Discounted Supermodular Stochastic Games: Theory and Applications[1]

Rabah Amir[2]

Received: ______________; *Accepted:* ______________

Abstract

This paper considers a general class of discounted Markov stochastic games characterized by multidimensional state and action spaces with an order structure, and one-period rewards and state transitions satisfying some complementarity and monotonicity conditions. Existence of pure-strategy Markov (Markov-stationary) equilibria for the finite (infinite) horizon game, with nondecreasing –and possibly discontinuous – strategies and value functions, is proved. The analysis is based on lattice programming, and not on concavity assumptions. Selected economic applications that fit the underlying framework are described: dynamic search with learning, long-run competition with learning-by-doing, and resource extraction.

JEL codes: C73, C61, D90.

Key words and phrases: Stochastic game, supermodularity, Markov strategy dynamic programming.

[1] I would like to thank Jean-Francois Mertens, Abraham Neyman, Matt Sobel, Zari Rachev and Sylvain Sorin for helpful exchanges concerning the subject of this paper.
[2] Department of Economics, University of Arizona, Tucson, AZ 85721 (E-mail: ramir@eller.arizona.edu)

1. Introduction

Stochastic games provide a natural framework for modelling competition over time in situations where agents' actions influence the economic environment in a way that can be captured by a state variable. Viewed as game-theoretic analogs of dynamic optmization problems, stochastic games fit as a tool of analysis in a variety of areas in economics, including in particular resource extraction and industrial organization. Introduced in a classic paper by Shapley (1953), stochastic games have been an active field of research in pure game theory[3] , in systems theory[4] and in economics[5] . Interestingly, while a fair amount of interaction on dynamic games has taken place between the latter two fields, the pure theory developed quite independently and, as a general rule, did not provide directly usable results in applications.

In the pure theory, there is an extensive literature dealing with the existence of subgame-perfect equilibrium in stochastic games, culminating with the work of Mertens and Parthasarathy (1987) who establish the aforementioned existence in strategies that are (partly) history-dependent when the transition law is continuous in the variation norm in the actions, a very strong assumption that rules out many economic applications of interest. Shifting focus away from Nash equilibrium, Nowak and Raghavan (1992) and Harris, Reny and Robson (1995) show existence of a type of correlated equilibrium using the strong continuity assumption described above (see also Duffie et.al., 1988). Recently, Nowak (2002) established the existence of Markov-stationary equilibrium for a class of games characterized by a transition law formed as the linear combination of finitely many measures on the state space.

[3] See Neyman and Sorin (2003) for a thorough series of papers covering the state of the art on the theory of stochastic games.

It is also worthwhile to point out that many of the basic results behind the theory of dynamic programming (such as the contraction property in value function space) were already unequivocally laid out in Shapley's (1953) seminal paper over a decade before being rediscovered again (Blackwell, 1965 and Denardo, 1967).

[4] A standard reference for this part of the literature referring to dynamic games is Basar and Olsder (1999).

[5] See Amir (2000) for a fairly thorough survey of the applications of stochastic games to economics and management science.

As for economic applications, they can be essentially classified in three different categories, as laid out in some detail in Amir (2003). The first consists of numerous studies relying on the well-known linear-quadratic[6] model (with deterministic or stochastic transitions) in various settings. The reason for selecting this choice is clearly tractability: In any finite-horizon, there is a unique Markov equilibrium with closed-form strategies that are linear in the state variable.[7]

The second restricts the players's strategy space to open-loop strategies. While the resulting game is then substantially easier to analyse in most cases, this restriction on the players' behavior has become less accepted in economics in recent years, in terms of approximating real-life behavior in most settings. Open-loop strategies simply entail an excessive level of commitment on the part of the players.

The third category considers Markov behavior and general specification of the primitives of the model. As in the other two categories, behavior is still limited to pure strategies, as is often the case in economic modelling. These papers have generally exploited the special structure dictated by the economic environment to prove existence of a Markov equilibrium and provide a characterization of its properties.

The present paper contributes both to the general theory and to the third category above. We consider a Markov-stationary discounted stochastic game with multidimensional state and action spaces, and impose minimal monotonicity and complementarity (i.e. supermodularity-type) assumptions on the reward and state transition functions that guarantee the existence of a Markov-stationary equilibrium. The associated strategies and value functions are all monotone nondecreasing in the state variable, as a consequence of the assumed monotonicity and complementarity structure. The resulting structured class of dy-

[6] That is, the state transition law is linearly additive in the state, actions and noise variable (if any), and the one-period reward is quadratic in the state and actions. The general properties of (multidimensional versions) of this model are analysed in detail in Basar and Olsder (1999).

[7] In the framework of resource extraction, Levhari and Mirman's (1980) well-known model has a solution sharing these same tractability features.

namic games may then be appropriately termed discounted supermodular stochastic games.

To relate this paper to the general theory, observe that the main result here is the most general existence result of Nash equilibrium in Markov-stationary strategies in the literature on discounted stochastic games with uncountable state and action spaces. Exploiting the rich structure of our setting, the existence result at hand requires continuity of the (distribution function) of the transition law only in the topology of uniform convergence in the actions in the infinite-horizon case, and of weak continuity of the same in the finite-horizon case.

This paper also closely relates to economic applications in that the structure at hand is general enough to encompass many of the stochastic game models in economics. To illustrate this point, a few specific applications of the set-up are presented at the end of the paper, some in full detail and others listed as possible extensions. While the reader may at first think that the result at hand relies on too many assumptions, these applications illustrate convincingly that the underlying assumptions are quite natural in a variety of settings, where clear economic interpretations can be appropriately provided. In this sense, this paper may be viewed as a first step towards a theory of structured stochastic games oriented towards economic applications. Finally, we stress that the equilibria at hand are always in pure strategies, which satisfies an important restriction imposed by economists' persistent reluctance to deal with mixed strategies.

Of all our assumptions, the most restrictive are the complementarity assumptions on the transition law, which, as we argue later, exclude deterministic transitions from being a special case of our set-up. Thus, it seems that circumventing the use of mixed strategies at this level of generality has a price. Indeed, one may think of these complementarity assumptions on the transitions as reflecting an assumption of sufficient exogenous noise in the system to replace the endogenous noise usually engendered by mixed strategies. To add some perspective, it is certainly worthwhile to point out that virtually all the studies of strategic dynamics conducted at a high level of generality required some assumption(s) of the same type as ours here on the transition law that rule out deterministic transitions. In particular, Amir (1996a-b) and Ericson and Pakes (1995) assume a strong notion of convexity on the transitions that is quite closely related to our assumptions here, as brought out precisely in our end applications here.

From a methodological perspective, it is hoped this paper will convey a sense that the lattice-theoretic approach is well-suited for analysing dynamic games in economics, as it provides a natural framework for turning a lot of natural economic structure into appealing monotonic relationships that survive the dynamic programming recursion while satisfying the pure-strategy restriction.

2. Existence of Pure-Strategy Markov Equilibrium

This section provides the formal description of our stochastic game, the assumptions needed for the underlying analysis, the main results and a discussion of the scope of the assumptions and of the results.

2.1 Problem Statement

Consider an n-player discounted stochastic game described by the tuple $\left\{S, A_i, \widetilde{A}_i, \lambda_i, r_i, p\right\}$ with the following standard meaning. The state space S and actions spaces A_i are all *Euclidean intervals*, with $S \subset R^k$ and $A_i \subset R^{k_i}$. Denote the joint action set by $A = A_1 \times ... \times A_n$ and a typical element $a = (a_1, ..., a_n) = (a_i, a_{-i})$, for any i. $\widetilde{A}_i$ is the feasibility correspondence, mapping S to the subsets of A_i, so that $\widetilde{A}_i(s)$ is player i's set of feasible actions when the state is s. The one-period reward function for player i is $r_i : S \times A \to R$, and his discount factor is $\lambda_i \in [0, 1)$. Finally, p denotes the transition probability from $S \times A$ to (the set of probability measures on) S.

