent, which then leads to the next period where the principal again makes an offer to which the agent responds also would be a model in which actions are taken in a sequence. It is possible that some kinds of networks may also have a structure similar to the models discussed here.

References

[1] C. D. Aliprantis and K. Border, Infinite Dimensional Analysis: A Hitchiker's Guide, Third Edition, Springer-Verlag, Berlin Heidelberg, 2005.

[2] P. Barelli and J. Duggan, A Note on Semi-Marprokov Perfect Equilibrium in Discounted Stochastic Games, *Journal of Economic Theory*, , 151, 596-604, 2014.

[3] C. Castaing and M. Valadier, Convex Analysis and Measzurable Multifunctions, Springer-Verlag, New York 1977.

[4] K. L. Chung, A Course in Probability Theory, Academic Press, New York, 1974.

[5] Dynkin, E. B. and I. V. Evstigneev, Regular conditional expectations of correspondences, *Theory of Probability and its Applications*, 21, 325 - 338, 1977.

[6] D. Duffie, J. Geanakoplos, A. Mas-Collel, A. McLennan, Stationary Markov equilibria, *Econometrica*, 62, 745 - 781, 1994.

[7] J. Duggan, Noisy Stochastic Games, *Econometrica*, 80, 2017 - 2045, 2012.

[8] Fink, A., Equilibrium in Stochastic n-person Games, *Journal of Science of the Hiroshima University*, 28, 89- 93, 1964

[9] Granas, A. and J. Dugundji, Fixed Point Theory, Springer-Verlag, New York, 2003.

[10] Halmos, P., Measure Theory, Springer-Verlag, New York, 1974.

[11] He, Wei and Y. Sun, Stationary Markov perfect equilibria in disciunted stochastic games, *Journal of Economic Theory*, 169, 15 - 61, 2017.

[12] E. Klein and A.C. Thompson, Theory of correspondences, John Wiley & Sons, New Yor, 1984.

[13] Y. Levy, Discounted Stochastic games with no stationary Nash equilibrium: two examples, *Econometrica*, 81, 1973 - 2007, 2013.

[14] Y. Levy and A. Mclennan, Corrigendum to "Discounted Stochastic Games with no Stationary Nash Equilibrium: Two Examples," *Econometrica*, 83, 1237 - 1252, 2015.

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

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

[17] Parthasarathy,K. R., Probability Measures on Metric Spaces, Academic Press, New York, 1967.

[18] Pakes, A. and McGuire, P., Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model, *Rand Journal of Economics*, 25(4), 555-589, 1994.

[19] Pakes A. and P. McGuire, Stochastic Algorithms, Symmetric Markov-Perfect Equilibrium and the "Curse" of dimensionality, *Econometrica*, 69, 1261-1281, 1994.

[20] S. K. Chakrabarti, Markov Equilibria in Discounted Stochastic Games, *Journal of Economic Theory*, 85, 294 - 327, 1999.

[21] S. K. Chakrabarti, Markov Equilibrium in Dynamic Economic Models. Mimeo 2017, Forthcoming in *Essays in Honor of Bernard Cornet*, Editors N. Yannelis, et.al.

[22] Shapley, L., Stochastic Games, *Proceedings of the National Academy of Sciences*, 39, 1095 - 1000, 1953.