Mean Field Equilibrium in Dynamic Games with Strategic Complementarities

Sachin Adlakha

Center for Mathematics of Information, California Institute of Technology, Pasadena, California 91125, adlakha@caltech.edu

Ramesh Johari

Department of Management Science and Engineering, Stanford University, Stanford, California 94305, ramesh.johari@stanford.edu

We study a class of stochastic dynamic games that exhibit *strategic complementarities* between players; formally, in the games we consider, the payoff of a player has increasing differences between her own state and the empirical distribution of the states of other players. Such games can be used to model a diverse set of applications, including network security models, recommender systems, and dynamic search in markets. Stochastic games are generally difficult to analyze, and these difficulties are only exacerbated when the number of players is large (as might be the case in the preceding examples).

We consider an approximation methodology called *mean field equilibrium* to study these games. In such an equilibrium, each player reacts to only the long-run average state of other players. We find necessary conditions for the existence of a mean field equilibrium in such games. Furthermore, as a simple consequence of this existence theorem, we obtain several natural monotonicity properties. We show that there exist a "largest" and a "smallest" equilibrium among all those where the equilibrium strategy used by a player is nondecreasing, and we also show that players converge to each of these equilibria via natural *myopic learning dynamics*; as we argue, these dynamics are more reasonable than the standard best-response dynamics. We also provide sensitivity results, where we quantify how the equilibria of such games move in response to changes in parameters of the game (for example, the introduction of incentives to players).

1. Introduction

This paper studies a class of games that exhibit *strategic complementarities* between players. A strategic complementarity exists if, informally, "higher" actions by other players increase the return to higher actions for a given player. Games with strategic complementarities are a powerful modeling tool, applicable in a wide range of situations, including: systems with positive network effects (such as network security models, recommender systems, and social networks); coordination problems; dynamic search in markets; social learning; and oligopoly models (for example, quantity or price competition with complementarities).

Our focus in this paper is on *dynamic* games with strategic complementarities. Strategic complementarities have long provided a fertile analytical ground for static game-theoretic models; see, for example, Milgrom and Roberts (1990), Vives (1990), and Topkis (1998). However, the literature on dynamic games with complementarities has emerged relatively recently by comparison. Much of the attention in prior work on such games has focused on developing existence proofs for equilibrium; see, for example, Curtat (1996), Amir (2002, 2005), Vives (2009), Horst (2005), Nowak (2007), and Sleet (2001) for such results.

In this paper we consider a class of dynamic games referred to as *stochastic games*; in these games, agents' actions directly affect underlying state variables that influence their payoff (Shapley 1953). The standard solution concept for stochastic games is *Markov-perfect equilibrium* (Fudenberg and Tirole 1991). Despite the previously cited existence results for Markov-perfect equilibria in games with complementarities, there remain two significant obstacles, particularly as the number of players grows large. First is *computability*: the state space of the preceding games expands in dimension with the number of players, and thus the "curse of dimensionality" kicks in, making computation of Markov-perfect equilibria essentially infeasible (Pakes and McGuire 2001, Doraszelski and Pakes 2007). Second is *plausibility*: as the number of players grows large, it becomes increasingly difficult to believe that individual players track the exact behavior of the other agents. Rather than treat the growth of the population as an impediment to analysis, this paper addresses these obstacles by exploiting an asymptotic regime where the number of players grows large to *simplify* analysis of equilibria.

We study an equilibrium concept we call *mean field equilibrium*. In a limiting "mean field" regime where the number of players becomes infinite, it becomes reasonable to suppose that individuals postulate that fluctuations in the

empirical distribution of other players' states have "averaged out" due to large scale, and thus optimize holding the state distribution of other players fixed. Accordingly, we consider an approximation methodology where agents optimize only with respect to long-run average estimates of the distribution of other players' states. On the other hand, there should be a consistency check: the postulated state distribution should arise from the optimal strategies that agents compute. The resulting equilibrium concept is mean field equilibrium (MFE); this notion has been utilized across a range of work in economics, operations research, and control (as we discuss below). Although it should be viewed as an equilibrium concept for a formal limit stochastic game model with an infinite number of players, we note that it can be justified via approximation results that relate mean field equilibria to equilibria of games with a finite but large number of players.

Our results provide valuable insight into the structure of mean field equilibria in stochastic games with complementarities, as well as computational tools to determine such equilibria. To motivate our results, we first provide several examples of stochastic games with complementarities where the approach taken in this paper applies. These examples often exhibit large numbers of players, and thus the benefits of mean field equilibrium are significant. We demonstrate in §7 that each of these examples can be analyzed using the results we develop in this paper.

EXAMPLE 1 (INTERDEPENDENT SECURITY). In *interdependent security* games, as introduced in Kunreuther and Heal (2003), a large number of agents make individual decisions about their own security. However, the ultimate security of an agent depends on the security decisions made by other agents. For example, imagine a network of computers where each individual user makes an investment in keeping her own machine secure. This investment may be in the form of advanced antivirus filters, firewalls, etc. Although these investments improve the security of the individual computer, it can still be affected if the other computers in the network are not properly secured. In the interdependent security games we consider, agents take actions at some cost to improve their own security level, and earn a payoff each period that depends on whether or not a security breach occurs. The fact that the probability of a security breach is influenced by others' security levels introduces strategic complementarities into the stochastic game.

EXAMPLE 2 (COLLABORATIVE FILTERING). Many large online recommendation systems, such as those used by Netflix and Amazon, rely on collaborative filtering. In such systems, if an individual puts forth greater effort in maintaining their profile, the recommendations *they* receive will improve. However, the recommendations *other* individuals receive improve as well, and typically other individuals will feel a stronger incentive to exert additional effort to maintain their profile in this case. In the absence of such effort, the profile of an agent becomes stale and useless both to

her and others in the system. Thus, collaborative filtering systems exhibit strong strategic complementarities.

EXAMPLE 3 (DYNAMIC SEARCH WITH LEARNING). In dynamic search models, traders in a market exert effort to find trading partners (Diamond 1982). Such models are commonly used to study, for example, decentralized matching in labor markets. We consider a model where, at each time step, traders also gain experience by exerting effort; this experience makes future effort more productive. Of course, traders' experience increases as *they* put forth more effort, but their experience also increases as *others* put forth more effort since this increases the likelihood of useful interactions per unit effort. This creates strategic complementarities between the players; such a model was considered by Curtat (1996). □

Our main results provide conditions that ensure existence of mean field equilibria in stochastic games with complementarities. We also establish that simple learning procedures converge to equilibria, and provide insight into sensitivity of equilibria to parameter changes. We consider a general class of models with parsimonious assumptions over model primitives that ensure strategic complementarities. In particular, our model class allows players to be coupled *both* via their payoff function and state transitions; that is, players' payoffs and state transitions can depend on states or actions of other players. We also discuss extensions of our results to models with multidimensional state and action spaces, and with heterogeneity among players. Details of our results follow.

1. *Structural characterization of mean field equilibrium.* We establish existence of a mean field equilibrium in a general stochastic game model using lattice-theoretic techniques. Lattice-theoretic methods are typically applied in games with complementarities; the key techniques we use are attributed to Tarski (1955), Kamae et al. (1977), Hopenhayn and Prescott (1992), Zhou (1994), and Topkis (1998). Despite the use of lattice-theoretic techniques in our analysis, existence of equilibria in our game cannot be inferred from existence results for other games in the literature. Moreover, we show that there exists a "largest" and "smallest" equilibrium among the set of all mean field equilibria with nondecreasing strategies. Thus, in particular, there is a natural dominance relationship among the mean field equilibria of a given stochastic game with complementarities. This is particularly valuable in dynamic games, because our characterization applies to the *distribution* of states of agents in equilibrium.

2. *Convergence to equilibrium.* We provide two convergence results. First, we study a standard *best-response dynamic* (BRD). In this algorithm, at each time step, each agent computes the stationary population state distribution that would be induced by the current strategies of others, and in turn computes the best response to that state distribution. Using monotonicity properties derived in establishing

the existence of mean field equilibrium, we show that BRD converges.

However, BRD is unsatisfying both computationally and practically. From a computational standpoint, BRD requires computation of a stationary distribution given the current strategy choices of agents in the system; this is in principle a complex procedure to execute at each iteration. More importantly, BRD is an implausible approach to play in an actual game: it is unlikely that agents would explicitly compute the stationary distribution their competitors would obtain.

Instead, we consider a more a natural form of *myopic learning dynamics* (MLD) among the players; convergence of MLD is a central insight of our paper. In particular, suppose that, initially, each agent starts at the lowest (respectively, highest) possible state. At each time step, agents observe the current empirical population state distribution, and conjecture that this distribution will *remain constant for all time*; with this conjecture they compute an optimal strategy, and play in the next period according to that strategy. At the next time step, the state distribution will evolve, and agents repeat the same heuristic. We show that this dynamic converges to the lowest (respectively, highest) mean field equilibrium among all equilibria with nondecreasing strategies.

Note that MLD resolves both the *computability* and *plausibility* issues raised above. First, it is a natural, simple, implementable algorithm for finding a mean field equilibrium; indeed, MLD has some similarities with model predictive control or receding horizon control (Garcia et al. 1989), both popular approaches to complex dynamic control problems. Second, it corresponds to a learning dynamic that demands only a weak form of rationality and forecasting from the players, and yet yields an equilibrium in the limit.

3. *Separable stochastic games.* Although appealing, the general theory does pose some significant issues in application: the complementarity requirements on model primitives may preclude important and interesting cases of practical interest. Complementarity is a strong requirement, but also brittle: a model that does not appear to satisfy the assumptions a priori may do so through a judicious change of variables. We employ this fact to show that a range of games that do not satisfy the assumptions of our baseline model can be studied by a suitable change of variables, provided that the payoff is separable in the state and action of a given player—often a relatively mild assumption. Notably, models with *linear* dynamics fall in this class. This greatly expands the set of models that can be analyzed within our framework.

4. *Sensitivity.* Finally, essentially for free, the complementarity structure allows us to analyze changes in the equilibrium in response to changes in parameters of the game. In particular, we can predict shifts (in a first-order stochastic dominance sense) of the equilibrium state distribution of players in response to exogenous parameter changes. Such sensitivity analysis, or *comparative statics*, allows our model to address, for example, the value of incentives to increase security levels, or the value of increasing the quality of recommendations by a given factor.

The remainder of the paper is organized as follows. In §2 we introduce our basic stochastic game model as well as the formal definition of mean field equilibrium. Notably, we also discuss a justification for the use of mean field equilibrium: that it approximates equilibria of finite games well. This approximation property has been developed in a variety of specific contexts in the past (see, for example, Glynn 2004, Huang et al. 2006, Weintraub et al. 2008, and Tembine et al. 2009), and in our context we apply the methodology developed in Adlakha et al. (2013) (inspired by Weintraub et al. 2008) to justify mean field equilibrium as a limiting notion of equilibrium.

Next, in §3, we define stochastic games with complementarities. We then prove our first main result: that a mean field equilibrium exists for a stochastic game with complementarities. In §3.2, we show that equilibria are "ordered," in the sense that there exist a smallest and largest mean field equilibrium among all those where the equilibrium strategy is nondecreasing. In §4, we prove convergence of both the BRD and MLD algorithms described above. We also discuss the performance of MLD in finite systems.

In §5, we provide comparative statics results for the games under consideration. In §6, we consider separable stochastic games with complementarities (as described above) and establish that these are a special case of our basic model of stochastic games with complementarities.

In §7, we revisit each of the examples described above. In particular, we provide formal verification that these examples satisfy the assumptions made in the paper to obtain existence and convergence results. Section 8 concludes with a discussion of extensions to include both player heterogeneity (that is, type information) and multidimensional state and/or action spaces. Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2013 .1192. Supplementary material and omitted proofs are deferred to the electronic companion; in particular, in §EC.1 there, we extend our results to cover games where players' payoffs and transition kernels may depend on the *actions* of others, rather than their states.

We conclude by briefly surveying related work in two different areas: on stochastic games with complementarities, and on mean field equilibrium. First, we note that prior literature has established existence of equilibrium in stochastic games with complementarities; however, these results are typically obtained under somewhat restrictive assumptions and do not yield explicit algorithms for computation of equilibria. For example, several such results require use of topological fixed-point theorems such as Kakutani's theorem (Curtat 1996; Amir 2002, 2005; Horst 2005). In Nowak (2007), the author provides computational methods for several two-player stochastic games with

complementarities, but only with specific transition dynamics: the transition kernel is an additive combination of finitely many measures on the state space. Another related work is Vives (2009), where the author derives conditions under which static complementarities lead to dynamic complementarities; however, he does not deal with the question of existence of equilibrium.