Throughout this paper, it is convenient to place assumptions on, and work with, the cumulative distribution function F associated with the transition probability p, defined by

$$F(s^{'}/s,a) \triangleq \Pr ob(s_{t+1} \leq s^{'}/s_t = s, a_t = a) \text{ for any } s, s^{'} \in S \text{ and } a \in A. \quad (1)$$

The standard definitions of pure Markov and Markov-stationary strategies, and expected discounted payoffs are now given. A general pure-strategy for Player i is a sequence $\Gamma_i = (\gamma_1, \gamma_2, ..., \gamma_t, ...)$ where γ_t specifies a (pure) action vector to be taken at stage t as a (Borel-measurable) function of the history of all states and actions up to stage t. If this history up to stage t is limited to the value of the current state, s_t, then the strategy is said to be Markov. If a Markov strategy $(\gamma_1, \gamma_2, ..., \gamma_t, ...)$ is time-invariant, i.e. such that $\gamma_j = \gamma_k \triangleq \gamma$ for all $j \neq k$, then the strategy is said to be Markov-stationary, and can then be designated by γ.

Given an n-tuple of general strategies $\Gamma = (\Gamma_1, \Gamma_2, ..., \Gamma_N)$ for the N players, and an initial state $s \in S$, there exists a unique probability distribution $m(\Gamma, s)$ that is induced on the space of all histories, according to Ionescu-Tulcea's Theorem (see e.g. Bertsekas and Shreve, 1978). Given a horizon of T periods (which may be finite or infinite, as will be specified), it can be shown via standard arguments that the expected discounted payoff of player i can be written as

$$U_i(\Gamma, s) = (1-\lambda) \sum_{t=0}^{T} \lambda^t R_i^t(\Gamma)(s), \quad (2)$$

where $R_i^t(\Gamma)$ is the stage-t expected reward for player i, or in other words, with m_i^t denoting the stage-t marginal of $m(\Gamma, s)$ on A_i,

$$R_i^t(\Gamma)(s) = \int r_i(s_t, a_t) dm_i^t(\Gamma, s).$$

An n-tuple of strategies $\Gamma^* = (\Gamma_1^*, \Gamma_2^*, ..., \Gamma_N^*)$ constitutes a Nash equilibrium if no player can strictly benefit from a unilateral deviation, or, for all $i = 1, 2, ..., N$,

$$U_i(\Gamma^*, s) \geq U_i(\Gamma_i, \Gamma_{-i}^*, s) \text{ for any strategy } \Gamma_i. \quad (3)$$

A Nash equilibrium is Markov (Markov-stationary) if the associated equilibrium strategies are Markov (Markov-stationary). In checking for Markov equilibrium, whether we restrict the space of allowable deviations to Markov strategies only or allow general (history-dependent) strategies is immaterial, in the sense that whatever payoff a *unilateral* deviation by one player can achieve using a general strategy can also be achieved relying only on Markov deviations. In other words, a Markov equilibrium obtained when considering only Markov deviations remains an equilibrium when more general strategies are allowed. Since Markov strategies are much easier to handle, this invariance property is very convenient.

When the horizon is infinite ($T = \infty$), we can define a Markov-stationary equilibrium in an analogous way, using Markov-stationary strategies as deviations. Such an equilibrium remains an equilibrium if the players are allowed to use arbitrarily more general strategies according to an analogous mechanism as for Markov strategies.[8] An important consequence of these facts is that the most general existence result for a Markov (infinite-horizon Markov-stationary) discounted stochastic game, i.e. one with reward and transition law that are Markov (Markov and time -invariant), is in Markov (Markov-stationary) strategies.

2.2 The Assumptions and their Scope

The following *Standard Assumptions* on the state and action spaces (S), rewards (R), transitions (T) and feasible action correspondence (A), for each $i = 1, ..., n$, are in effect *throughout this paper*, without further reference. Let R^k_+ denote the positive orthant of k-dimensional Euclidean space. Upper semi-continuity will always be abbreviated by u.s.c. for functions and u.h.c. for correspondences. A brief summary of all the lattice-theoretic notions and results invoked here is provided in the Appendix.

·On the basic spaces and the feasibility correspondence of the game:

[8] This follows from general results in dynamic programming theory: Given all the other players use Markov (Markov-stationary) strategies, the player under consideration faces a Markov (Markov-stationary) dynamic program, for which it is known that an optimal strategy (achieving the global maximum of the overall payoff) exists within the space of Markov (Markov-stationary) strategies: See e.g. Bertsekas and Shreve (1978).

(A1) The state space S is an interval in R^k_+.

(A2) The actions spaces A_i are all *compact Euclidean intervals*, with $A_i \subset R^{k_i}_+$.

(A3) $\widetilde{A}_i(s)$ is a compact sublattice of A_i for each $s \in S$.

(A4) $\widetilde{A}_i$ is ascending and upper hemi-continuous in s.

(A5) $\widetilde{A}_i$ is expanding, i.e. $\widetilde{A}_i(s_2) \subset \widetilde{A}_i(s_1)$ whenever $s_1 \geq s_2$.

·On the reward function:

(R1) r_i is jointly continuous in (a_i, a_{-i}) for fixed s and u.s.c in s for fixed (a_i, a_{-i}).

(R2) r_i is increasing in (s, a_{-i}), for each a_i.

(R3) r_i is supermodular in a_i and has strictly nondecreasing differences in $(a_i; a_{-i}, s)$.

(R4) r_i is uniformly bounded, i.e. $\exists K > 0$ such that $|r_i(s, a)| \leq K$, for all $(s, a) \in S \times A$.

·On the transition law:

(T1) p is *weak**-continuous in (s, a_i, a_{-i}) for each $s' \in S$, i.e. for every Borel set $E \subset S$,

$$p(E/s^k, a^k) \to p(E/s, a) \text{ whenever } (s^k, a^k) \to (s, a) \text{ and } p(\partial E/s, a) = 0,$$

where ∂E is the boundary of E.[9]

(T2) F is increasing in (s, a) in the sense of first-order stochastic dominance.

(T3) F is supermodular in a and has increasing differences in (a, s).

We next discuss the scope and limitations of this set of assumptions. For the sake of brevity, we will skip assumptions that are either self-evident in content, or made for purely technical reasons, in a standard sense. (A4) and (A5) essentially say that as the state variable increases, new higher actions become feasible while no actions are lost on the low end of the feasible set. For any given player, a higher value of the state variable and/or of

[9] Equivalently, Assumption (T1) may be restated as: $F(s^{'}/s^k, a^k) \longrightarrow F(s^{'}/s, a)$ as $(s^k, a^k) \to (s, a)$ at every point s' of continuity of the limit function $F(s'/s, a)$, where $F(\cdot/s, a)$ is the c.d.f. associated with $p(\cdot/s, a)$. An alternative characterization of Assumption (T1), most useful in proofs, is: For every continuous bounded function v,

$$\int v(s')dF(s'/(s^k, a^k)) \longrightarrow \int v(s')dF(s'/(s, a)) \text{ as } (s^k, a^k) \longrightarrow (s, a).$$

the rivals' actions increases the reward today (R2), the probability of a higher state in the next period (T2), the marginal returns to an increase in the player's actions (R3), and the marginal increase (with respect to an increase in own actions) in the probability of a higher value of the state in the next period (T3). Similarly, a higher value of a subset of a player's actions increases the marginal reward (R3) and the marginal (probabilistic) increase in the next state (T3), due to higher values of the remaining actions.

The assumption of continuity specifying the dependence of the state transition probabilities on the state-action pair is always a central assumption in the theory of stochastic games. Here, (T1), often referred to as weak convergence, is essentially the most general assumption possible in such a context. In particular, it is compatible with having deterministic transitions that are continuous functions of the state and action variables (see (4) below from the expression for F then). In the present paper, (T1) is sufficient to analyze the finite-horizon game, but not the infinite-horizon game, which will require a stronger notion of continuity (discussed below).