The preceding works all consider games with finitely many players, whereas mean field equilibrium is defined in a limiting regime where the number of players grows large. The notion of mean field equilibrium is inspired by mean field models in physics, where large systems exhibit macroscopic behavior that is considerably more tractable than their microscopic description. (See, for example, Mézard and Montanari 2009 for background, and Blume 1993 and Morris 2000 for related ideas applied to static games.) In the context of stochastic games, mean field equilibrium and related approaches have been proposed under a variety of monikers across economics and engineering; see, for example, studies of anonymous sequential games (Jovanovic and Rosenthal 1988, Bergin and Bernhardt 1995, Chakrabarti 2003); stationary equilibrium (Hopenhayn 1992); dynamic stochastic general equilibrium in macroeconomic modeling (Stokey et al. 1989); Nash certainty equivalent control (Huang et al. 2006, 2007); mean field games (Lasry and Lions 2007); oblivious equilibrium (Weintraub et al. 2008, 2011); and dynamic user equilibrium (Friesz et al. 1993, Wunderlich et al. 2000). Mean field equilibrium has also been studied in recent works on information percolation models (Duffie et al. 2009), sensitivity analysis in aggregate games (Acemoglu and Jensen 2013), coupling of oscillators (Yin et al. 2010), scaling behavior of markets (Bodoh-Creed 2012), and on power control in wireless communications (Wiecek et al. 2011). Most closely related to our paper is the work of Sleet (2001), who considers mean field equilibria of a dynamic price-setting game with stochastic, exogenous firm-specific demand shocks per period, that exhibits strategic complementarities. The general analytical techniques in this paper can be applied to recover the existence result for that game.

2. Model and Definitions

In this section we begin with preliminaries. We define a general model of a stochastic game in §2.1. In §2.2, we present the standard solution concept for stochastic games with finitely many players, *Markov-perfect equilibrium*. By contrast, in §2.3 we define *mean field equilibrium*, and in §2.4 we provide a formal justification for mean field equilibrium as an approximation to Markov-perfect equilibria in games with a large finite number of players. Finally, in §2.5, we discuss lattice-theoretic preliminaries necessary for the analysis in the sequel.

2.1. Stochastic Games

Stochastic games are a general model for dynamic strategic interactions among multiple agents. The stochastic games we consider are *anonymous*: the interaction between players is via aggregate information about the state.

Formally, an *anonymous stochastic game* is a tuple $\Gamma = (m, \mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$ defined as follows.

Time. The game is played in discrete time, with time periods by $t = 0, 1, 2, \ldots$.

Players. There are m players in the game; we use i to denote a particular player.

State. The state of player i at time t is denoted by $x_{i,t} \in \mathcal{X}$, where $\mathcal{X} \subset \mathbb{R}$ is compact.

Population state. Let $f_{-i,t}(y)$ denote the fraction of players (excluding player i) that have their state as y at time t, that is,

$$f_{-i,t}(y) = \frac{1}{m-1} \sum_{j \neq i} \mathbf{1}_{\{x_{j,t}=y\}}, \tag{1}$$

where $\mathbf{1}_{\{x_{j,t}=y\}}$ is the indicator function that the state of player j at time t is y. We refer to $f_{-i,t}$ as the *population state* at time t (from player i's point of view). Note that it is a discrete probability measure on the state space $\mathcal{X}$.

Action. The action taken by player i at time t is denoted by $a_{i,t}$. The set of feasible actions when the player is in state x is a compact set $\mathcal{A}(x) \subset \mathbb{R}$. We let $\mathsf{A} = \cup_{x \in \mathcal{X}} \mathcal{A}(x)$, and assume that A is compact as well.

Transition probabilities. The state of a player evolves according to the following Markov process. If the state of player i at time t is $x_{i,t} = x$, the player takes an action $a_{i,t} = a \in \mathcal{A}(x)$ at time t, and the population state is $f_{-i,t} = f$, then the next state is distributed according to the Borel probability measure $\mathbb{P}(\cdot \mid x, a, f)$, where for Borel sets $S \subset \mathcal{X}$,

$$\mathbb{P}(S \mid x, a, f)$$
$$= \mathrm{Prob}(x_{i,t+1} \in S \mid x_{i,t} = x, a_{i,t} = a, f_{-i,t} = f). \tag{2}$$

Further, given $x_{i,t}$, $a_{i,t}$, and $f_{-i,t}$, the next state $x_{i,t+1}$ is conditionally independent of all other past history of the game.

Payoff. The single-period payoff to player i at time t is $\pi(x_{i,t}, a_{i,t}, f_{-i,t}) \in \mathbb{R}$. Note that all players have the same payoff function, and it is independent of the actions taken by other players; we relax the latter assumption in §EC.1 in the e-companion.

Discount factor. The players discount their future payoff by a discount factor $0 < \beta < 1$. Thus, a player i's infinite-horizon payoff is given by: $\sum_{t=0}^{\infty} \beta^t \pi(x_{i,t}, a_{i,t}, f_{-i,t})$.

The players are coupled to each other through the transition probabilities and payoff function. Note that the transition probability kernel and the payoff function depend on the number of players only via the population state $f_{-i,t}$; the fact that this is aggregate (identity-independent) information about other players captures the notion of anonymity. The examples discussed in the introduction naturally belong to the class of anonymous stochastic games. For example, in the interdependent security model (Example 1), it is natural to assume that a single player's payoff is affected by the empirical distribution of

security levels of other players in the network, but not by their specific identity. The same assumption is also plausible for the other examples presented earlier (see §7 for details on how the examples discussed in the Introduction can be analyzed within our framework).

Before we proceed, we introduce some additional useful notation. Let $\mathfrak{F}$ be the set of all Borel probability measures on $\mathscr{X}$: $\mathfrak{F} = \{f \colon \mathscr{X} \to [0,1] \,|\, f(x) \geqslant 0, \sum_{x \in \mathscr{X}} f(x) = 1\}$. Throughout the paper, we endow $\mathfrak{F}$ with the topology of weak convergence (Billingsley 1968). In addition, we let $\mathfrak{F}^{(m)} \subset \mathfrak{F}$ be the set of probability measures with support on at most m distinct states. Note that in an m-player anonymous stochastic game, the population state distribution $f_{-i,t}$ is an element of $\mathfrak{F}^{(m)}$.

We conclude with the following continuity assumption over the action sets, transition kernel, and payoff function, *which remains in force for the entire paper.*

ASSUMPTION 1. *The point-to-set correspondence $x \mapsto \mathscr{A}(x)$ has a closed graph. Further, the payoff function $\pi(x, a, f)$ and the transition probability measure $\mathbb{P}(\cdot \,|\, x, a, f)$ are both jointly continuous on $\mathscr{X} \times \mathsf{A} \times \mathfrak{F}$ (where we endow $\mathfrak{F}$ with the topology of weak convergence).*

2.2. Markov-Perfect Equilibrium

In studying stochastic games, attention is typically focused on a class of *Markov* strategy spaces, where the action of a player at each time is a function of only the current state of every player (Fudenberg and Tirole 1991, Maskin and Tirole 1988). In the context of anonymous stochastic games, a Markov strategy depends on the current state of the player as well as the current population state. Because a player using such a strategy tracks the evolution of the other players, we refer to such strategies in our context as *cognizant* strategies.

DEFINITION 1. Let $\mathfrak{M}^{(m)}$ be the set of feasible *cognizant strategies* available to a player in an m-player anonymous stochastic game, that is, all Borel-measurable maps from $\mathscr{X} \times \mathfrak{F}^{(m)}$ to $\mathscr{A}$.

Consider an m-player anonymous stochastic game. Letting $\mu_i \in \mathfrak{M}^{(m)}$ denote the cognizant strategy used by player i, we have $a_{i,t} = \mu_i(x_{i,t}, f_{-i,t})$. The next state of player i is randomly drawn according to the kernel $\mathbb{P}$. We let $\boldsymbol{\mu}^{(m)}$ denote the strategy vector where every player has chosen cognizant strategy μ.

Let $V^{(m)}\big(x, f \,|\, \mu', \boldsymbol{\mu}^{(m-1)}\big)$ be the expected discounted payoff for a player with initial state x and with initial population state $f \in \mathfrak{F}^{(m)}$, given that the player follows a cognizant strategy μ' and every other player follows the cognizant strategy μ. In particular, we have

$$V^{(m)}\big(x, f \,|\, \mu', \boldsymbol{\mu}^{(m-1)}\big)$$

$$\triangleq \mathbb{E}\bigg[\sum_{t=0}^{\infty} \beta^t \pi(x_{i,t}, a_{i,t}, f_{-i,t}) \,|\, x_{i,0} = x, f_{-i,0} = f;$$

$$\mu_i = \mu', \boldsymbol{\mu}_{-i} = \boldsymbol{\mu}^{(m-1)}\bigg], \quad (3)$$

where $\boldsymbol{\mu}_{-i}$ denotes the strategies employed by every player except i.

We focus our attention on a *symmetric Markov-perfect equilibrium* (MPE), where all players use the same cognizant strategy μ. In an abuse of notation, we write $V^{(m)}(x, f \,|\, \boldsymbol{\mu}^{(m)})$ to refer to the expected discounted value as given in Equation ((3)) when every player follows the same cognizant strategy μ.

DEFINITION 2 (MARKOV-PERFECT EQUILIBRIUM). The vector of cognizant strategies $\boldsymbol{\mu}^{(m)} \in \mathfrak{M}^{(m)}$ is a *symmetric Markov perfect equilibrium* (MPE) if for all initial states $x \in \mathscr{X}$ and population states $f \in \mathfrak{F}^{(m)}$ we have

$$\sup_{\mu' \in \mathfrak{M}^{(m)}} V^{(m)}(x, f \,|\, \mu', \boldsymbol{\mu}^{(m-1)}) = V^{(m)}(x, f \,|\, \boldsymbol{\mu}^{(m)}).$$

Thus, an MPE is a profile of cognizant strategies that simultaneously maximize the expected discounted payoff for every player, given the strategies of other players. Standard results in dynamic programming ensure that under our assumptions, each player i has an optimal cognizant strategy as long as all other players use a cognizant strategy, even if we allow player i to optimize over possibly history-dependent strategies.[1]

2.3. Mean Field Equilibrium

MPE is complex because players must forecast and track the exact behavior of their competitors; indeed, as a result, guaranteeing existence of MPE can be quite challenging, and computation is typically intractable. On the other hand, in a game with a large number of players, we might expect that fluctuations of players' states "average out," and hence the actual population state remains roughly constant over time. Because the effect of other players on a single player's payoff is only via the population state, it is intuitive that, as the number of players increases, a single player has negligible effect on the outcome of the game. This intuition is formalized through the notion of *mean field equilibrium* (Jovanovic and Rosenthal 1988; Bergin and Bernhardt 1995; Hopenhayn 1992; Stokey et al. 1989; Friesz et al. 1993; Huang et al. 2006, 2007; Lasry and Lions 2007; Weintraub et al. 2008, 2011; Adlakha et al. 2013; Bodoh-Creed 2012).

In mean field equilibrium, each player optimizes its payoff based on only a conjecture of the long-run average population state with an infinite population. Thus, rather than forecasting and tracking the exact population state, a single player's action depends only on her own current state as well as the conjectured long-run average population state. This is motivated by the fact that in a large population in steady state, a single player need not concern herself with the fine scale dynamics of competitors' specific states. Given this simplified player behavior, note that each player must solve a dynamic program to determine their optimal strategy; the strategy chosen by each player then leads to a long-run average state distribution induced by the resulting

transition dynamics. Mean field equilibrium requires that the latter distribution matches the original population state conjecture made by the players.

In this section we formally define mean field equilibrium. We study a formal limit model with an infinite population of agents, where a single agent optimizes against a conjecture of the population state of the others. In the next section, we provide justification of this limit model by establishing that in an asymptotic regime where the number of players approaches infinity, a mean field equilibrium is approximately a MPE (in a sense we make precise).

Note that in this limit model, because players optimize holding the population state constant, their optimal strategies will depend only on their current state. We call such players *oblivious* and refer to their strategies as *oblivious strategies* (following Weintraub et al. 2008). This approach means players *ignore* each others' exact states, and hold only a conjecture of the long-run average population state.

We let $\mathfrak{M}_O$ denote the set of oblivious strategies.

DEFINITION 3. Let $\mathfrak{M}_O$ be the set of feasible *oblivious strategies* available to a player, that is, Borel measurable maps from $\mathscr{X}$ to $\mathscr{A}$.

Given an oblivious strategy $\mu \in \mathfrak{M}_O$, a player i takes an action $a_{i,t} = \mu(x_{i,t})$ at time t. If the player conjectures the aggregate population state to be $f \in \mathfrak{F}$, then she also conjectures that her next state is randomly distributed according to the transition probability measure $\mathbb{P}$:

$$x_{i,t+1} \sim \mathbb{P}(\cdot \mid x_{i,t}, \mu(x_{i,t}), f), \tag{4}$$

where $f \in \mathfrak{F}$ is the conjectured long-run average population state. Note that f is a Borel probability measure over the state space; since f is modeling an infinite population, it need not have finite support (as in games with finitely many players).

We define the *oblivious value function* $V(x \mid \mu, f)$ to be the expected net present value for any player with initial state x, when the long-run average population state is conjectured to be f, and the player uses an oblivious strategy μ. We have

$$V(x \mid \mu, f) \triangleq \mathbb{E}\left[\sum_{t=0}^{\infty} \beta^t \pi(x_t, a_t, f) \,\middle|\, x_0 = x;\ \mu\right]. \tag{5}$$

Given a population state f, a player computes an optimal strategy by maximizing their oblivious value function. Note that because the oblivious value function does not track the evolution of the population state, we should expect a player's optimal strategy to depend only on their current state—that is, it must be oblivious. We capture this optimization step via the operator $\mathscr{P}$ defined next.

Define $V^*(x \mid f)$ as $V^*(x \mid f) = \sup_{\mu' \in \mathfrak{M}_O} V(x \mid \mu', f)$.

DEFINITION 4. The operator $\mathscr{P}: \mathfrak{F} \twoheadrightarrow \mathfrak{M}_O$ maps a distribution $f \in \mathfrak{F}$ to the set of optimal oblivious strategies; thus, $\mathscr{P}$ is a point-to-set correspondence. That is, $\mu \in \mathscr{P}(f)$ if and only if $V(x \mid \mu, f) = V^*(x \mid f)$, $\forall x \in \mathscr{X}$.

Note that in principle, $\mathscr{P}(f)$ may be empty, although under our assumptions standard dynamic programming results ensure that this does not occur (Maitra 1968, Dutta et al. 1994).

Now suppose that *all* players use the oblivious strategy μ, and the long-run average population state f drives their state dynamics. In this scenario, we expect the long-run population state to be an invariant distribution of the strategy μ under the dynamics

$$x_{t+1} \sim \mathbb{P}(\cdot \mid x_t, \mu(x_t), f). \tag{6}$$

We capture this relationship via the operator $\mathscr{D}$, defined next.

DEFINITION 5. The operator $\mathscr{D}: \mathfrak{M}_O \times \mathfrak{F} \twoheadrightarrow \mathfrak{F}$ maps an oblivious strategy μ and a distribution f to the set of invariant distributions associated with the dynamics (6); thus, $\mathscr{D}$ is a point-to-set correspondence.

Thus, $g \in \mathscr{D}(\mu, f)$ if and only if for all Borel sets $S \subset \mathscr{X}$,

$$g(S) = \int_{\mathscr{X}} \mathbb{P}(S \mid x, \mu(x), f) \, g(dx). \tag{7}$$

Note that the image of the operator $\mathscr{D}$ is empty if the strategy does not result in an invariant distribution, although again, we show under our assumptions that this does not occur.

We can now define mean field equilibrium. If every agent conjectures that f is the long-run average population state, then every agent would prefer to play an optimal oblivious strategy μ. On the other hand, if every agent plays μ, and the long-run average population state is indeed f, then f must also be an invariant distribution of (6). Thus, mean field equilibrium requires a consistency condition: the invariant distribution under μ and f should be exactly f.

DEFINITION 6 (MEAN FIELD EQUILIBRIUM). A strategy μ and a distribution f constitute a mean field equilibrium (MFE) if $\mu \in \mathscr{P}(f)$ and $f \in \mathscr{D}(\mu, f)$.

2.4. MFE as an Approximation to MPE

In the previous section we introduced MFE through a formal limit model. A natural question that arises in the context of MFE is whether it is a good approximation to a game with finitely many players. Here we formalize the approximation property of interest: informally, we show that an MFE is (in a sense made precise) approximately an MPE as the number of players grows large. This provides a formal justification for the notion of mean field equilibrium by considering a limiting regime where the number of players grows large.

The following proposition establishes the approximation result. Our result closely follows similar results by Weintraub et al. (2008) and Adlakha et al. (2013) (where state spaces are discrete and infinite rather than continuous and compact), but we also provide a self-contained proof sketch in the e-companion.

PROPOSITION 1. *Let (μ, f) be an MFE. Then, for all states x and sequences of cognizant strategies $\nu^{(m)} \in \mathfrak{M}$, we have*:

$$\limsup_{m \to \infty} V^{(m)}(x, f^{(m)} \mid \nu^{(m)}, \boldsymbol{\mu}^{(m-1)})$$
$$- V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)}) \leqslant 0, \tag{8}$$

almost surely, where the initial state of the given player is x, and the population state $f^{(m)}$ is derived by sampling states of the other $m - 1$ players independently from the probability mass function f.

Note that $V^{(m)}(x, f^{(m)} \mid \nu^{(m)}, \boldsymbol{\mu}^{(m-1)})$ is the *actual* value function of a player as defined in Equation (3), when the player uses a cognizant strategy $\nu^{(m)}$ and every other player plays an oblivious strategy μ. Similarly, $V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)})$ is the actual value function of a player as defined in Equation (3), when every player is playing the oblivious strategy μ.

To understand this approximation notion, suppose that we consider a sequence of games with $m \to \infty$, where all players other than player i use the oblivious strategy μ; and where the initial state of all players other than player i is sampled i.i.d. from f. Then, the proposition states that as $m \to \infty$, the difference between the payoff player i achieves by playing μ and the maximum possible payoff player i can achieve by playing any cognizant strategy approaches zero almost surely, for all initial states x of player i. Thus, in particular, μ is approximately optimal for player i in a large finite game. A weaker version of this property was introduced by Weintraub et al. (2008); a similar notion is also studied by Glynn (2004), Huang et al. (2005), Tembine et al. (2009), and Bodoh-Creed (2012).

Based on the preceding proposition, MFE is a reasonable approximation to MPE with a finite large number of players. The remainder of the paper is devoted to studying the properties of MFE. In particular, we focus on existence of, and convergence to, MFE.

We conclude with a notational note: because in the remainder of the paper, we focus only on MFE of stochastic games, we drop the number of players m from the tuple defining a stochastic game Γ.

2.5. Lattice-Theoretic Preliminaries

This section contains an overview of some basic definitions and notation used in the remainder of the paper. Our development requires some basic concepts from the theory of lattices. Given a partially ordered set X with order $\succeq$, an element x is called an *upper bound* of S if $x \succeq y$ for all $y \in S$; similarly, x is called a lower bound of S if $y \succeq x$ for all $y \in S$. We say that x is a *supremum* or *least upper bound* of S in X if x is an upper bound of S, and for any other upper bound x' of S, we have $x' \succeq x$. In this case, we write $x = \sup S$. We similarly define infimum (or greatest lower bound) and denote it by $\inf S$. The partially ordered

set $(X, \succeq)$ is a *lattice* if for all pairs $x, y \in X$, the elements $\sup\{x, y\}$ and $\inf\{x, y\}$ exist in X. The lattice $(X, \succeq)$ is a *complete lattice* if, in addition, for all nonempty subsets $S \subset X$, the elements $\sup S$ and $\inf S$ exist in X.

If X is a lattice, a function $f: \mathcal{X} \to \mathbb{R}$ is *supermodular* if $f(\sup\{x, x'\}) + f(\inf\{x, x'\}) \geqslant f(x) + f(x')$ for every pair x, x'. If Y is also a lattice, a function $f: X \times Y \to \mathbb{R}$ has *increasing differences* in x and y if for all $x' \succeq x$, $y' \succeq y$, there holds $f(x', y') - f(x', y) \geqslant f(x, y') - f(x, y)$. Finally, a correspondence $T: X \to Y$ is *nondecreasing* if whenever $x' \succeq x$, $y \in T(x)$, and $y' \in T(x')$, there holds $\sup\{y, y'\} \in T(x')$, and $\inf\{y, y'\} \in T(x)$. (For more detail on lattice programming, the reader is referred to Topkis 1998.)

Throughout this paper, we view $\mathcal{X}$ and $\mathcal{A}$ as lattices in the usual ordering; since these spaces are both compact subsets of $\mathbb{R}$, the corresponding lattices are complete (Topkis 1998). We also view the set of strategies $\mathfrak{M}_O$ as a lattice, under the coordinate ordering $\trianglerighteq$: that is, $\mu' \trianglerighteq \mu$ if and only if $\mu'(x) \geqslant \mu(x)$ for all x.

In addition, recall that we let $\mathfrak{F}$ denote the set of all Borel probability measures on $\mathcal{X}$. We view $\mathfrak{F}$ as a lattice with the (*first-order*) *stochastic dominance* ordering; formally, we write $f' \succeq_{SD} f$ if and only if:

$$\int_{\mathcal{X}} g(x) f'(dx) \geqslant \int_{\mathcal{X}} g(x) f(dx)$$

for all nondecreasing, bounded, measurable functions g on $\mathcal{X}$ (where the integral is the Riemann-Stieltjes integral). It is straightforward to show that this condition is equivalent to $F(x) \leqslant F'(x)$, where F (respectively, F') is the cumulative distribution function of f (respectively, f'). It is well known that $\mathfrak{F}$ is a lattice: the lattice supremum $\sup_{SD}\{f, f'\}$ (respectively, the lattice infimum $\inf_{SD}\{f, f'\}$) is found by the pointwise infimum (respectively, supremum) of the corresponding distribution functions. Because $\mathcal{X}$ is compact, it is straightforward using an analogous argument to verify that $\mathfrak{F}$ is a complete lattice (Echenique 2003).

We conclude by defining some properties of parameterized distributions we require in the sequel. Let $f(\cdot \mid y)$ denote a family of measures in $\mathfrak{F}$, parameterized by $y \in Y$, where Y is a lattice. Then we say f is *stochastically nondecreasing* in y if whenever y' is larger than y, $f(\cdot \mid y') \succeq_{SD} f(\cdot \mid y)$. Similarly, let $f(\cdot \mid y, z)$ denote a family of measures in $\mathfrak{F}$ parameterized by $y \in Y$ and $z \in Z$, where both Y and Z are lattices. Then we say that f has *stochastically increasing differences* in y and z if the expectation $\int_{x \in \mathcal{X}} g(x) f(dx \mid y, z)$ has increasing differences in y and z, for every nondecreasing, bounded, measurable function g on $\mathcal{X}$.

3. Existence of MFE

In this section and the following section, we consider a baseline model of stochastic games with complementarities, in which we prove existence and convergence results. In this section, we establish our first main result: that there

exists an MFE for the stochastic game with complementarities. We also show an ordering result: there exists a "largest" and a "smallest" equilibrium among the set of all MFE with nondecreasing strategies.

We have the following definition.

DEFINITION 7. A *stochastic game with complementarities* is a stochastic game $\Gamma = (\mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$ that satisfies the following properties.

1. *Nondecreasing and supermodular payoff.* The payoff $\pi(x, a, f)$ is nondecreasing in x, and supermodular in (x, a).

2. *Payoff complementarity.* The payoff function $\pi(x, a, f)$ has increasing differences in (x, a) and f.

3. *Monotone and supermodular transition kernel.* The transition kernel $\mathbb{P}(\cdot \mid x, a, f)$ is stochastically supermodular in (x, a) and is stochastically nondecreasing in each of x, a, and f.

4. *Transition kernel complementarity.* The transition kernel $\mathbb{P}(\cdot \mid x, a, f)$ has stochastically increasing differences in (x, a) and f.

5. *Monotone action set.* The correspondence $\mathcal{A}(x)$ is nondecreasing in x. Further, $\sup_{a \in \mathcal{A}(x)} \pi(x, a, f)$ is nondecreasing in x for all fixed f.

6. *Continuity.* Assumption 1 holds.

The first assumption is natural for a range of models—if larger states are more valuable, then the payoff function will be nondecreasing in the state. The boundedness assumption on the payoff will be trivially satisfied if, for example, $\mathcal{X}$ is an interval and the payoff is continuous in x. The second assumption ensures that there are complementarities between the state and action of a single player and the population state of other players. The next three assumptions create complementarities between state and action, as well as ensure that larger states and/or larger actions now are more likely to lead to larger states in the future. The last property simply reiterates the continuity properties stated in Assumption 1; we use these properties to simplify dynamic programming arguments.

Whereas it may be straightforward to verify whether a payoff function exhibits the desired complementarity properties, the same verification is somewhat more challenging for the transition kernel. Thus, before continuing, we provide examples of transition kernels that exhibit the complementarity conditions required in Definition 7.

EXAMPLE 4 (ADDITIVELY SEPARABLE DYNAMICS). Suppose that $\mathbb{P}(\cdot \mid x, a, f)$ is defined as follows:

$$\mathbb{P}(\cdot \mid x, a, f) = cQ(\cdot \mid x, a) + (1 - c)R(\cdot \mid f). \quad (9)$$

Here c is constant such that $0 \leqslant c \leqslant 1$, and Q and R are both distributions. If $Q(\cdot \mid x, a)$ is stochastically nondecreasing in x and a, stochastically supermodular in (x, a), and if $R(\cdot \mid f)$ is stochastically nondecreasing in f, then it can be checked that the transition kernel of (9) satisfies all the conditions of Definition 7. This kernel is inspired by additively separable transition models in classical stochastic games; see, for example, Filar and Vrieze (1996). □

EXAMPLE 5 (MIXTURE DYNAMICS). Suppose that $\mathbb{P}(\cdot \mid x, a, f)$ is defined as follows:

$$\mathbb{P}(\cdot \mid x, a, f) = q(x, a, f)F(\cdot) + (1 - q(x, a, f))G(\cdot). \quad (10)$$

Here F and G are both distributions on $\mathcal{X}$, such that F first order stochastically dominates G, and $0 \leqslant q(x, a, f) \leqslant 1$. If $q(x, a, f)$ is nondecreasing in x, a, and f, supermodular in (x, a), and has increasing differences in (x, a) and f, then it can be checked that the expectation of (10) against any nondecreasing function satisfies all the conditions of Definition 7. As one example of a q that satisfies these properties, suppose:

$$q(x, a, f) = \frac{x + a + \eta(f)}{2 \sup \mathcal{X} + \sup \mathsf{A}},$$

where $\eta(f) = \int_{\mathcal{X}} x' f(dx')$ is the mean of f. Such dynamics are commonly used in the context of games with strategic complementarities (Curtat 1996, Nowak 2007). □

Informally, how might we expect players to behave in such a game? Observe that if other players have a larger population state, this increases the return to a larger state for a given player. To achieve a larger state, a player must take a larger action, but this also increases the likelihood of larger states in the future. All these effects conspire to create a situation where, when players are confronted with larger population states, they are likely to take higher actions. This monotonicity drives our analysis.