Of all the above assumptions, the most restrictive is arguably the supermodularity assumptions on the transitions, (T3). Indeed, it rules out (nondegenerate) deterministic transitions, as argued below. Before doing so, it is insightful to consider (T3) for the special case of real state and action spaces. It is then easily shown (Topkis, 1968) that (T3) is equivalent to[10] ,[11]

$$F(s'/s, a_i, a_{-i}) \text{ being submodular in } (s, a), \text{ for each } s' \in S.$$

A transition probability of the Dirac type cannot satisfy either component of Assumption (T3). To see this, consider a deterministic transition law given by $s^{t+1} = f(s^t, a_i^t, a_{-i}^t)$, where f is a continuous function. The distribution function of the corresponding transition

[10]The requirement of supermodularity of the transitions with respect to a_i is trivially satisfied when a_i is a scalar.

[11]To avoid confusion, note that the supermodularity of F as defined in Appendix is equivalent to the submodularity of the function $F(s'/s, a_i, a_{-i})$ in the indicated arguments. This is only valid in the scalar case.

probability can be written as

$$F(s'/s, a_i, a_{-i}) = \begin{cases} 0 \text{ if } s' < f(s, a_i, a_{-i}) \\ 1 \text{ if } s' \geq f(s, a_i, a_{-i}) \end{cases} \tag{4}$$

Assume, for simplicity (and for the sake of the present argument only), that there are only two players (i and $-i$) and that the state and the action spaces are all given by $[0, 1]$. Then it is easy to verify that, as defined by (4), $F(s'/s, a_i, a_{-i})$ cannot be submodular (say) in (a_i, a_{-i}) for fixed s, unless f is actually independent of one of the a's. To see this, simply graph F on the (a_i, a_{-i})-unit square, and observe that unless the zero-one discontinuity of F happens along a vertical or a horizontal line, F will not be submodular in $(a_i, a_{-i}) \in [0, 1]^2$ for fixed s.[12] A similar argument holds for the other pairs of arguments.

Nonetheless, the exclusion of deterministic transitions notwithstanding, Assumptions (T1)-(T3) are general enough to allow for a wide variety of possible transition probabilities, including rich families that can be generated by mixing autonomous distribution functions, ordered by stochastic dominance, according to mixing functions satisfying the complementarity and monotonicity conditions contained in (T2)-(T3). Specifically, let F_1 and F_2 be distribution functions such that $F_1 \succ F_2$, where $\succ$ stands for first-order stochastic dominance, and let $g : S \times A_i \times A_{-i} \longrightarrow [0, 1]$ be nondecreasing in s, supermodular in a and have increasing differences in (s, a). Then the transition probability given by

$$F(s'/s, a_i, a_{-i}) = g(s, a_i, a_{-i})F_1(s') + [1 - g(s, a_i, a_{-i})]F_2(s')$$

is easily seen to satisfy Assumption (T2)-(T3). Indeed, for any nondecreasing function $v : S \to R$,

$$\int v(s')dF(s'/s, a_i, a_{-i}) = g(s, a_i, a_{-i}) \int v(s')dF_1(s') + [1 - g(s, a_i, a_{-i})] \int v(s')dF_2(s')$$

so that, invoking Theorem ?, the verification follows immediately from Theorem ? for the smooth case, and upon standard inequality manipulations without smoothness assumptions.

[12]To perform the verification, simply check the usual inequality on the four vertices of a rectangle in R_+^2, i.e. with $a_i' \geq a_i$ and $a_{-i}' \geq a_{-i}$: $F(a_i', a_{-i}') - F(a_i, a_{-i}') \leq F(a_i', a_{-i}) - F\ (a_i, a_{-i})$.

2.3 The Main Results and their Scope

The main results of this paper are:

Theorem 1 *Under the Standard Assumptions, for every finite horizon, the discounted stochastic game has a Markov equilibrium, with strategy components and corresponding value functions that are upper semi-continuous and nondecreasing in the state vector.*

The infinite-horizon game requires a stronger notion of continuity in the actions (but not in the state) than **(T1)** for the transition probability p, that is best expressed on the associated distribution function F :

(T1)* $F(\cdot/s, a)$ is $weak^*$-continuous in s for each $a \in A$, and continuous in a in the topology of uniform convergence for each $s \in S$, i.e.

$$\sup_{s' \in C} \left| F(s'/s, a^k) - F(s'/s, a) \right| \to 0 \text{ as } a^k \to a \text{ for any } s \in S \text{ and compact subset } C \subset S.$$

Theorem 2 *Under the Standard Assumptions and **(T1)***, the infinite-horizon discounted stochastic game has a Markov-stationary equilibrium, with strategies and corresponding value functions that are upper semi-continuous and nondecreasing in the state vector.*

3. Proofs

This section provides the proofs of our two main results, breaking up the underlying arguments into a sequence of lemmas. Additional notation is introduced as the need for it arises. It is actually more convenient to start with the steps leading to the proof of Theorem 2, and then move on to those of Theorem 1. We begin with setting the various spaces of interest.

For any compact Euclidean set E, let $BV(S, E)$ be the Banach space of right-continuous functions of bounded variation from S to E endowed with the variation norm[13] . Denote by $M(S, E)$ the subset of $BV(S, E)$ consisting of nondecreasing (right-continuous) functions, and by $M_K(S, R)$ the subset of functions in $M(S, R)$ taking values in $[-K, K]$, where R stands for the reals and K is the upper bound on the one-period rewards (Assumption (A.4).)

[13]This banach space is isomorphic to the space of signed bounded regular measures with the variation norm (see e.g. Luenberger, 1968).

The feasible strategy space for player i in the infinite-horizon game is the following subset of the set of all stationary strategies:

$$\widetilde{M_i}(S, A_i) \triangleq \left\{ \gamma \in M(S, A_i) \text{ such that } \gamma(s) \in \widetilde{A}_i(s) \right\}.$$

For a finite T-period horizon , a Markovian strategy for player i consists of a sequence of length T of elements of $\widetilde{M_i}(S, A_i)$. Let $\widetilde{M}(S, A) \triangleq \widetilde{M_1}(S, A) \times ... \times \widetilde{M_n}(S, A) = \widetilde{M_i}(S, A) \times \widetilde{M}_{-i}(S, A)$.

By Assumption (R.4) and the discounted nature of the payoffs, all feasible payoffs in this game are $\leq K$. Hence, the space of all possible value functions in this game is a subset of $M_K(S, R)$.

It is well-known that $BV(S, E)$ is the dual of the Banach space $C(S, E)$ of bounded continuous functions with the sup norm. Throughout the proof, we will endow the effective strategy and value function spaces, $\widetilde{M}(S, A)$ and $M_K(S, R)$ respectively, with the weak* topology of the corresponding BV space. The well-known characterizations of convergence in this topology are given in Footnote 7.

Player i's best-response problem for the infinite-horizon stochastic game may be defined as follows, given the rivals' stationary strategies γ_{-i} (note that we write $V_\gamma(s)$ instead of $V_{\gamma_{-i}}(s)$ for notational simplicity):

$$V_\gamma(s) \triangleq \sup E \left\{ (1 - \lambda_i) \sum_{t=0}^{\infty} \lambda_i^t r(s^t, a_i^t, \gamma_{-i}(s^t)) \right\} \tag{5}$$

subject to

$$s^{t+1} \sim p(\cdot / s^t, a_i^t, \gamma_{-i}(s^t)) \text{ with } s^0 = s, \tag{6}$$

where the expectation $E\{\cdot\}$ is over the unique probability measure on the space of all histories that is induced by s, $\gamma_{-i}(\cdot)$ and a stationary strategy by player i. Furthermore, the supremum may be taken over the space of stationary strategies without any loss of value since, as discussed earlier, given the other players' strategies $\gamma_{-i}(\cdot)$, (5)-(6) is a Markov-stationary dynamic programming problem.

We begin with some preliminary lemmas of a technical nature (recall that the Standard Assumptions are in effect throughout the paper, without further mention.)