For the remainder of the section, we fix a stochastic game with complementarities $\Gamma = (\mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$. Let $\Phi: \mathfrak{F} \to \mathfrak{F}$ denote the composition of $\mathscr{P}$ and $\mathscr{D}$ for the game Γ: $\Phi(f) = \mathscr{D}(\mathscr{P}(f), f)$. A fixed point of Φ identifies an MFE of Γ. Intuitively, under the assumptions we have made, we might expect Φ to be a *monotone* map; that is, larger initial conjectures about the population state should lead players to take higher actions, which should in turn lead to a larger invariant distribution. Tarski's fixed-point theorem ensures that monotone functions on a lattice have a fixed point.[2]

THEOREM 1 (TARSKI 1955). *Suppose that $\mathcal{L}$ is a nonempty complete lattice, and $T: \mathcal{L} \to \mathcal{L}$ is a nondecreasing function. Then the set of fixed points of T is a nonempty complete lattice.*

We proceed to show that Φ is monotone by showing that each of two correspondences $\mathscr{P}$ and $\mathscr{D}$ are monotone (with respect to the coordinate ordering on strategies in $\mathfrak{M}_O$, and the first-order stochastic dominance ordering on $\mathfrak{F}$).

Our main result in this section is the following theorem.

THEOREM 2. *There exists an MFE for the stochastic game with complementarities Γ.*

In the next section, we sketch a proof of this theorem; and in §3.2, we show that if we restrict attention to equilibria where the strategy is nondecreasing, then there exists a "largest" equilibrium and a "smallest" equilibrium.

3.1. Theorem 2: Proof Sketch

We sketch the proof of Theorem 2; each step is filled in by the lemmas in the appendix.

Step 1. We show $\mathscr{P}(f)$ is nonempty, and that optimal strategies can be identified via Bellman's equation (Lemma 2).

Step 2. We show that the value function $V^*(x \mid f)$ is nondecreasing in x and has increasing differences in x and f. We use this fact to show that

$$\pi(x, a, f) + \beta \int_{\mathscr{X}} V^*(x' \mid f)\mathbb{P}(dx' \mid x, a, f)$$

is supermodular in (x, a) and has increasing differences in (x, a) and f (Lemmas 3, 4, and 5).

Step 3. We use the complementarity properties of the previous step to show that the strategies $\bar{p}(f)$ and $\underline{p}(f)$ are nondecreasing in the state x, where

$$\bar{p}(f) = \sup \mathscr{P}(f); \quad \text{and} \quad \underline{p}(f) = \inf \mathscr{P}(f). \tag{11}$$

We also show that $\bar{p}$ and $\underline{p}$ are nondecreasing in f. (These facts are shown in Lemma 6.)[3]

Step 4. We show that when restricted to strategies μ that are nondecreasing in state, $\bar{d}(\mu, f)$ and $\underline{d}(\mu, f)$ are nondecreasing in μ and f, where

$$\bar{d}(\mu, f) = \sup \mathscr{D}(\mu, f) \quad \text{and}$$
$$\underline{d}(\mu, f) = \inf \mathscr{D}(\mu, f). \tag{12}$$

(This is shown in Lemmas 7 and 8.)

Step 5. We conclude that the functions $\bar{\Phi}(f)$ and $\underline{\Phi}(f)$ are nondecreasing in f, where

$$\bar{\Phi}(f) = \bar{d}(\bar{p}(f), f); \quad \underline{\Phi}(f) = \underline{d}(\underline{p}(f), f). \tag{13}$$

Thus, both $\bar{\Phi}(f)$ and $\underline{\Phi}(f)$ possess fixed points by Tarski's theorem (Lemma 9). These fixed points identify MFE.

3.2. Largest and Smallest Equilibria

Typically, in games with supermodular structure, it is possible to show various ordering relationships among the equilibria. In particular, there is typically a "largest" and "smallest" equilibrium (Milgrom and Roberts 1990). In our setting, we might conjecture that the largest fixed point of $\bar{\Phi}$ (respectively, the smallest fixed point of $\underline{\Phi}$) is the largest (respectively, the smallest) MFE of the stochastic game Γ. However, this need not be the case: as seen above, monotonicity properties of the map $\mathscr{D}$ are only inferred on the subset of strategies that are *nondecreasing* in the state. In general, such monotonicity properties might not hold over the entire strategy set—that is, $\underline{d}(\mu, g)$ and $\bar{d}(\mu, g)$ may not be nondecreasing over the entire set $\mathfrak{M}_O$. These monotonicity properties are necessary for establishing the ordering of equilibria in classical supermodular game theory.

From the discussion in the preceding paragraph, however, observe that if we restrict attention to nondecreasing strategies, then indeed an ordering result can be proven. In particular, the following corollary shows that any MFE where the strategy is nondecreasing is bounded above by the largest fixed point of $\bar{\Phi}$, and bounded below by the smallest fixed point of $\underline{\Phi}$. (See e-companion for the proof.)

COROLLARY 1. *Let $\bar{f}$ be the largest fixed point of $\bar{\Phi}$, and let $\underline{f}$ be the smallest fixed point of $\underline{\Phi}$, that is,*

$$\bar{f} = \sup\{f : \bar{\Phi}(f) = f\}; \quad \underline{f} = \inf\{f : \underline{\Phi}(f) = f\}. \tag{14}$$

Let (μ, f) be any MFE of the stochastic game with complementarities Γ, where μ is nondecreasing. Then $\underline{f} \preceq_{\mathrm{SD}} f \preceq_{\mathrm{SD}} \bar{f}$, and thus $\underline{p}(\underline{f}) \trianglelefteq \mu \trianglelefteq \bar{p}(\bar{f})$.

4. Convergence to Equilibrium

In this section, we show that an MFE can be obtained using a natural form of learning dynamics among the players. We start by considering a simple form of best-response dynamics to compute equilibria, where we iteratively apply the maps $\bar{\Phi}$ and $\underline{\Phi}$ defined in (13). We argue that this process is unsatisfactory, both from a computational and modeling standpoint, and instead propose an alternate process we refer to as *myopic learning dynamics*; these dynamics are both computationally simpler and correspond to a natural learning behavior among the agents. We show that this process converges to MFE.

We fix a stochastic game with complementarities $\Gamma = (\mathscr{X}, \mathscr{A}, \mathbf{P}, \pi, \beta)$. Throughout this section, we study Γ in the limit of a continuum of agents, consistent with our definition of MFE.

4.1. Best-Response Dynamics

We start by considering the following algorithm.

ALGORITHM *L*-BRD:

1. Initialize the state of every agent to $\underline{x} = \inf \mathscr{X}$, and let f_0 denote the resulting population state—that is, f_0 places all its mass on $\underline{x}$.

2. At time t, let $\mu_{t+1} = \underline{p}(f_t)$, and let $f_{t+1} = \underline{d}(\mu_t, f_t)$, cf. (11) and (12).

3. Repeat (2).

Here *L*-BRD denotes *lower best-response dynamics*. Given a current population state, we compute the lowest best response of a player, and then compute the smallest invariant distribution corresponding to the resulting strategy. This is the simplest dynamic we might consider; since $\underline{\Phi}(f) = \underline{d}(\underline{p}(f), f)$, we have $f_{t+1} = \underline{\Phi}(f_t)$. In spirit, this algorithm is similar to other best-response dynamics that are common in the literature on supermodular games (Milgrom and Roberts 1990, Vives 1990).

We now show that this algorithm converges, and further, the limit point is the smallest mean field equilibrium.

The next proposition shows *L*-BRD converges; the proof follows by exploiting monotonicity of $\underline{\Phi}$.

PROPOSITION 2. *Let Γ be a stochastic game with complementarities. Define f_t and μ_t iteratively according to Algorithm L-BRD. Then, $f_0 \preceq_{SD} f_1 \preceq_{SD} f_2 \cdots$, and $\mu_0 \unlhd \mu_1 \unlhd \mu_2 \cdots$. Further, there exists a distribution f^* and a strategy μ^*, nondecreasing in x, such that f_t converges weakly to f^* as $t \to \infty$, and μ_t converges pointwise to μ^* as $t \to \infty$.*

Further, (μ^*, f^*) is an MFE, and f^* is $\underline{f}$, the smallest fixed point of $\underline{\Phi}$ (cf. (14)).

Thus, best-response dynamics converge to an MFE. Further, the limit point is the smallest MFE among all those where the equilibrium strategy is nondecreasing.

We conclude by noting that we can analogously define an *upper* best-response dynamic algorithm (*U*-BRD) where each agent's initial state is set to $\bar{x} = \sup \mathcal{X}$. The next step population state is given by the dynamics: $f_{t+1} = \bar{\Phi}(f_t) = \bar{d}(\bar{p}(f_t), f_t)$. The same conclusion as Proposition 2 holds for *U*-BRD as well, except that the limit point is the largest fixed point of $\bar{\Phi}$; that is, $f^* = \bar{f}$ (cf. 14).

We note that one alternative to *L*-BRD and *U*-BRD is presented by Sleet (2001). He suggests an algorithm based on iterative value and policy iteration to compute an MFE of a dynamic price-setting game with stochastic, exogenous firm-specific demand shocks per period. The setting considered there is specialized, but the convergence proof also exploits monotonicity properties induced by complementarity conditions in that specific model.

4.2. Myopic Learning Dynamics

The preceding section establishes the desirable result that best-response dynamics converge. However, in a dynamic context, iterative application of $\bar{\Phi}$ and $\underline{\Phi}$ is not completely satisfactory, whether viewed from a computational or modeling standpoint. First, given f, computing $\bar{\Phi}(f)$ or $\underline{\Phi}(f)$ requires computing the invariant distribution of the Markov chain induced by $\bar{p}(f)$ or $\underline{p}(f)$, introducing additional complexity. Second, the process of iteratively applying $\bar{\Phi}$ or $\underline{\Phi}$ does not naturally correspond to any reasonable dynamic process that agents are likely to follow in practice: it is difficult to imagine an agent first computing the invariant distribution of the current strategy in use by her competitors, and then solving a dynamic program given that invariant distribution.

By contrast, in this section we present a pair of *myopic learning dynamics* that address these considerations. The algorithms presented in this section are simple and easy to implement. Furthermore, they demand only a weak form of rationality from the players, thereby resolving the two main issues of *computability* and *plausibility* associated with the standard solution concept of MPE (as discussed in the introduction).

In the myopic learning dynamic, at each time t, each agent computes a best response to the current population state distribution f_t, *assuming that the population state will remain at f_t at all future times*. (This step is similar to

model predictive control or receding horizon control; see, for example, Garcia et al. 1989.) In other words, agents play according to a strategy in $\mathcal{P}(f_t)$. This play yields a new population state f_{t+1} at the next time step according to the transition kernel.

The algorithms we consider are reasonable in settings where agents are not likely to predict future learning by other agents. Indeed, such an assumption seems plausible precisely in the large systems that mean field equilibrium is meant to model. In such systems, myopic behavior is simple computationally; by contrast, solving a dynamic program with full knowledge of future strategies that other agents will employ places unreasonable informational requirements on the agents.

We first consider an algorithm where agents play actions induced by $\underline{p}$.

ALGORITHM *L*-MLD:

1. Every agent initializes their state to $\underline{x} = \inf \mathcal{X}$ at time $t = 0$.

2. Agents observe the population state f_t.

3. An agent with state x chooses the action a_t so that $a_t = \mu_t(x)$, where $\mu_t(x) = \underline{p}(f_t)(x)$. The agent's next state is distributed according to $\overline{\mathbb{P}}(\cdot \mid x, a_t, f_t)$.

4. Repeat (2)–(3).

Here *L*-MLD denotes *lower myopic learning dynamics*. Observe that agents compute a new strategy based on the observed *current* population state—not based on the invariant distribution associated with the last strategy chosen. This means that two simultaneous dynamic processes are taking place: strategy revision on the part of the players, but also state update via the system dynamics (4). Due to this intertwined dynamic, novel arguments are required to prove convergence of best-response dynamics (relative to usual proofs of convergence for such dynamics in supermodular games, for example, Milgrom and Roberts 1990, Vives 1990). We also note that although the same strategy is computed by every agent, the particular action chosen will vary depending on their current state.

The preceding description yields a simple recursion for the population state at the next time step; for all Borel sets S:

$$f_{t+1}(S) = \int_{\mathcal{X}} \mathbb{P}(S \mid x', \mu_t(x'), f_t) f_t(dx')$$
$$= Q_{\mu_t, f_t}(f_t)(S), \tag{15}$$

where $Q_{\mu, f}(f)$ is defined as follows:

$$Q_{\mu, f}(g)(S) = \int_{\mathcal{X}} \mathbb{P}(S \mid x, \mu(x), f) g(dx). \tag{16}$$

Our goal is to understand the behavior of the sequence of population states $f_0, f_1, f_2, \ldots$, as well as the sequence of policies $\mu_0, \mu_1, \mu_2, \ldots$. We have the following proposition, which mirrors Proposition 2.

PROPOSITION 3. *Let Γ be a stochastic game with complementarities. Define f_t and μ_t iteratively, according to Algorithm L-MLD. Then, $f_0 \preceq_{SD} f_1 \preceq_{SD} f_2 \cdots$, and $\mu_0 \trianglelefteq \mu_1 \trianglelefteq \mu_2 \cdots$. Further, there exists a distribution f^* and a strategy μ^*, nondecreasing in x, such that f_t converges weakly to f^* as $t \to \infty$, and μ_t converges pointwise to μ^* as $t \to \infty$.*

Further, (μ^*, f^*) is an MFE, and f^* is $\underline{f}$, the smallest fixed point of $\underline{\Phi}$ (cf. (14)).

Thus, we find the same result as for *L*-BRD: the dynamics converge to the smallest MFE among all those where the equilibrium strategy is nondecreasing.

The proof of Proposition 3 proceeds as follows. We exploit two key monotonicity properties established in the course of proving existence of an equilibrium (Theorem 2): first, that $\underline{p}(f)$ is monotone in f (Lemma 6 in the appendix); and second, that $Q_{\mu,f}(g)(S)$ is monotone in μ, f, and g (Lemma 7 in the appendix). These two properties together allow us to establish that μ_t and f_t form monotone sequences—even though players are reacting only to the current population state, the population state over time moves monotonically towards an equilibrium.

Note that *L*-MLD initializes players to the lowest state, inf $\mathcal{X}$. This behavior of *L*-MLD is particularly meaningful for several of the applications described in the introduction; for example, in an interdependent security setting, we might envision a scenario where a new, more efficient technology for security is introduced. In this case, the "low" initial population state might correspond to the status quo, and then the myopic learning dynamics track the adaptation of the population to a new equilibrium configuration.

A similar convergence result also holds if instead every agent starts at the *largest* state $\bar{x} = \sup \mathcal{X}$, and follows the strategy $\bar{p}(f_t)$ at each time step. We call this Algorithm *U*-MLD. The same conclusion as Proposition 3 holds for *U*-MLD as well, except that the limit point is the *largest* fixed point of $\bar{\Phi}$, that is, $f^* = \bar{f}$ (cf. (14)).

We emphasize that the interpretation of these dynamics that we provide in this section is only valid in the formal mean field limit model. When there are only *finitely* many agents, of course, the population state will not exactly follow the dynamics described here. We conclude this section by discussing the behavior of myopic learning dynamics in finite systems. In particular, suppose that in a game consisting of *m* players, each player follows the dynamic prescribed by *L*-MLD. Each player starts in the lowest state, and then at each time step, observes the current population state and plays one step according to the optimal oblivious strategy given that population state. Because the system is finite, additional error is introduced due to the randomness in state transitions of individual agents; in particular, due to this randomness, it is not immediately guaranteed that myopic learning dynamics will converge to an MFE in a finite game. However, if the state space is discrete, then using techniques similar to Proposition 1, it can be shown

that $f_t^{(m)} \to f_t$ weakly, almost surely, where $f_t^{(m)}$ is the population state after *t* time steps with *m* players, and f_t is the population state in the *L*-MLD dynamic after *t* time steps in the mean field limit. Thus (in this precise sense), *L*-MLD behavior in large finite systems is "close" to the mean field model.

5. Comparative Statics

In this section we discuss *sensitivity* analysis of equilibria, also known as *comparative statics* results. Our goal is to understand how the equilibrium distribution and optimal strategy are altered in response to changes in parameters. These results allow us to evaluate changes in equilibrium with respect to changes in a parameter.

In this section we consider a family of stochastic games with complementarities, parameterized by $\theta \in \Theta$, where Θ is a complete lattice. In the context of security games, this parameter could, for example, represent the effectiveness of a particular security technology. Alternatively, in the context of recommendation systems, Θ might represent the effectiveness of the collaborative filtering engine in improving recommendations to one agent based on the profiles of other agents.

Formally, suppose we are given a family of stochastic games $\Gamma(\theta)$ for $\theta \in \Theta$ with common strategy spaces, action spaces, and discount factors, where for each θ, $\Gamma(\theta)$ is a stochastic game with complementarities, that is, $\Gamma(\theta)$ satisfies Definition 7 for each $\theta \in \Theta$. We refer to Γ as a *parametric family* of stochastic games with complementarities. Let $\pi(x, a, f; \theta)$ and $\mathbf{P}(\cdot \mid x, a, f; \theta)$ be the payoff and transition kernel, respectively, in $\Gamma(\theta)$. We make the following assumption.

ASSUMPTION 2. *The payoff $\pi(x, a, f; \theta)$ has increasing differences in (x, a, f) and θ. The transition kernel $\mathbf{P}(\cdot \mid x, a, f; \theta)$ has stochastically increasing differences in (x, a, f) and θ, and is stochastically nondecreasing in θ for fixed x, a, f.*

Under the preceding assumption, we can give a directional characterization of the movement of equilibrium in response to parameter changes; the proof can be found in the e-companion.

THEOREM 3. *Let Γ be a parametric family of stochastic games with complementarities, and suppose that Assumption 2 holds. Let $\underline{f}(\theta)$ and $\bar{f}(\theta)$ denote the "smallest" and "largest" MFE in the game $\Gamma(\theta)$, cf. (14). Then $\bar{f}(\theta)$ and $\underline{f}(\theta)$ are both nondecreasing in θ.*

Such comparative statics results are commonly applied in the context of games with complementarities, but it is worth noting that in a dynamic context this result provides additional insight, because it quantifies how the *distribution* of agents' states will respond as a parameter changes. This kind of insight is particularly valuable for system designers, regulators, and policy makers, where changes in equilibrium behavior due to control decisions may be challenging

to characterize. As one simple consequence of the preceding theorem, suppose that in security games an incentive is introduced for agents to invest in security as a linear rebate in the payoff, proportional to an agent's security level x. It is straightforward to check that this results in more players opting for higher investment, and thus the equilibrium population state tends to shift towards higher security levels.

6. Separable Stochastic Games

As the preceding sections illustrate, stochastic games with complementarities possess a number of properties that make them amenable to equilibrium analysis. One potential concern, however, is that the set of models admitted by Definition 7 may be somewhat limiting. Consider the following example.

EXAMPLE 6 (LINEAR DYNAMICS). Consider a simple model where the distribution of the next state of an agent is "linear" in x and a. Let W be a zero mean random variable that takes countably many values, and fix positive constants A and B. We consider a state space $\mathcal{X} = [-\underline{M}, \bar{M}]$, for some large positive constants $\underline{M}, \bar{M}$, and let $\mathcal{A}(x) = [\underline{a}, \bar{a}]$ for all x, where $\underline{a} \leqslant \bar{a}$. Define $\mathbb{P}$ as follows:

$$\mathbb{P}(x' \,|\, x, a)$$
$$= \begin{cases} \text{Prob}(Ax + Ba + W \geqslant \bar{M}), & x' = \bar{M}; \\ \text{Prob}(Ax + Ba + W = x'), & -\underline{M} < x' < \bar{M}; \quad (17) \\ \text{Prob}(Ax + Ba + W \leqslant -\underline{M}), & x' = -\underline{M}. \end{cases}$$

In this model, the state dynamics are essentially linear, except at the boundaries of the state space (where the state is truncated to lie within $[-\underline{M}, \bar{M}]$). Such a model might naturally arise in a wide range of examples, such as Examples 1 or 2 (see §7 for details).

Unfortunately, such a kernel *does not* exhibit stochastically increasing differences in general. To see this, we consider a simple instance where $\text{Prob}(W = 1) = \text{Prob}(W = -1) = 1/2$, and $\underline{M} = \bar{M} = M \gg 1$. Consider any nondecreasing function $\phi(x)$, and fix x and a such that $|Ax + Ba| < M - 1$. Then,

$$\mathbb{E}[\phi(x') \,|\, x, a] = \tfrac{1}{2}\phi(Ax + Ba + 1) + \tfrac{1}{2}\phi(Ax + Ba - 1).$$

In general, the right-hand side exhibits increasing differences in x and a only if ϕ is locally *convex*. This is easiest to see for differentiable ϕ; in that case, the cross-partial derivative $\partial^2 \phi(Ax + Ba + 1)/\partial x \partial a$ has to be nonnegative to ensure increasing differences, which only holds if $\phi''(Ax + Ba + 1) \geqslant 0$. For general nondecreasing ϕ, therefore, the expectation $\mathbb{E}[\phi(x') \,|\, x, a]$ need not exhibit increasing differences in x and a. $\square$

The preceding example highlights a deficiency in stochastic games with complementarities: while a rich class analytically, they do present some restrictions from a modeling standpoint. In this sense, complementarity can appear to be a *brittle* property.

However, this same brittleness can actually become an advantage. Although at first glance it may appear that complementarity fails, often simple transformations can lead to games that admit analysis via complementarity methods even if the original game did not. (A common example is the class of *log-supermodular* games used extensively in oligopoly theory, where the logarithm of the profit function may be supermodular; see, for example, Milgrom and Roberts 1990 and Vives 1990 for details.)

In this section we demonstrate that a wide range of models, including those with dynamics similar to Example 6, can be transformed to standard stochastic games with complementarities. Further, the class of models we develop has the benefit that the assumptions are typically easier to check in practice. This significantly widens the applicability of our theory to models where the desired monotonicity properties may not be immediately apparent.

The class of games we consider in this section feature a payoff that is *separable* in the state and action. We have the following definition.

DEFINITION 8. A *separable stochastic game* is a stochastic game $\Gamma = (\mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$ with the following additional properties.

1. *Actions.* There exist $\underline{a}, \bar{a}$, such that $\mathcal{A}(x) = [\underline{a}, \bar{a}]$ for all x.

2. *Payoff.* The single-period payoff to player i at time t can be written as $\pi(x_{i,t}, a_{i,t}, f_{-i,t}) = v(x_{i,t}, f_{-i,t}) - c(a_{i,t})$, where we refer to $v(x, f)$ as the *utility* at state x and population state f, and $c(a)$ as the *cost* for action a.

3. *Transition probabilities.* The state of a player evolves according to a Markov process with the following transition probabilities. If the state of player i at time t is $x_{i,t} = x$ and the player takes an action $a_{i,t} = a$ at time t, then the next state is distributed according to the Borel probability measure $\mathbb{P}(\cdot \,|\, h(x, a), f)$, where for Borel sets $S \subset \mathcal{X}$,

$$\mathbb{P}(S \,|\, h(x, a), f)$$
$$= \text{Prob}(x_{i,t+1} \in S \,|\, x_{i,t} = x, a_{i,t} = a, \mathbf{f}_{-i,t} = f). \quad (18)$$

Note that $\mathbb{P}$ depends on x and a only through the function $h(x, a)$; we refer to $h(x, a)$ as the *kernel parameter*. We assume that h takes values in a compact interval $H = [\underline{h}, \bar{h}] \subset \mathbb{R}$.

In this section, we provide insight into *separable stochastic games with complementarities*. We have the following definition.

DEFINITION 9. A *separable stochastic game with complementarities* is a separable stochastic game $\Gamma = (\mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$ with the following properties.

1. *Nondecreasing payoff and convex cost.* The utility function $v(x, f)$ is nondecreasing in x, and the cost function $c(a)$ is nondecreasing and convex in a.

2. *Payoff complementarity.* The utility function $v(x, f)$ has increasing differences in x and f.

3. *Monotone transition kernel.* The transition kernel $\mathbb{P}(\cdot \mid \hat{h}, f)$ is stochastically nondecreasing in $\hat{h}$ and f.

4. *Transition kernel complementarity.* The transition kernel $\mathbb{P}(\cdot \mid \hat{h}, f)$ has stochastically increasing differences in $\hat{h}$ and f.

5. *Kernel parameter monotonicity and complementarity.* The kernel function $h(x, a)$ is supermodular in x and a, nondecreasing in the state x, and concave and nondecreasing in the action a.

6. *Continuity.* The primitives $v(x, f)$, $\mathbb{P}(\cdot \mid \hat{h}, f)$, and $h(x, a)$ are continuous on their respective domains.

We proceed by reparameterizing the strategy in terms of the kernel parameter; under this reparametrization, the resulting model is revealed to be a special case of the general model studied earlier in this paper.

Formally, suppose we are given a separable stochastic game with complementarities $\Gamma = (\mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$. Before we proceed, we require some additional notation. For each x, define

$$H(x) = \{\hat{h}: h(x, a) = \hat{h} \text{ for some } a \in \mathcal{A}\}. \qquad (19)$$

Thus $H(x)$ is the image of $\mathcal{A}$ under $h(x, \cdot)$. In addition, for each $\hat{h} \in H(x)$, define

$$C(x, \hat{h}) = \inf_{a \in A: \ h(x, a) = \hat{h}} c(a). \qquad (20)$$

Thus, $C(x, \hat{h})$ is the minimum cost incurred to achieve kernel parameter $\hat{h}$ when at state x.

The next lemma establishes some basic properties of H and C. It uses the assumption that the cost function is a convex function of action a.

LEMMA 1. *Suppose Γ is a separable stochastic game with complementarities. Suppose $H(x)$ and $C(x, \hat{h})$ are defined as in* (19) *and* (20)*, respectively. Then, for each x, $H(x)$ is a compact interval, and the sets $H(x)$ are nondecreasing in x.*

The function $C(x, \hat{h})$ is convex and nondecreasing in $\hat{h}$ on $H(x)$ for each x, and nonincreasing in x for each $\hat{h}$ as long as $\hat{h} \in H(x)$. Further, for all x:

$$\inf_{\hat{h} \in H(x)} C(x, \hat{h}) = c(\underline{a}). \qquad (21)$$

If $x' > x$, $\hat{h}', \hat{h} \in H(x') \cap H(x)$, and $\hat{h}' > \hat{h}$, then

$$C(x', \hat{h}') - C(x, \hat{h}') \leqslant C(x', \hat{h}) - C(x, \hat{h}).$$

In other words, $C(x, \hat{h})$ has decreasing differences in x and $\hat{h}$.