Lemma 3 *Let $v \in M_K(S,R)$. Then $\int v(s')dF(s'/s,a_i,a_{-i})$ is jointly u.s.c. in (s,a_i,a_{-i}).*

Proof. Recall that a function is u.s.c. if and only if it is the pointwise limit of a decreasing sequence of continuous functions (see e.g. Goffman, 1953). From this fact, we know here that since v is u.s.c., there exists a sequence of continuous functions $v^m \downarrow v$. For each m, $\int v^m(s')dF(s'/s,a_i,a_{-i})$ is continuous in (s,a_i,a_{-i}) due to the continuity of v^m, Assumption (T.1) and the well-known characterization of *weak** convergence via integrals. Furthermore, by the Monotone Convergence Theorem, since $v^m \downarrow v$, we have $\int v^m(s')dF(s'/s,a_i,a_{-i}) \downarrow \int v(s')dF(s'/s,a_i,a_{-i})$. Hence, being the limit of a decreasing sequence of continuous functions, $\int v(s')dF(s'/s,a_i,a_{-i})$ is u.s.c. in (a_i,a_{-i}), using again the fact stated at the start of this proof. ■

Lemma 4 *Let γ_i be a sublattice-valued correspondence from S to A_i that is u.h.c. from above*[14] *and such that every one of its selection is nondecreasing. Then γ_i has a maximal selection, $\overline{\gamma}_i$, which is u.s.c. and continuous from above. Furthermore, $\overline{\gamma}_i$ is the only selection of γ_i satisfying these properties.*

Proof. The existence of the maximal selection $\overline{\gamma}_i$ follows from the fact that $\gamma_i(s)$ is a sublattice of A_i for each $s \in S$. $\overline{\gamma}_i$ is nondecreasing in s by assumption. To show $\overline{\gamma}_i$ is u.s.c., observe that for any $s_0 \in S$, $\limsup_{s \downarrow s_0} \overline{\gamma}_i(s) \leq \max \gamma_i(s_0) = \overline{\gamma}_i(s_0)$, where the inequality follows from γ_i being u.h.c. from above. Since $\overline{\gamma}_i$ is nondecreasing, it follows that $\limsup_{s \to s_0} \overline{\gamma}_i(s) \leq \limsup_{s \downarrow s_0} \overline{\gamma}_i(s) \leq \overline{\gamma}_i(s_0)$. Hence, $\overline{\gamma}_i$ is u.s.c.

To show $\overline{\gamma}_i$ is continuous from above, consider for any $s_0 \in S$ and any sequence $s_n \downarrow s_0$: $\overline{\gamma}_i(s_n) \geq \overline{\gamma}_i(s_0)$ since $\overline{\gamma}_i$ is nondecreasing, so that $\liminf \overline{\gamma}_i(s_n) \geq \overline{\gamma}_i(s_0)$. Furthermore, lim sup $\overline{\gamma}_i(s_n) \leq \overline{\gamma}_i(s_0)$ since $\overline{\gamma}_i$ is u.s.c. Combining the two inequalities yields $\lim \overline{\gamma}_i(s_n) = \overline{\gamma}_i(s_0)$ whenever $s_n \downarrow s_0$, or that $\overline{\gamma}_i$ is continuous from above.

[14] *This is defined as: $s_k \downarrow s$ and $a_i^k \to a_i$ with $a_i^k \in \gamma_i(s_k) \Rightarrow a_i \in \gamma_i(s)$.*

We now show $\overline{\gamma}_i$ is the unique selection that is continuous from above. Since every selection of γ_i is nondecreasing, all selections of γ_i must coincide on a dense set of points where γ_i is single-valued and thus continuous as a function. It follows that any other selection that is continuous from above must coincide with $\overline{\gamma}_i$, since the values of $\overline{\gamma}_i$ on S are then completely determined by its values on a dense subset, as $\overline{\gamma}_i$ is continuous from above . ■

Lemma 5 *Let $\gamma_{-i} \in \widetilde{M}_{i-i}(S, A_i)$. Then V_γ, as defined by (7), is in $M_K(S, R)$ and is the unique solution to the functional equation*

$$V_\gamma(s) = \max_{a_i \in \widetilde{A}_i(s)} \left\{ (1-\lambda_i) r(s, a_i, \gamma_{-i}(s)) + \lambda_i \int V_\gamma(s') dF(s'/s, a_i, \gamma_{-i}(s)) \right\}. \tag{7}$$

Proof. For $\gamma_{-i} \in \widetilde{M}_{-i}(S, A)$, define an operator T on $M_K(S, R)$ by

$$Tv(s) \triangleq \sup_{a_i \in \widetilde{A}_i(s)} \left\{ (1-\lambda_i) r(s, a_i, \gamma_{-i}(s)) + \lambda_i \int v(s') dF(s'/s, a_i, \gamma_{-i}(s)) \right\}. \tag{8}$$

We show that T maps $M_K(S, R)$ into itself. To this end, we first show that Tv is nondecreasing in s. Let $v \in M_K(S, R)$ and $s_1 \geq s_2$. Then, by Assumption (R.2) and (T.2) and the fact that γ_{-i} is nondecreasing, we have

$$\begin{aligned} & (1-\lambda_i) r(s_1, a_i, \gamma_{-i}(s_1)) + \lambda_i \int v(s') dF(s'/s_1, a_i, \gamma_{-i}(s_1)) \\ \geq \; & (1-\lambda_i) r(s_2, a_i, \gamma_{-i}(s_2)) + \lambda_i \int v(s') dF(s'/s_2, a_i, \gamma_{-i}(s_2)). \end{aligned} \tag{9}$$

Since $A_i(s_2) \subset A_i(s_1)$ by Assumption (A.5), the conclusion that $Tv(s_1) \geq Tv(s_2)$ follows from taking sups on both sides of (9).

The next step is to show that Tv is u.s.c. To this end, we first show that the maximand in (8) is jointly u.s.c. in (a_i, s). In view of the fact that γ_{-i} is nondecreasing and right-continuous, an argument similar to the proof of Lemma 3 shows that $\int v(s') dF(s'/s, a_i, \gamma_{-i}(s))$ is u.s.c. in (s, a_i). The term $r(s, a_i, \gamma_{-i}(s))$ is also u.s.c. in (s, a_i), by Assumptions (R.1) and (R.2) and the fact that γ_{-i} is right-continuous and nondecreasing. Hence, being the sum of two u.s.c. functions, the maximand in (8) is also u.s.c. in (s, a_i), by the subadditivity of the $\liminf$ operator. Since $A_i(s)$ is u.h.c., Tv is u.s.c. by the Maximum Theorem. Since Tv

is nondecreasing in s, it is also right-continuous in s (such a step is shown in the proof of Lemma 4.)

It is clear that $Tv(\cdot) \leq K$. Hence we have established that T maps $M_K(S,R)$ into itself. $M_K(S,R)$ is a norm-closed subset of the Banach space of bounded Borel measurable functions with the sup norm. Hence, $M_K(S,R)$ is a complete metric space in the sup norm. A standard argument in discounted dynamic programming shows that T is a contraction, with a unique fixed-point, $V_\gamma(\cdot)$, which then clearly satisfies 7. ■

Lemma 6 *Let $\gamma_{-i} \in \widetilde{M}_{-i}(S, A_{-i})$. Then a maximal best-response $\overline{\gamma}_i$ exists, and is the only best-response in $\widetilde{M}_i(S, A_i)$.*

Proof. We first show that the maximand in (7) is supermodular in a_i and has nondecreasing differences in (a_i, s). The supermodularity in a_i follows directly from Assumptions (R.3) and (T.3) , Theorem 12 (in Appendix) and the fact that $V_\gamma(\cdot)$ is nondecreasing. To show strictly increasing differences for the r term in (7), let $a_i^{'} > a_i$ and $s' > s$ and consider,

$$\begin{aligned} r(s', a_i', \gamma_{-i}(s')) - r(s', a_i, \gamma_{-i}(s')) \quad &> r(s, a_i', \gamma_{-i}(s')) - r(s, a_i, \gamma_{-i}(s')) \\ &\geq r(s, a_i', \gamma_{-i}(s)) - r(s, a_i, \gamma_{-i}(s)) \end{aligned}$$

where the first inequality follows from the strictly increasing differences of $r(\cdot, \cdot, \gamma_{-i}(s'))$ in (s, a_i) from Assumption (R.3), and the second is from the nondecreasing differences of $r(s, \cdot, \cdot)$ in (a_i, a_{-i}) and the fact that $\gamma_{-i}(\cdot)$ is nondecreasing so that $\gamma_{-i}(s') \geq \gamma_{-i}(s)$. Increasing differences for the integral term in follows from analogous steps, upon invoking Theorem 12 (in Appendix). The details are omitted. Hence, the maximand in (7) has strictly increasing differences in (a_i, s), since this property is preserved by summation.

Since the maximand in (7) is also u.s.c. in a_i (from Lemma 5), and the feasible set $\widetilde{A}_i(s)$ is compact-valued and ascending , it follows from Topkis's Theorem 13 (in Appendix) that the maximal best-response $\overline{\gamma}_i$ of the best-response correspondence γ_i^* exists and (along with all the other best-response selections) is nondecreasing in s.

We now show that $\overline{\gamma}_i$ is u.s.c. and continuous from above in $s \in S$. From Lemma 3 and the proof of Lemma 4, we know that $\int V_\gamma(s')dF(s'/s, a_i, \gamma_{-i}(s))$ is u.s.c. and continuous from above in (s, a_i). Furthermore, by Assumptions (R.1)-(R.2) and the fact that γ_{-i} is nondecreasing and continuous from above, $r(s, a_i, \gamma_{-i}(s))$ is also continuous from above in (s, a_i). Hence, the maximand in (7) is also continuous from above in (s, a_i). Now, let $s^k \downarrow s$ and $a_i^k \downarrow a_i$ with $a_i^k \in \gamma_i^*(s^k)$. Towards showing $a_i \in \gamma_i^*(s)$ or that γ_i^* is u.h.c. from above, consider

$$\begin{aligned}
& (1-\lambda_i)r(s, a_i, \gamma_{-i}(s)) + \lambda_i \int V_\gamma(s')dF(s'/s, a_i, \gamma_{-i}(s)) \\
\geq \;& \limsup_{k\to\infty} \left\{ (1-\lambda_i)r(s^k, a_i^k, \gamma_{-i}(s^k)) + \lambda_i \int V_\gamma(s')dF(s'/s^k, a_i^k, \gamma_{-i}^k(s^k)) \right\} \\
= \;& (1-\lambda_i)r(s, a_i, \gamma_{-i}(s)) + \lambda_i \int V_\gamma(s')dF(s'/s, a_i, \gamma_{-i}(s)) \\
= \;& V_\gamma(s),
\end{aligned}$$

where the inequality is due to the fact that the bracketed term is u.s.c. in (s, a_i), and the first equality to the fact that it is continuous from above in (s, a_i). This shows that $a_i \in \gamma_i^*(s)$, so that $\gamma_i^*(s)$ is upper hemi-continuous from above at s. Since the maximal selection $\overline{\gamma}_i$ is nondecreasing, it is u.s.c. in s, and hence also continuous from above in s (lemma 4). That $\overline{\gamma}_i$ is the only best response that is continuous from above also follows from Lemma 4. Hence, $\overline{\gamma}_i$ is the unique best-response in $\widetilde{M_i}(S, A_i)$. ■

An important step in the proof of continuity of the best-response map is contained in the following intermediate result.

Lemma 7 *For any $v_i^k \to^* v$ in $M_K(S, R)$, $a_i^k \to a_i$ and $a_{-i}^k \to a_{-i}$, we have for each fixed $s \in S$,*

$$\int v_i^k(s')dF(s'/s, a_i^k, a_{-i}^k) \to \int v_i(s')dF(s'/s, a_i, a_{-i}) \tag{10}$$

provided $v_i^k(\overline{s}) \to v_i(\overline{s})$, where $\overline{s} \triangleq \sup S$.

Proof. By the integration by parts formula (see ?), we have

$$\int v_i^k(s')dF(s'/s,a_i^k,a_{-i}^k) = \left[v_i^k(s')F(s'/s,a_i^k,a_{-i}^k)\right]_{s'=0}^{s'=\infty} - \int F(s'/s,a_i^k,a_{-i}^k)dv_i^k(s') \quad (11)$$
$$= v_i^k(\overline{s}) - \int F(s'/s,a_i^k,a_{-i}^k)dv_i^k(s'),$$

since we always have $F(\overline{s}/s,a_i^k,a_{-i}^k)=1$ and $F(\inf S/s,a_i^k,a_{-i}^k)=0$. Likewise,

$$\int v_i(s')dF(s'/s,a_i,a_{-i}) = v_i(\overline{s}) - \int F(s'/s,a_i,a_{-i})dv_i(s'). \quad (12)$$

As $(a_i^k,a_{-i}^k)\to(a_i,a_{-i})$, we have, by Assumption (T1)*, for each fixed $s\in S$,

$$F(s'/s,a_i^k,a_{-i}^k)\to F(s'/s,a_i,a_{-i}), \textit{ uniformly in } s' \textit{ on compact subsets of } S. \quad (13)$$

The fact that $v_i^k(s')\to^* v_i(s')$ together with (13) implies that (see e.g. Billingsley, 1968, p.34), for each fixed $s\in S$,

$$\int F(s'/s,a_i^k,a_{-i}^k)dv_i^k(s') \to \int F(s'/s,a_i,a_{-i})dv_i(s'). \quad (14)$$

Since $v_i^k(\overline{s})\to v_i(\overline{s})$ by assumption, (10) follows from (11), (12)and (14). ■

We are now ready to define the single-valued best-response map B for our stochastic game, that associates to each n-vector γ of stationary strategies the *unique maximal* best response, $\overline{\gamma}$, in the sense of Lemma 6, i.e.

$$B:\widetilde{M}_1(S,A)\times...\times\widetilde{M}_n(S,A) \longrightarrow \widetilde{M}_1(S,A)\times...\times\widetilde{M}_n(S,A)$$
$$(\gamma_1,\gamma_2,...,\gamma_n) \longrightarrow (\overline{\gamma}_1,\overline{\gamma}_2,...,\overline{\gamma}_n).$$

Lemma 8 *B is continuous in the product weak*-topology.*

Proof. It suffices to show continuity along one coordinate, i.e. of the map $\gamma_{-i}\longrightarrow\overline{\gamma}_i$. Let $\gamma_{-i}^k\longrightarrow\gamma_{-i}$ and assume (by going to a subsequence if needed, which is possible since $\widetilde{M}_i(S,A_i)$ is *weak**-compact by the Alaoglu-Bourbaki Theorem) that $\overline{\gamma}_i^k\longrightarrow\overline{\gamma}_i$. We must show that $\overline{\gamma}_i$ is the maximal best-response to γ_{-i}. Denoting $V_{\gamma_{-i}^k}$ by V_i^k, we have

$$\begin{aligned} V_i^k(s) &= (1-\lambda_i)r(s,\overline{\gamma}_i^k(s),\gamma_{-i}^k(s)) + \lambda_i\int V_i^k(s')dF(s'/s,\overline{\gamma}_i^k(s),\gamma_{-i}^k(s)) \\ &= \max_{a_i\in\widetilde{A}_i(s)}\left\{(1-\lambda_i)r(s,a_i,\gamma_{-i}^k(s)) + \lambda_i\int V_k(s')dF(s'/s,a_i,\gamma_{-i}^k(s))\right\}. \end{aligned} \quad (15)$$

By Helly's Selection Theorem (or the Alaoglu-Bourbaki Theorem) , the sequences of functions $\gamma_{-i}^k, \overline{\gamma}_i^k$ and V_i^k all have *weak** convergent subsequences, each with a nondecreasing right-continuous limit. By iterating if necessary, take a common convergent subsequence for all three sequences, that has the further property that $V_i^k(\overline{s}) \to^* V_i(\overline{s})$, where $\overline{s} = \sup S$ (see Lemma 7). W.l.o.g., relabel this subsequence with the index k (for simpler notation.) Thus, we have $\gamma_{-i}^k \to^* \gamma_{-i}, \overline{\gamma}_i^k \to^* \overline{\gamma}_i$, $V_i^k \to^* V_i$ and $V_i^k(\overline{s}) \to^* V_i(\overline{s})$.