We now use Lemma 1 to define a *new* stochastic game, which is in fact a stochastic game with complementarities as in Definition 7.

PROPOSITION 4. *Suppose that $\Gamma = (\mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$ is a separable stochastic game with complementarities. Define a new game $\hat{\Gamma} = (\hat{\mathcal{X}}, \hat{\mathcal{A}}, \hat{\mathbb{P}}, \hat{\pi}, \beta)$, where:*
 1. $\hat{\mathcal{X}} = \mathcal{X}$;
 2. $\hat{\mathcal{A}}(x) = H(x)$ *for all* $x \in \mathcal{X}$;
 3. $\hat{\mathbb{P}}(x' \mid x, \hat{h}, f) = \mathbb{P}(x' \mid \hat{h}, f)$; *and*
 4. $\hat{\pi}(x, \hat{h}, f) = v(x, f) - C(x, \hat{h})$,
with $H(x)$ and $C(x, \hat{h})$ defined in (19) *and* (20)*, respectively. Then $\hat{\Gamma}$ is a stochastic game with complementarities, cf. Definition 7.*

Based on the preceding proposition, we have the following theorem.

THEOREM 4. *Any separable stochastic game with complementarities Γ has an MFE.*

The preceding result can be extended, of course, to provide analogs of Corollary 1 (existence of a largest and smallest equilibrium), as well as Propositions 2 and 3 (convergence of best-response dynamics and myopic learning dynamics, respectively). Note, however, that the dynamics defined here are in the *modified* strategy space, where the "action" is the kernel parameter chosen. In particular, the dynamics in the original action space may not be monotone at all; nevertheless, the eventual limit point is an MFE.

It is also straightforward to generalize the comparative statics result in Theorem 3 to separable stochastic games using the same transformation as the preceding result. In addition, the definition of a separable stochastic game with complementarities can be naturally extended to separable action-coupled stochastic games with complementarities (simply by replacing the population state f by the population action distribution α in the payoff and transition kernel), and an argument similar to Proposition 4 shows that such a game can be transformed to a standard action-coupled stochastic game with complementarities.

We conclude this section by noting that the preceding results continue to hold in a setting where the payoff is not necessarily monotone, as long as dynamics are decoupled. Formally, suppose that $\Gamma = (\mathcal{X}, \mathcal{A}, \mathbb{P}, \pi, \beta)$ is a stochastic game that satisfies all the conditions in Definition 9, *except* that v is not necessarily nondecreasing in x. Suppose in addition that $\mathbb{P}(\cdot \mid \hat{h}, f)$ does not depend on f; thus, we denote the kernel simply as $\mathbb{P}(\cdot \mid \hat{h})$. In this model it can again be shown that an MFE exists, as we now describe.

The proof of Theorem 2 (and subsequent results on ordering of equilibria and convergence) use the fact that the payoff is nondecreasing in x to show that $\int_{\mathcal{X}} V^*(x' \mid f)\mathbb{P}(dx' \mid x, a, f)$ is supermodular in (x, a) and has increasing differences in (x, a) and f (see Lemma 3). For the expectation to preserve these properties, the integrand must be nondecreasing in state; this is why we require the payoff to be nondecreasing. However, if $\mathbb{P}$ only depends on the kernel parameter, then we can show that $\int_{\mathcal{X}} V^*(x' \mid f)\mathbb{P}(dx' \mid \hat{h})$ has increasing differences in x and $\hat{h}$, even if the payoff is not necessarily nondecreasing. For details, we refer the reader to Lemma EC.1 in the

e-companion. Substitution of this lemma in the proof of Theorem 2 yields the desired result.

7. Examples

In this section we revisit two of the examples mentioned in the introduction: interdependent security and collaborative filtering. We show that each of these examples can be formalized within the framework developed in this paper, so that the existence and convergence results we have proven apply. (The third example, dynamic search with learning, is discussed in §EC.1 in the e-companion.)

7.1. Example 1: Interdependent Security

We consider a dynamic model of interdependent security in a computer cluster, where the state x gives the security level of a player. Players can improve their security level through investment; an investment a incurs a cost $c(a)$ that is convex and nondecreasing in a. A higher action leads to improvement in the security level, and with no or little investment the security level deteriorates due to depreciation. Thus, a reasonable model for the dynamic evolution of the security level might be the linear dynamics in (17), where $\underline{M} = 0$ and $\bar{M} \gg 0$, and $A = 1$, $B > 0$, and W has a negative expected value. Let $p(x)$ be the probability of a bad event occurring when an individual computer is at the security level x, and let L be the cost of this bad event to the host. Assume p is decreasing and continuous.

We consider a simplified model where at each time step, an individual computer "talks" to a randomly selected computer in the network. (This talk can be in the form of establishing a TCP connection, exchanging data, emails, etc.) Thus, at each time, there is a probability that an individual computer will suffer a bad event because of the security level of the rest of the network. Let $f_{-i,t}(y)$ be the fraction of all computers (except computer i) that have their security level at y at time t. Then, at each time step, computer i receives an expected value that is given as:

$$v(x_{i,t}, \mathbf{x}_{-i,t}) = -p(x_{i,t})L - (1 - p(x_{i,t}))\left(\int_{\mathcal{X}} p(y) f_{-i,t}(dy)\right).$$

The first part of the payoff reflects the security of host i. The scaling factor $1 - p(x_{i,t})$ in the second term is the probability that no bad event happens because of the individual security level. The term $\int_{\mathcal{X}} p(y) f_{-i,t}(dy)$ represents the average security level of the rest of the network. Because p is decreasing and continuous, it is straightforward to verify that the product of these two terms exhibits strategic complementarities between the security level of agent i and the security level of every other agent. It follows that this is a *separable stochastic game with complementarities*.

7.2. Example 2: Collaborative Filtering

As a canonical example, we consider the collaborative filtering system used by a recommendation engine on a movie rental site such as Netflix. We let the state x be the quality of a user's profile, and assume that x takes values in a compact interval. The action a represents the effort put forth in updating her profile, for example, through rating more movies; actions are costly, with $c(a)$ denoting the cost incurred by action a. We assume c is convex. If user i does not put forth any effort at time t, then the profile becomes "stale"; that is, the quality of the profile drops over time. Thus, in this model the quality can be modeled via dynamics as in (17) as well, where $A = 1$, $B > 0$, and W has negative expected value.

Based on the quality of a user's profile x as well as the profile of other users in the system (captured by the population state f), the recommendation system suggests a movie to a user. Let $v(x, f)$ denote the expected desirability of the movie recommended to a user, given their profile quality x and the population state f. Observe that v will increase if x increases, since a more accurate profile results in more accurate recommendations. However, for most collaborative filtering systems, it is also the case that *if others have higher-quality profiles, then the marginal return to a higher-quality profile is higher;* for example, this would be the case under a nearest neighbor algorithm as is commonly used by a variety of online recommendation systems. Thus, such a model is a separable stochastic game with strategic complementarities. Collaborative filtering systems are one example of a setting with *positive network effects*; games with strategic complementarities are commonly used to model settings with positive network effects.

8. Conclusion

This paper has considered existence of MFE in games that exhibit *strategic complementarities* in the states of the players. Our proofs exploit monotonicity and complementarity properties of the model primitives to demonstrate that there exist both a "largest" and "smallest" MFE among all equilibria where the strategy is nondecreasing in the state. Further, we demonstrate that there exist natural myopic learning dynamics that converge to these equilibria. Finally, we apply our results in the context and illustrate how specific examples of games with complementarities may be analyzed using our techniques.

We conclude by noting two extensions that can be developed for the models described here.

1. *Types.* In our model players are homogeneous; however, this is not a consequential restriction, and is made primarily for convenience. More generally, we can extend the definition of a stochastic game by assuming that there exists a finite *type space* Δ, with $\pi(x, a, f; \delta)$ and $\mathbf{P}(\cdot \mid x, a, f; \delta)$ the payoff and transition kernel, respectively, of a type δ player. Further, we assume that the probability a player is of type δ is given by $\psi(\delta)$. With this extension, as long as the conditions of Definition 7 are satisfied for each δ, it is straightforward to extend our existence, convergence, and comparative statics results. The main technical issue is

that now an MFE must provide an optimal strategy μ_δ and population state f_δ for each δ. We omit the details.

2. *Multidimensional state and action spaces.* A more difficult extension involves models where the state and action spaces may be multidimensional lattices. The main challenge here arises because the set of distributions on a multidimensional compact lattice $\mathscr{X}$ is *not* generally a lattice in the first-order stochastic dominance ordering; see Kamae et al. (1977) for details. However, first-order stochastic dominance *does* give a closed partial order on the set of distributions on $\mathscr{X}$.

We can leverage this fact as follows. Suppose that in addition to the conditions of Definition 7, the action set is a fixed lattice $\mathscr{A}$ for all x, that is, $\mathscr{A}(x) = \mathscr{A}$ for all x.[4] Further, suppose the model primitives (payoff and transition kernel) are all continuous in state, action, and population state—that is, Assumption 1 is satisfied. Then Kleene's fixed-point theorem (Kleene 1971) can be used to establish existence of, and convergence to, an MFE. Kleene's fixed-point theorem states that if X is a space with a closed partial order and a smallest element, then any monotone continuous function from X to itself possesses a fixed point. We omit the details of this argument, because it is essentially identical to our preceding development.

We do emphasize, however, that our analysis of *separable* stochastic games is intimately tied to the assumption that state and action spaces are single dimensional. In particular, our proof techniques rely heavily on the scalar nature of the action and kernel parameter spaces (cf. Lemma 1); relaxing these conditions remains an open direction.

Supplemental Material

Supplemental material to this paper is available at http://dx.doi.org/10.1287/opre.2013.1192.

Appendix A. Proofs: Section 3.1

We start with the lemma that demonstrates that optimal strategies exist and can be identified via Bellman's equation; the proof uses standard results from dynamic programming. See the e-companion (supplementary material) for the proof.

LEMMA 2. *For each $f \in \tilde{\mathfrak{F}}$, $\mathscr{P}(f)$ is nonempty. Furthermore, $\mu \in \mathscr{P}(f)$ if and only if for each x:*

$$\mu(x) \in \underset{a \in \mathscr{A}(x)}{\arg\max}\left\{\pi(x, a, f) + \beta \int_{\mathscr{X}} V^*(x' \mid f)\mathbb{P}(dx' \mid x, a, f)\right\}$$

The next three lemmas combine to show that the value function $V^*(x \mid f)$ has increasing differences in x and f.

LEMMA 3. *Suppose that $U(x \mid f)$ is a nondecreasing bounded function in x and has increasing differences in x and f. Define*

$$T(x, a, f) = \int_{\mathscr{X}} U(x' \mid f)\mathbb{P}(dx' \mid x, a, f). \tag{A1}$$

Then, $T(x, a, f)$ is nondecreasing in x and a, supermodular in (x, a), and has increasing differences in (x, a) and f.

PROOF. By Definition 7, $\mathbb{P}(\cdot \mid x, a, f)$ is stochastically nondecreasing in x and a and stochastically supermodular in (x, a). Since $U(x \mid f)$ is a nondecreasing bounded function, it follows from the definition of stochastically nondecreasing and stochastically supermodular that $T(x, a, f)$ is nondecreasing in x and a and supermodular in (x, a) for fixed f.

Fix $\hat{x} \geqslant x$, $\hat{a} \geqslant a$, and $\hat{f} \succeq_{SD} f$ and define $\hat{T}(x, a, f, g) = \int_{\mathscr{X}} U(x' \mid f)\mathbb{P}(dx' \mid x, a, g)$. To prove $T(x, a, f)$ has increasing differences in (x, a) and f, it suffices to show that

$$\hat{T}(\hat{x}, \hat{a}, \hat{f}, \hat{f}) - \hat{T}(x, a, \hat{f}, \hat{f})$$
$$\geqslant \hat{T}(\hat{x}, \hat{a}, f, f) - \hat{T}(x, a, f, f). \tag{A2}$$

Let us fix g; since $U(x \mid f)$ has increasing differences in x and f, $U(x \mid \hat{f}) - U(x \mid f)$ is a nondecreasing function of x. Since $\mathbb{P}(\cdot \mid \hat{x}, a, g) \succeq_{SD} \mathbb{P}(\cdot \mid x, a, g)$ by Definition 7, it follows that:

$$\hat{T}(\hat{x}, a, \hat{f}, g) - \hat{T}(\hat{x}, a, f, g)$$
$$\geqslant \hat{T}(x, a, \hat{f}, g) - \hat{T}(x, a, f, g). \tag{A3}$$

Also by Definition 7, $\mathbb{P}(\cdot \mid \hat{x}, \hat{a}, g) \succeq_{SD} \mathbb{P}(\cdot \mid \hat{x}, a, g)$, which implies that

$$\hat{T}(\hat{x}, \hat{a}, \hat{f}, g) - \hat{T}(\hat{x}, \hat{a}, f, g)$$
$$\geqslant \hat{T}(\hat{x}, a, \hat{f}, g) - \hat{T}(\hat{x}, a, f, g). \tag{A4}$$

Using Equations (A3) and (A4) and rearranging the terms, we get that

$$\hat{T}(\hat{x}, \hat{a}, \hat{f}, g) - \hat{T}(x, a, \hat{f}, g)$$
$$\geqslant \hat{T}(\hat{x}, \hat{a}, f, g) - \hat{T}(x, a, f, g). \tag{A5}$$

Now let $\hat{g} \succeq_{SD} g$ and note that $\mathbb{P}(\cdot \mid x, a, g)$ has increasing differences in (x, a) and g by Definition 7. Also, note that $U(x \mid f)$ is a bounded nondecreasing function of x. This implies that $\hat{T}(x, a, f, g)$ has increasing differences in (x, a) and g. That is,

$$\hat{T}(\hat{x}, \hat{a}, \hat{f}, \hat{g}) - \hat{T}(x, a, \hat{f}, \hat{g})$$
$$\geqslant \hat{T}(\hat{x}, \hat{a}, \hat{f}, g) - \hat{T}(x, a, \hat{f}, g). \tag{A6}$$

From Equations (A5) and (A6), we get that for any $\hat{x} \geqslant x$, $\hat{a} \geqslant a$, $\hat{f} \succeq_{SD} f$, and $\hat{g} \succeq_{SD} g$ we have

$$\hat{T}(\hat{x}, \hat{a}, \hat{f}, \hat{g}) - \hat{T}(x, a, \hat{f}, \hat{g}) \geqslant \hat{T}(\hat{x}, \hat{a}, f, g) - \hat{T}(x, a, f, g).$$

Taking $\hat{f} = \hat{g}$ and $f = g$ in the above equation shows that Equation (A2) is true, which proves the lemma. $\square$

LEMMA 4. *Suppose that $U(x \mid f)$ is a nondecreasing bounded function in x that has increasing differences in x and f. Define*

$$U^*(x \mid f) = \sup_{a \in \mathscr{A}(x)}\left\{\pi(x, a, f) + \beta \int_{\mathscr{X}} U(x' \mid f)\mathbb{P}(dx' \mid x, a, f)\right\}.$$

Then $U^*(x \mid f)$ is nondecreasing in x and has increasing differences in x and f.