The rest of the proof will consist of taking *weak** limits, term-by-term, on both sides of (15) along the subsequence just identified. Since *weak** convergence is equivalent to pointwise convergence on the subset of points of continuity of the limit function (Billingsley, 1968), and since the latter is dense for a nondecreasing function, there is a dense subset, call it S_C, of S such that $\gamma_{-i}, \overline{\gamma}_i$ and V_i are all continuous on S_C.

For any fixed $s \in S_C$, we have $\overline{\gamma}_i^k(s) \to \overline{\gamma}_i(s)$ and $\gamma_{-i}^k(s) \to \gamma_{-i}(s)$. Also, $V_i^k \to^* V_i$. Hence, by Lemma 7, $\int V_i^k(s')dF(s'/s, \overline{\gamma}_i^k(s), \gamma_{-i}^k(s)) \to \int V_i(s')dF(s'/s, \overline{\gamma}_i(s), \gamma_{-i}(s))$. Likewise, by Assumption (R.1), $r(s, \overline{\gamma}_i^k(s), \gamma_{-i}^k(s)) \to r(s, \overline{\gamma}_i(s), \gamma_{-i}(s))$. Since V_i is continuous at s and $V_i^k \to^* V_i$, we must have $V_i^k(s) \to V_i(s)$. All together then, we have from (15),

$$V_i(s) = (1-\lambda_i)r(s, \overline{\gamma}_i(s), \gamma_{-i}(s)) + \lambda_i \int V_i(s')dF(s'/s, \overline{\gamma}_i(s), \gamma_{-i}(s)) \text{ for every } s \in S_C. \quad (16)$$

Recall that the values of a right-continuous function are all determined by its values on a dense subset of its domain. Since $V_i(s), \overline{\gamma}_i(s)$ and $\gamma_{-i}(s)$ are all right continuous, (16) must hold for every $s \in S$.

It follows from (16) and standard results in discounted dynamic programming that $\overline{\gamma}_i(s)$ is a best response to $\gamma_{-i}(s)$. To terminate the proof, it remains only to show that $\overline{\gamma}_i(s)$ is the largest best-response to $\gamma_{-i}(s)$. Recall from the proof of Lemma (6) that the best-response correspondence to $\gamma_{-i}(\cdot)$ is u.h.c. from above and has the property that all its selections are nondecreasing. Hence, being u.s.c., $\overline{\gamma}_i$ must be the (unique) largest best-response by Lemma 4. ■

We are now ready for the

Proof of Theorem 2. It is easy to see that a pair of (Markov-stationary) strategies is a stationary equilibrium if it is a fixed point of the mapping B. Since B is a continuous operator in the $weak^*$ topology from $\widetilde{M_1}(S,A) \times ... \times \widetilde{M_n}(S,A)$ to itself, and since the latter is compact in the product $weak^*$ topology (by the Alaoglu-Bourbaki theorem) and also clearly convex, the existence of a fixed-point follows directly from Shauder's fixed-point theorem. ■

We now move on to the proof of Theorem 1, and the argument proceeds in several steps here as well. We define the following auxiliairy games. Let $v = (v_1, ..., v_n) \in M(S, R)^n$ be an n-vector of continuation values, and consider an n -person one-shot game G_v parametrized by the state variable, where Player i has as strategy space the set of all Borel measurable functions from S to A_i, and as payoff function

$$\Pi_i(v, s, a_i, a_{-i}) \triangleq (1 - \lambda_i) r_i(s, a_i, a_{-i}) + \lambda_i \int v_i(s') dF(s'/s, a_i, a_{-i}). \tag{17}$$

For each fixed $s \in S$, let the game where Player i has action set A_i and payoff (17) be denoted by G_v^s.

Lemma 9 *For any $v = (v_1, ..., v_n) \in M(S, R)^n$ and any fixed $s \in S$, G_v^s is a supermodular game.*

Proof. We first prove that $\Pi_i(v, s, a_i, a_{-i})$ has the requisite complementarity properties. By Theorem 12 and Assumption (T.3), since v is nondecreasing, $\int v_i(z') dF(z'/z, a_i, a_{-i})$ is supermodular in a_i, and has strictly nondecreasing differences in (a_i, a_{-i}). Since both these properties are preserved under scalar multiplication and addition, it follows from Assumption (R.3) that Π_i is supermodular in a_i and has increasing differences in $(a_i; a_{-i})$.

Next, it follows from Lemma 3 that $\Pi_i(v, s, a_i, a_{-i})$ is jointly u.s.c. in (a_i, a_{-i}). Finally, since each $A_i(s)$ is compact, G_v^s is a supermodular game for each $s \in S$. ■

Lemma 10 *For any $v = (v_1, ..., v_n) \in M(S, R)^n$, the game G_v has a largest Nash equilibrium $\overline{a}^v(s) = (\overline{a}_1^v(s), ..., \overline{a}_n^v(s))$, which is such that each $\overline{a}_i^v(s)$ is a nondecreasing u.s.c. function of s.*

Proof of Lemma ?. Since G_v^s is a supermodular game for each $s \in S$, it has a largest Nash equilibrium for each s, by Tarski's fixed-point theorem. Call it $\overline{a}^v(s) = (\overline{a}_1^v(s), ..., \overline{a}_n^v(s))$. By Assumptions (T.3) and (R.3), $\int v_i(s')dF(s'/s, a_i, a_{-i})$ and $r_i(s, a_i, a_{-i})$ have nondecreasing differences in (s, a_i) for each a_{-i}. Hence, so does $\Pi_i(v, s, a_i, a_{-i})$. By Theorem 15 (ii), $\overline{a}^v(s) = (\overline{a}_1^v(s), ..., \overline{a}_n^v(s))$ is nondecreasing in $s \in S$.

We now show that each $\overline{a}_i^v(s)$ is u.s.c. and continuous from above in $s \in S$. Suppose not. Then there is some $s \in S$ such that $\overline{a}_i^v(.)$ is not continuous from above at s. Let $\sigma_i(\cdot)$ coincide with $\overline{a}_i^v(\cdot)$ everywhere except (possibly) at s where $\sigma_i(\cdot)$ is continuous from above, for each i.

The argument in the proof of Lemma 6 is clearly valid here (the difference being that v is "exogenous" here, and "endogenous" there) and it shows that Player i's best-response to $\sigma_{-i}(s)$ is u.h.c. from above at s, so that its maximal selection at s must be $\sigma_i(s)$, cf. Lemma 4. Hence, $\sigma(s)$ is a Nash equilibrium of the game G_v^s, which is larger than $\overline{a}^v(s)$, by construction. Since this is a contradiction to the definition of $\overline{a}^v(s)$, we conclude that $\overline{a}^v(s)$ is continuous from above and u.s.c. at all $s \in S$. ■

Let $\Pi_i^*(v, s)$ denote the equilibrium payoff set of the game G_v. In other words,

$$\Pi_i^*(v, s) = \left\{(1 - \lambda_i)r_i(s, a^v(s)) + \lambda_i \int v_i(s')dF(s'/s, a^v(s)) : a^v(\cdot) \text{ is Nash equilibrium of } G_v\right\}$$

Lemma 11 *For all $v \in M(S, R)$, the maximal selection $\overline{\Pi}_i^*(v, s)$ of $\Pi_i^*(v, s)$ is well-defined and satisfies*

$$\overline{\Pi}_i^*(v, s) = (1 - \lambda_i)r_i(s, \overline{a}^v(s)) + \lambda_i \int v_i(s')dF(s'/s, \overline{a}^v(s)). \quad (18)$$

Furthermore, $\overline{\Pi}_i^*(v, s) \in M(S, R)$.

Proof of Lemma. From Assumptions (R2) and (T2), and the well-known characterization of first-order stochastic dominance, we know that each player's payoff in the game G_v^s is nondecreasing in the rivals' actions. Hence, by applying Theorem 15 (i) for each $s \in S$, we deduce that the equilibrium $\overline{a}^v(s)$ is the Pareto-dominant equilibrium. In other words,

(18) must hold, with $\overline{\Pi}_i^*(v,s)$ being the largest equilibrium payoff in the game G_v^s, for each $s \in S$.

We now show that $\overline{\Pi}_i^*(v,s)$ is nondecreasing in s. Let $s_1 \geq s_2$. Then

$$\begin{aligned}
\overline{\Pi}_i^*(v,s_1) &= (1-\lambda_i)r_i(s_1,\overline{a}^v(s_1)) + \lambda_i \textstyle\int v_i(s')dF(s'/s_1,\overline{a}^v(s_1)) \\
&\geq (1-\lambda_i)r_i(s_1,\overline{a}_i^v(s_2),\overline{a}_{-i}^v(s_1)) + \lambda_i \textstyle\int v_i(s')dF(s'/s_1,\overline{a}_i^v(s_2),\overline{a}_{-i}^v(s_1)) \\
&\geq (1-\lambda_i)r_i(s_1,\overline{a}_i^v(s_2),\overline{a}_{-i}^v(s_2)) + \lambda_i \textstyle\int v_i(s')dF(s'/s_1,\overline{a}_i^v(s_2),\overline{a}_{-i}^v(s_2)) \\
&= \overline{\Pi}_i^*(v,s_2),
\end{aligned}$$

where the first inequality follows from the Nash property and Assumption (A5), and the second from Assumptions (R2) and (T2).

To show that $\overline{\Pi}_i^*(v,s)$ is u.s.c. in s, consider

$$\overline{\Pi}_i^*(v,s_1) = \max_{a_i \in \widetilde{A}_i(s)} \left\{ (1-\lambda_i)r_i(s,a_i,\overline{a}_{-i}^v(s)) + \lambda_i \int v_i(s')dF(s'/s,a_i,\overline{a}_{-i}^v(s)) \right\}$$

Since the maximand is jointly u.s.c. in (a,s) and $\widetilde{A}(\cdot)$ is u.h.c., $\overline{\Pi}_i^*(v,s)$ is u.s.c. in s by the Maximum Theorem.

Finally, the fact that $\overline{\Pi}_i^*(v,s) \leq K$ being obvious, we have overall shown that $\overline{\Pi}_i^*(v,s) \in M(S,R)$ whenever $v \in M(S,R)$. ■

Proof of Theorem 1. The argument follows by backward induction, based on iteration of the mapping $v = (v_1,...,v_n) \longrightarrow \Pi^*(v,s) = (\Pi_1^*(v,s),...,\Pi_n^*(v,s))$. Clearly, with $v_0 \equiv (0,0,...,0)$, $v^1 = (v_1^1,...,v_n^1) \triangleq \Pi^*(v_0,s)$ is the equilibrium value function vector for the one-shot game, with player i's payoff function given by $(1-\lambda_i)r_i(s,a_i,a_{-i})$. Likewise, $v^2 = (v_1^2,...,v_n^2) \triangleq \Pi^*(v_1,s)$ is the equilibrium value function vector for the two-period game, with player i's payoff function given by $(1-\lambda_i)r_i(s,a_i,a_{-i}) + \int v_i^1 dF(s,a_i,a_{-i})$, and so on until the last period in the horizon T. By Lemmas 10 and 11, this process clearly generates a Markov equilibrium with strategy components in $\widetilde{M}(S,A)$ and corresponding value functions in $M_K(S,R)$. This completes the proof of Theorem 1. ■

4. On some Economic Applications

While the list of assumptions required for our main resuts is long and seemingly overly restrictive, we argue in this section that the results are actually relatively widely applicable in economics, in view of the natural monotonicity and complementarity conditions that commonly characterize many problems in strategic economic dynamics. The presentation below is somewhat informal, in that various regularity conditions conveniently used in the theory will not be dealt with systematically here.

4.1 Dynamic search with learning.

Consider the following infinite-horizon search model, which generalizes the model devised by Curtat (1996) as a dynamic extension of Diamond's (1982) static model. At every stage, each of N traders expands effort or resources searching for trading partners. Denoting by $a_i \in [0, 1]$ the effort level of agent i, by $C_i(a_i)$ the corresponding search cost, and by s the current productivity level of the search process, i's one-stage reward and the state transition probability are given by

$$r_i(s, a) = sa_i \sum_{j \neq i} a_j - C_i(a_i) \text{ and } \; s' \sim F(\cdot/s, a)$$

It is easy to verify that the one-period reward satisfies Assumptions (R1)-(R4). It is clearly natural to have $F(\cdot/s, a)$ stochastically increasing in (s, a) as in Assumption (T2). Given the scalar nature of the state and actions here, Assumption (T3) requires $1 - F(s'/\cdot)$ to be supermodular in (s, a) for every s', which has the following natural complementarity interpretation: The probability that the next productivity is higher than any given level increases more due to a change in a player's search level when the other players search harder and/or current productivity is higher.

As a special case of this transition law, one may consider $s' \sim \widetilde{F}(\cdot/s + \sum_j a_j)$ and assume that $\widetilde{F}(s'/\cdot)$ is decreasing and concave, for each $s' \in S$. This transition law is easily seen to satisfy Assumptions (T1)-(T3), the verification details being left out. The assumptions of monotonicity and concavity on $F(s'/\cdot)$ have the following economic "increasing returns" interpretation in this context: The probability of the next search productivity index being greater than or equal to any given level s' increases at an increasing rate with the current index and agents' effort levels. In other words, $1 - F(s'/\cdot)$ is increasing and convex, for each $s' \in S$.

Adding Assumption (T1)*, this model with discounted rewards fits as a special case of our general framework. We conclude that a pure-strategy Markov-stationary equilibrium exists, and has the property that effort levels are nondecreasing functions of the current search productivity index.

4.2 Bertrand Competition with Learning-by-doing

Consider the following model of price competition with substitute goods, and constant unit costs of production that are lowered over time as a consequence of learning-by-doing. Let s denote the state of know-how in the industry, common to all firms, and $c_i(s)$ the cost per unit of output of firm i. Let a_i denote firm i's price and $D_i(a)$ its demand function. Here, firm i's per-period profit and the state transition law are given by

$$r_i(s,a) = (a_i - c_i(s))D_i(a) \text{ and } s' \sim F(\cdot/s,a)$$

Assuming[15] that $c_i^{'}(s) \leq 0$ and that firm i's demand satisfies the standard assumptions $\frac{\partial D_i(a)}{\partial a_j} > 0$ and $\frac{\partial D_i(a)}{\partial a_j} + [a_i - c_i(s)]\frac{\partial^2 D_i(a)}{\partial a_j \partial a_i} \geq 0$, it is easily verified that the one-period reward r_i is supermodular in $(-a,s)$. It is clearly natural to have $F(\cdot/s,a)$ stochastically increasing in $(-a,s)$ as lower prices lead to higher demands overall, and thus higher production levels, or higher industry-wide learning-by-doing for the firms (Assumption (T2)). Given the scalar nature of the state and actions here, Assumption (T3) requires $1-F(s'/\cdot)$ to be supermodular in $(-a,s)$ for every s', which has the following natural complementarity interpretation: The

[15]For the sake of brevity, we omit the description of some regularity conditions (such as boundedness, compactness,...) on the primitives of the model here.

probability that the next industry-wide know-how level is higher than any given target increases more due to an decrease in a firm's price when the other firms' prices are lower and/or current know-how is higher. Since firm i's price set when the state is s is given by $[c_i(s), \infty)$, Assumption (A.

Adding Assumption (T1)*, this model with discounted rewards fits as a special case of our general framework. Hence, a pure-strategy Markov-stationary equilibrium exists, and has the property that prices are nonincreasing functions (due to the sign change in $(-a, s)$) of the current know-how level.