PROOF SKETCH. Define $W(x, a, f) = \pi(x, a, f) + \beta \int_{\mathscr{X}} U(x' \mid f) \cdot \mathbb{P}(dx' \mid x, a, f) = \pi(x, a, f) + \beta T(x, a, f)$. From Lemma 3, we know that $T(x, a, f)$ is nondecreasing in x and a, supermodular in (x, a), and has increasing differences in (x, a) and f. From Definition 7, we get that $W(x, a, f)$ is nondecreasing in x, supermodular in (x, a), and has increasing differences in (x, a) and f.

To complete the proof, observe that if $\mathscr{A}(x) = A$ for all x, then it would be trivial that $U^*(x \mid f) = \sup_{a \in \mathscr{A}(x)} W(x, a, f)$ is nondecreasing in x. The general argument is slightly more technical, but follows the same logic. To show that $U^*(x \mid f)$ has increasing differences, we follow an argument inspired by Lemma A.1 of Hopenhayn and Prescott (1992). To see the complete details of these derivations, we refer the reader to the e-companion. $\square$

LEMMA 5. *$V^*(x \mid f)$ is nondecreasing in x and has increasing differences in x and f.*

PROOF. Let $V_0(x \mid f) = 0$ for all x, and let:

$$V_{k+1}(x \mid f) = \sup_{a \in \mathscr{A}(x)} \left\{ \pi(x, a, f) + \beta \int_{\mathscr{X}} V_k(x' \mid f) \mathbb{P}(dx' \mid x, a, f) \right\}; \quad (A7)$$

this is value iteration. By the preceding lemma, every V_k is nondecreasing in x and has increasing differences in x and f. Under our assumptions, value iteration converges starting from the zero function (Bertsekas 2001), that is, for all x, $V_k(x \mid f) \to V^*(x \mid f)$ as $k \to \infty$. Since monotonicity and increasing differences are preserved upon taking limits, we conclude $V^*(x \mid f)$ is nondecreasing in x and has increasing differences in x and f. $\square$

We now apply Topkis' Theorem in the next lemma to conclude that the set of optimal strategies is monotone; see the ecompanion for the proof.

LEMMA 6. *For each x and f, define the set $\Omega(x, f)$ as:*

$$\Omega(x, f) = \operatorname*{argmax}_{a \in \mathscr{A}(x)} \left\{ \pi(x, a, f) + \beta \int_{\mathscr{X}} V^*(x' \mid f) \mathbb{P}(dx' \mid x, a, f) \right\}. \quad (A8)$$

Then, Ω is nondecreasing in (x, f).
Further,

$$\bar{p}(f)(x) = \sup \Omega(x, f); \quad \underline{p}(f)(x) = \inf \Omega(x, f),$$

where $\bar{p}$ and $\underline{p}$ are the strategies defined in (11). Both $\bar{p}(f)$ are $\underline{p}(f)$ are nondecreasing in f, and for fixed f both strategies are also nondecreasing in x.

We now turn our attention to $\mathscr{D}$. Given any strategy μ and population state f, define a map $Q_{\mu, f} : \tilde{\mathfrak{F}} \to \tilde{\mathfrak{F}}$ according to the kernel induced by μ and f, that is, for all Borel sets S:

$$Q_{\mu, f}(g)(S) = \int_{\mathscr{X}} \mathbb{P}(S \mid x, \mu(x), f)\, g(dx).$$

(This is Equation (16).)

LEMMA 7. *Suppose $f' \succeq_{SD} f$, $g' \succeq_{SD} g$, and $\mu' \trianglerighteq \mu$, and both μ' and μ are nondecreasing, then $Q_{\mu', f'}(g') \succeq_{SD} Q_{\mu, f}(g)$.*

PROOF. Let ϕ be a bounded nondecreasing real-valued function on $\mathscr{X}$. We need to show that for every $x \in \mathscr{X}$, we have

$$\int_{\mathscr{X}} \int_{\mathscr{X}} \phi(y) \mathbb{P}(dy \mid x, \mu'(x), f')\, g'(dx)$$
$$\geqslant \int_{\mathscr{X}} \int_{\mathscr{X}} \phi(y) \mathbb{P}(dy \mid x, \mu(x), f)\, g(dx). \quad (A9)$$

Let us define $H(x; \mu, f) = \int_{\mathscr{X}} \phi(y)\, \mathbb{P}(dy \mid x, \mu(x), f)$. Observe that

$$\int_{\mathscr{X}} \int_{\mathscr{X}} \phi(y) \mathbb{P}(dy \mid x, \mu(x), f)\, g(dx) = \int_{\mathscr{X}} H(x; \mu, f)\, g(dx).$$

Let $x' \geqslant x$ and note that μ is a nondecreasing function of x. From Definition 7, we know that $\mathbb{P}(\cdot \mid x, a, f)$ is stochastically nondecreasing in (x, a), which implies that $\mathbb{P}(\cdot \mid x', \mu(x'), f) \succeq_{SD} \mathbb{P}(\cdot \mid x, \mu(x), f)$. Since ϕ is a nondecreasing function, we get that $H(x'; \mu, f) \geqslant H(x; \mu, f)$.

From Definition 7, we know that $\mathbb{P}(\cdot \mid x, a, f)$ is nondecreasing in a and f. Thus, for any fixed x, we have $\mathbb{P}(\cdot \mid x, \mu'(x), f') \succeq_{SD} \mathbb{P}(\cdot \mid x, \mu(x), f)$. This, along with the fact that ϕ is a nondecreasing function, implies that $H(x; \mu', f') \geqslant H(x; \mu, f)$ for every fixed $x \in \mathscr{X}$.

We now reason as follows:

$$\int_{\mathscr{X}} H(x; \mu', f')\, g'(dx) \geqslant \int_{\mathscr{X}} H(x; \mu, f)\, g'(dx)$$
$$\geqslant \int_{\mathscr{X}} H(x; \mu, f)\, g(dx).$$

Here the first inequality follows from the fact that $H(x; \mu', f') \geqslant H(x; \mu, f)$ and the second inequality follows from that fact that $H(x'; \mu, f) \geqslant H(x; \mu, f)$ for $x' \geqslant x$, and that $g' \succeq_{SD} g$. This proves Equation (A9) and hence proves the lemma. $\square$

LEMMA 8. *Fix $\mu \in \mathfrak{M}_O$ and $f \in \tilde{\mathfrak{F}}$, and suppose μ is nondecreasing in x. Then $\mathscr{D}(\mu, f)$ is a nonempty complete lattice. Further, $\underline{d}(\mu, f)$ and $\bar{d}(\mu, f)$ (as defined in (12)) exist and are both invariant distributions of the Markov process induced by μ and f (cf. (7)).*

Finally, if $f' \succeq_{SD} f$ and $\mu' \trianglerighteq \mu$ and both μ and μ' are nondecreasing, then $\underline{d}(\mu', f') \succeq_{SD} \underline{d}(\mu, f)$ and $\bar{d}(\mu', f') \succeq_{SD} \bar{d}(\mu, f)$.

PROOF. By the preceding lemma, $Q_{\mu, f}(g)$ is nondecreasing in g. By Tarski's theorem, the set of fixed points of $Q_{\mu, f}$ is a nonempty complete lattice. View the set of nondecreasing strategies μ as a partially ordered set $\mathfrak{M}_O$ with the coordinate-wise ordering $\trianglerighteq$. Then $Q_{\mu, f}(\cdot)$ is a nondecreasing function on $\mathfrak{M}_O \times \tilde{\mathfrak{F}} \times \tilde{\mathfrak{F}}$, by the preceding lemma. Theorem 2.5.2 in Topkis (1998) generalizes Tarski's theorem to fixed points of a nondecreasing function parameterized by a partially ordered set $(\mathfrak{M}_O \times \tilde{\mathfrak{F}})$; one consequence of this generalization is that the largest and smallest fixed points are nondecreasing in the parameter. This generalization directly implies that both $\underline{d}(\mu, f)$ and $\bar{d}(\mu, f)$ are nondecreasing in μ and f. $\square$

In the next lemma we establish existence of fixed points of Φ, thus proving Theorem 2.

LEMMA 9. *Let $\underline{\Phi}(f)$ and $\bar{\Phi}(f)$ be defined as in (13). Then $\underline{\Phi}(f), \bar{\Phi}(f) \in \Phi(f)$. Further, both are nondecreasing in f, and thus the sets of their fixed points are each nonempty complete lattices.*

Thus, there exists an MFE for the stochastic game with complementarities Γ: in particular, if f is a fixed point of $\underline{\Phi}$ (respectively, $\bar{\Phi}$), then $(\underline{p}(f), f)$ (respectively, $(\bar{p}(f), f)$) is an MFE.

PROOF. That $\Phi(f)$ is nonempty follows by Lemmas 2 and 8. Observe that if $f' \succeq_{SD} f$, then $\underline{p}(f') \unrhd \underline{p}(f)$ by Lemma 6. Further, $\underline{p}(f')$ and $\underline{p}(f)$ are both nondecreasing in x as well, so by Lemma 8, $\underline{d}(\underline{p}(f'), f') \succeq_{SD} \underline{d}(\underline{p}(f), f)$, establishing that $\underline{\Phi}$ is monotone. That $\underline{\Phi}(f) \in \Phi(f)$ follows from the definition. The proof for $\bar{\Phi}(f)$ is identical. The conclusion regarding fixed points follows from Tarski's theorem. $\square$

Appendix B. Proofs: Section 4

We start with two essential lemmas; proofs of both can be found in the e-companion.

LEMMA 10. *Suppose that* $f_0 \preceq_{SD} f_1 \preceq_{SD} f_2 \cdots$, *and* $\mu_0 \unlhd \mu_1 \unlhd \mu_2 \cdots$. *Then there exists a distribution* f^* *and a strategy* μ^* *such that* f_t *converges weakly to* f^* *as* $t \to \infty$, *and* μ_t *converges pointwise to* μ^* *as* $t \to \infty$.

LEMMA 11. *The strategies* $\underline{p}(f)$ *and* $\bar{p}(f)$ (*cf.* (11)) *are both continuous in* f, *and* $Q_{\mu,g}(f)$ (*cf.* (16)) *is continuous in* μ, g, *and* f, *where we endow* $\mathfrak{F}$ *with the topology of weak convergence, and* $\mathfrak{M}_O$ *with the topology of pointwise convergence.*

PROOF OF PROPOSITION 2. Because $f_{t+1} = \underline{\Phi}(f_t)$ and $\underline{\Phi}$ is monotone by Lemma 9, it follows that $f_0 \preceq_{SD} f_1 \preceq_{SD} f_2 \cdots$. Since $\mu_t = \underline{p}(f_t)$, and $\underline{p}$ is monotone in f_t by Lemma 6, it follows that $\mu_0 \unlhd \mu_1 \unlhd \mu_2 \cdots$. Finally, because $\underline{p}(f)(x)$ is nondecreasing in x for every f, it follows that μ_t is nondecreasing. By Lemma 10, there exists a limit (μ^*, f^*). Since every μ_t is nondecreasing in x, the limit μ^* must be nondecreasing in x as well.

We now show that because Assumption 1 holds, then the limit point (μ^*, f^*) is the smallest MFE. By Lemma 11, both $\underline{p}(f)$ and $Q_{\mu,g}(f)$ are continuous. This implies that $\mu_t = \underline{p}(f_t) \to \underline{p}(f^*)$ as $t \to \infty$, so $\mu^* = \underline{p}(f^*)$. Further, because $f_{t+1} = \underline{d}(\mu_t, f_t)$, it follows that $Q_{\mu_t, f_t}(f_{t+1}) = f_{t+1}$. Taking limits on the left and right, we have $Q_{\mu^*, f^*}(f^*) = f^*$, that is, $f^* \in \mathscr{D}(\mu^*, f^*)$. Thus, we conclude (μ^*, f^*) is an MFE.