Two versions of dynamic Cournot competition with learning-by-doing can be accomodated within our general framework. One is Curtat's (1996) model with complementary products. Omitting the diagonal dominance conditions given by Curtat's (1996), but keeping all his other assumptions, our main result would apply to his model. The second model would consider homogeneous products and rely on change-of-order arguments to fit the framework at hand.

4.3 Resource Extraction

Consider two agents noncooperatively exploiting a natural resource or some other common-property stochastically productive asset. Each agent seeks to maximize his discounted sum of utilities over time, depending only on his own consumption levels c_t^i. The one-period reward of agent i and the state transition or growth law are given by

$$\sum_{t=0}^{\infty} \lambda_i^t u_i(c_t^i) \text{ and } s' \sim F(\circ/s - c_t^1 - c_t^2)$$

This problem was considered by Amir (1996a) who showed existence of a Markov-stationary equilibrium with strategies having slopes in $[0, 1]$. It can be shown that this simple model is not a special case of our general framework, nor of Curtat's. However, once one restricts the strategy space of one player to the class of Liptschitz-continuous functions just described, one can use exactly the same approach as in Curtat (1996) to carry the analysis through.

Thus while this model is not a special case, it can be approached in essentially the same way. Details may be found in Amir (1996a).

5. Appendix

A brief summary of the lattice-theoretic notions and results is presented here.

Throughout, S will denote a partially ordered set and A a lattice, and all cartesian products are endowed with the product order. A function F: $A \to R$ is (strictly) supermodular if $F(a \vee a') + F(a \wedge a') \geq (>) F(a) + F(a')$ for all $a, a' \in A$. If $A \subset R^m$ and F is twice continuously differentiable, F is supermodular if and only if $\frac{\partial^2 F}{\partial a_i \partial a_j} \geq 0$, for all $i \neq j$. A function $G : A \times S \to R$ has (strictly) increasing differences in s and a if for $a_1 (>) \geq a_2$, $G(a_1, s) - G(a_2, s)$ is (strictly) increasing in s. If $A \subset R^m$, $S \subset R^n$ and G is smooth, this is equivalent to $\frac{\partial^2 G}{\partial a_i \partial s_j} \geq 0$, for all $i = 1, ..., m$ and $j = 1, ..., n$.

A set I in R^n is increasing if $x \in I$ and $x \leq y \Rightarrow y \in I$. With $S \subset R^n$ and $A \subset R^m$, a transition probability F from $S \times A$ to S is supermodular in a (has increasing differences in s and a) if for every increasing set $I \subset R^n$, $\int 1_I(t) dF(t/s, a)$ is supermodular in a (has increasing differences in s and a) where 1_I is the indicator function of I. A characterization of these properties, using first-order stochastic dominance, follows (see Athey, 1998-1999 for an extensive study of this class of results, including ordinal ones):

Theorem 12 *(Topkis, 1968). A transition probability F from $S \times A$ to $S \subset R^n$ is supermodular in s (has increasing differences in s and a) if and only if for every integrable increasing function $v : S \to R$, $\int v(t) dF(t/s, a)$ is supermodular in s (has increasing differences in s and a).*

Let $L(A)$ denote the set of all sublattices of A. A set-valued function $H : S \to L(A)$ is ascending if for all $s \leq s'$ in S, $a \in A_s$, $a' \in A_{s'}$, $a \vee a' \in A_{s'}$ and $a \wedge a' \in A_s$. Topkis's main monotonicity result follows (also see Milgrom and Shannon, 1994):

Theorem 13 *(Topkis, 1978). Let $F : S \times A \to R$ be upper semi-continuous and supermodular in a for fixed s, and have increasing (strictly increasing) differences in s and a, and $H : S \to L(A)$ be ascending. Then the maximal and minimal (all) selections of* $\arg\max \{F(s, a) : a \in H(s)\}$ *are increasing functions of s.*

A game with action sets that are compact Euclidean lattices and payoff functions that are u.s.c. and supermodular in own action, and have increasing differences in (own action, rivals' actions) is a supermodular game. By Theorem 5.2, such games have minimal and maximal best-responses that are monotone functions, so that a pure-strategy equilibrium exists by (see also Vives, 1990):

Theorem 14 *(Tarski, 1955). An increasing function from a complete lattice to itself has a set of fixed points which is itself a nonempty complete lattice.*

The last result deals with comparing equilibria.

Theorem 15 *A.5 (Milgrom and Roberts, 1990).*

(i) If a supermodular game is such that each payoff is nondecreasing in rivals'actions, then the largest (smallest) equilibrium is the Pareto-best (worst) equilibrium.

(ii) Consider a parametrized supermodular game where each payoff has increasing differences in the parameter (assumed real) and own action. Then the maximal and minimal equilibria are increasing functions of the parameter.

REFERENCES

Amir, R. (1996a), Continuous stochastic games of capital accumulation with convex transitions, *Games and Economic Behavior*, 16, 111-131.

Amir, R. (1996b), Strategic intergenerational bequests with stochastic convex technology, *Economic Theory*, 8, 367-376.

Amir, R. (2003), Stochastic games in economics and related fields I: A general overview, in Neyman, A. and Sorin, S., eds., *Stochastic Games and Applications*. Dordrecht and Boston: Kluwer.

Athey, S. (2002), Comparative statics under uncertainty: Log supermodularity and single-crossing properties, *Quarterly Journal of Economics*,

Athey, S. (2001), Single-crossing properties and existence of pure-strrategy equilibria in games of incomplete information, *Econometrica*, 69, 861-89.

Bertsekas, D. and S. Shreve (1978), *Stochastic optimal control: The discrete-time case*, New-York: Academic Press.

Basar, T. and G. Olsder (1999), *Dynamic Noncooperative Game Theory*, Academic Press, New-York.

Curtat, L. (1996), Markov equilibria of stochastic games with complementarities, *Games and Economic Behavior*, 17, 177-199.

Diamond, P. (1982), Aggregate demand management in search equilibrium, *Journal of Political Economy*, 90, 881-894.

Duffie, D., J. Geanakoplos, A. Mas-Collell and A. Mc Lennan (1988), Stationary Markov equilibria, *Econometrica*, 62, 745-781.

Ericson, R. and A. Pakes (1995), Markov-perfect industry dynamics: A framework for empirical work, *Review of Economic Studies,* 62, 53-82.

Goffman, C. (1953), Real Functions, Boston: Prindle, Weber and Schmidt.

Granot, F. and Veinott, A. (1985), Substitutes, complements, and ripples in network flows, *Mathematics of Operations Research*, 10, 471-497.

Leininger, W. (1986), Existence of perfect equilibrium in a model of growth with altruism between generations, *Review of Economic Studies,* 53, 349-67.

Levhari, D. and L. Mirman (1980), The great fish war: An example using a dynamic Cournot-Nash solution, *Bell Journal of Economics*, 322-344.

Milgrom, P. and J. Roberts(1990), "Rationalizability, learning, and equilibrium in games with strategic complementarities", *Econometrica*, 58, 1255-78.

Milgrom, P. and C. Shannon (1994), "Monotone comparative statics", *Econometrica*, 62, 157-180.

Neyman, A. and S. Sorin, eds. (2003), *Stochastic Games and Applications.* Dordrecht and Boston: Kluwer.

Nowak, A. (2002), On a new class of nonzero sum discounted stochastic games having stationary Nash equilibrium points, *International Journal of Game Theory*, 32, 121-132.

Nowak, A. and T. Raghavan(1992), Existence of stationary correlated equilibria with symmetric information for discounted stochastic games, *Mathematics of Operations Research* 17, 519-526.

Shapley, L. (1953), Stochastic games, *Proceedings of the National Academy of Sciences of the USA*, 39, 1095-1100.

Tarski, A. (1955), A lattice-theoretical fixpoint theorem and its applications, *Pacific Journal of Mathematics*, 5, 285-309.

Topkis, D. (1968), Ordered optimal solutions, PhD thesis, Stanford Univ.

Topkis, D. (1978), "Minimizing a submodular function on a lattice, *Operations Research*, 26, 305-321.

Vives, X. (1990), "Nash equilibrium with strategic complementarities", *Journal of Mathematical Economics*, 19, 305-321.