Let $\underline{f}$ be the smallest fixed point of $\underline{\Phi}(f)$, as defined in (14). Observe that at time 0, $f_0 \preceq_{SD} \underline{f}$, since f_0 is the smallest distribution in the lattice $\mathfrak{F}$. Because $\underline{\Phi}$ is monotone, $f_t \preceq_{SD} \underline{f}$ for all t. Since f_t converges weakly to f^*, we conclude $f^* \preceq_{SD} \underline{f}$. On the other hand, observe that μ^* is nondecreasing, so by Corollary 1, we have $\underline{f} \preceq_{SD} f^*$—that is, $f^* = \underline{f}$, as required. $\square$

PROOF OF PROPOSITION 3. We proceed by induction. First note that by Lemma 6, $\underline{p}(f)(x)$ is a nondecreasing strategy in x for each f, so every μ_t is nondecreasing. We start by observing that f_0 is the smallest distribution in $\mathfrak{F}$ in the $\succeq_{SD}$ ordering, so $f_1 \succeq_{SD} f_0$ trivially. Since $\underline{p}$ is monotone in f by Lemma 6, we have $\mu_0 = \underline{p}(f_0) \unlhd \underline{p}(f_1) = \mu_1$.

So now suppose that $f_0 \preceq_{SD} f_1 \preceq_{SD} \cdots \preceq_{SD} f_t$, and $\mu_0 \unlhd \mu_1 \unlhd \cdots \unlhd \mu_t$. Define Q_{μ_t, f_t} according to (16). Then by Lemma 7, Q_{μ_t, f_t} is nondecreasing; because $f_{t+1} = Q_{\mu_t, f_t}(f_t)$, we conclude $f_{t+1} \succeq_{SD} f_t$. The same argument as the preceding paragraph then yields $\mu_{t+1} \unrhd \mu_t$, as required. Applying Lemma 10 yields the convergence result; note that μ^* must be nondecreasing, since every μ_t is nondecreasing.

From Lemma 11, because Assumption 1 holds, we conclude that $\mu^* = \underline{p}(f^*)$, and $Q_{\mu^*, f^*}(f^*) = f^*$—that is, $f^* \in \mathscr{D}(\mu^*, f^*)$. Thus, (μ^*, f^*) is an MFE.

Let $\underline{f}$ be the smallest fixed point of $\underline{\Phi}(f)$, as defined in (14). Observe that at time 0, $f_0 \preceq_{SD} \underline{f}$, because f_0 is the smallest distribution in the lattice $\mathfrak{F}$. Thus, $\mu_0 = \underline{p}(f_0) \unlhd \underline{p}(\underline{f})$, so

$f_1 = Q_{\mu_0, f_0}(f_0) \preceq_{SD} Q_{\underline{p}(f), f}(\underline{f}) = \underline{f}$, where the last equality follows since $\underline{f}$ must be an invariant distribution associated with $\underline{p}(\underline{f})$ and $\underline{f}$. Proceeding inductively, we have $f_t \preceq_{SD} \underline{f}$ for all t. Because f_t converges weakly to f^*, we conclude $f^* \preceq_{SD} \underline{f}$. On the other hand, observe that μ^* is nondecreasing, so by Corollary 1, we have $\underline{f} \preceq_{SD} f^*$—that is, $f^* = \underline{f}$, as required. $\square$

Appendix C. Proofs: Section 6

PROOF OF LEMMA 1. Because h is continuous in a for fixed x, and because $\mathscr{A}$ is a compact interval, it follows by the intermediate value theorem that $H(x)$ is a compact interval. Now if $x' \geqslant x$, then $h(x', a) \geqslant h(x, a)$ for all a (by Definition 9), so we conclude $H(x)$ is nondecreasing in x.

Since $h(x, a)$ is nondecreasing in both x and a, and $c(a)$ is nondecreasing in a, it follows that $C(x, \hat{h})$ is nondecreasing in $\hat{h}$ when x is fixed, and nonincreasing in x when $\hat{h}$ is fixed. Equation (21) also follows because $c(a)$ is nondecreasing in a. Convexity in $\hat{h}$ follows by standard results in convex optimization: since we restrict attention to $\hat{h} \in H(x)$ and h and c are both nondecreasing in a, we can rewrite the constraint $h(x, a) = \hat{h}$ as $h(x, a) \geqslant \hat{h}$ in the definition of $C(x, \hat{h})$, that is, for $\hat{h} \in H(x)$ we have:

$$C(x, \hat{h}) = \inf_{a \in \mathscr{A}: h(x, a) \geqslant \hat{h}} c(a).$$

Now, because $C(x, \hat{h})$ is defined via minimization of a convex objective function over a convex feasible region parameterized by $\hat{h}$, it is convex in $\hat{h}$ (Bertsekas 2009).

Finally, we establish the claim of decreasing differences. Fix $x, x', \hat{h}$, and $\hat{h}'$ as in the statement of the lemma. Define $\alpha_1, \alpha_2, \alpha_3$, and α_4 as optimizing values of a in the definition of $C(x, \hat{h})$, $C(x', \hat{h})$, $C(x, \hat{h}')$, and $C(x', \hat{h}')$, respectively. We have $h(x, \alpha_1) = h(x', \alpha_2) = \hat{h}$, and $h(x, \alpha_3) = h(x', \alpha_4) = \hat{h}'$.

Observe that since $\hat{h}' > \hat{h}$ and h is nondecreasing in action, $\alpha_4 > \alpha_2$, and $\alpha_3 > \alpha_1$. Further, because h is nondecreasing in x, we have $\alpha_4 \leqslant \alpha_3$ and $\alpha_2 \leqslant \alpha_1$. Let $\delta = \alpha_4 - \alpha_2$. Define $g(a) = -h(x, -a)$ for $a \in -\mathscr{A}$; then observe that g is a convex, nondecreasing function on $-\mathscr{A}$. By Lemma 12 (see below), we have $g(-\alpha_2) - g(-\alpha_4) \geqslant g(-\alpha_1) - g(-\alpha_1 - \delta)$. In terms of h, this implies:

$$h(x, \alpha_4) - h(x, \alpha_2) \geqslant h(x, \alpha_1 + \delta) - h(x, \alpha_1). \tag{C1}$$

We can now show that $\alpha_4 - \alpha_2 \leqslant \alpha_3 - \alpha_1$. We have $\hat{h}' - \hat{h} = h(x', \alpha_4) - h(x', \alpha_2) \geqslant h(x, \alpha_4) - h(x, \alpha_2) \geqslant h(x, \alpha_1 + \delta) - h(x, \alpha_1)$. Here, the first inequality follows by supermodularity of h in (x, a) (Definition 9), and the second inequality follows by (C1). Since $h(x, \alpha_1) = \hat{h}$ and $h(x, \alpha_3) = \hat{h}'$, and h is nondecreasing in action, it follows that $\alpha_1 + \delta \leqslant \alpha_3$; that is, $\alpha_3 - \alpha_1 \geqslant \alpha_4 - \alpha_2$.

The result now follows by another application of Lemma 12 (see below), which implies: $c(\alpha_3) - c(\alpha_1) \geqslant c(\alpha_4) - c(\alpha_2)$, or equivalently, $C(x, \hat{h}') - C(x, \hat{h}) \geqslant C(x', \hat{h}') - C(x', \hat{h})$, as required. $\square$

The proof of the following lemma is in the e-companion.

LEMMA 12. *Let* $S \subset \mathbb{R}$ *be convex, and suppose* $g \colon S \to \mathbb{R}$ *is a nondecreasing convex function. Fix* $x, x', y, y' \in S$, *such that* $y \geqslant x$, $y' > y$, $x' > x$, *and* $y' - y \geqslant x' - x$. *Then:*

$$g(y') - g(y) \geqslant g(x') - g(x).$$

PROOF OF PROPOSITION 4. We simply check the conditions outlined in Definition 7. Observe that $v(x, f)$ is nondecreasing in x and has increasing differences in x and f by assumption. In addition, $C(x, \hat{h})$ is convex in $\hat{h}$ and nonincreasing in x, and has decreasing differences in x and $\hat{h}$ by Lemma 1. It follows that $\hat{\pi}(x, \hat{h}, f)$ is nondecreasing in x, supermodular in $(x, \hat{h})$ (the latter as it is separable in x and $\hat{h}$), and has increasing differences in $(x, \hat{h})$ and f. Furthermore, for fixed $\hat{h}$ and f, $\sup_{x \in \mathcal{X}} \pi(x, \hat{h}, f) < \infty$. Thus, the first two conditions of Definition 7 are satisfied.

Next, we consider the transition kernel. Here the desired properties follow by assumption: $\hat{\mathbb{P}}(\cdot \mid x, \hat{h}, f)$ is trivially stochastically supermodular in $(x, \hat{h})$, since it does not depend on x and $\hat{h}$ is scalar. By assumption, the kernel is stochastically nondecreasing in $\hat{h}$ and f, and has increasing differences in $\hat{h}$ and f.

Finally, note that $H(x)$ is a compact interval, H is nondecreasing, and $H(x) \subset [\underline{h}, \bar{h}]$ for all x, which is also compact. Further, observe that for all x and f,

$$\sup_{\hat{h} \in H(x)} \hat{\pi}(x, \hat{h}, f) = v(x, f) - \inf_{\hat{h} \in H(x)} C(x, \hat{h}) = v(x, f) - c(\underline{a}),$$

where the last step follows by Lemma 1. Since v is nondecreasing in x, it follows that $\sup_{\hat{h} \in H(x)} \pi(x, \hat{h}, f)$ is nondecreasing in x.

The requisite continuity properties follow from the continuity properties assumed on v, $\mathbb{P}$, and h. It follows that Γ is a stochastic game with complementarities, as required. □

PROOF OF THEOREM 4. Let $\hat{\Gamma}$ be the equivalent stochastic game with complementarities constructed in Proposition 4. Suppose that $(\hat{\mu}, \hat{f})$ is an MFE of $\hat{\Gamma}$; note that in this case $\hat{\mu}$ is a strategy where $\hat{\mu}(x) \in H(x)$ for all $x \in \mathcal{X}$.

Define a new strategy $\mu: \mathcal{X} \to \mathcal{A}$ as follows. For each x, let $\mu(x)$ be an action such that $h(x, \mu(x)) = \hat{\mu}(x)$. That is, we choose the action $\mu(x)$ to yield exactly the kernel parameter $\hat{\mu}(x)$. Then observe that $\pi(x, \mu(x), g) = \hat{\pi}(x, \hat{\mu}(x), g)$ for all x and g. Since $\hat{\mu}$ is an optimal oblivious strategy for a player given population state f in $\hat{\Gamma}$, by construction of $\hat{\Gamma}$ the strategy μ maximizes the expected discounted payoff to a player given f in the original game Γ. Further, because $\mathbb{P}(\cdot \mid x, \mu(x), g) = \hat{\mathbb{P}}(\cdot \mid x, \hat{\mu}(x), g)$, it follows that f is an invariant distribution of the strategy μ. Thus, (μ, f) is an MFE of the game Γ, as required. □

Endnotes

1. Note that guaranteeing existence of MPE may require the use of randomized strategies. To keep the presentation simple, we do not consider randomization.

2. Note that although Tarski's theorem applies to functions, in our case Φ is a correspondence. Zhou (1994) provides a generalization of Tarski's theorem to correspondences.

3. See also Hopenhayn and Prescott (1992), Topkis (1998), and Smith and McCardle (2002) for other conditions that yield monotonicity of optimal solutions to dynamic programs.

4. We employed the total ordering of $\mathcal{A}(x)$ in proving that the value function $V^*(x \mid f)$ is nondecreasing in x, via Lemma 4; however, if $\mathcal{A}$ does not depend on x, then it is straightforward to check that $V^*(x \mid f)$ is nondecreasing in x even if $\mathcal{A}$ is multidimensional.

References

Acemoglu D, Jensen MK (2013) Aggregate comparative statics. *Games Econom. Behav.* 81:27–49.

Adlakha S, Johari R, Weintraub GY (2013) Equilibria of dynamic games with many players: Existence, approximation, and market structure. *J. Econom. Theory.* http://dx.doi.org/10.1016/j.jet.2013.07.002.

Amir R (2002) Complementarity and diagonal dominance in discounted stochastic games. *Ann. Oper. Res.* 114(1–4):39–56.

Amir R (2005) Discounted supermodular stochastic games: Theory and applications. Technical report, Mimeo, University of Arizona, Tucson.

Bergin J, Bernhardt D (1995) Anonymous sequential games: Existence and characterization of equilibria. *Econom. Theory* 5(3):461–489.

Bertsekas DP (2001) *Dynamic Programming and Optimal Control*, Vol. 2 (Athena Scientific, Nashua, NH).

Bertsekas DP (2009) *Convex Optimization Theory* (Athena Scientific, Nashua, NH).

Bertsekas DP, Shreve SE (1978) *Stochastic Optimal Control: The Discrete-Time Case* (Academic Press, New York).

Billingsley P (1968) *Convergence of Probability Measures* (John Wiley & Sons, Inc., New York).

Blume LE (1993) The statistical mechanics of best-response strategy revision. *Games Econom. Behav.* 11(2):111–145.

Bodoh-Creed A (2012) Approximation of large games with applications to uniform price auctions. *Auctions, Market Mechanisms, and Their Applications* (Springer, Berlin), 54.

Chakrabarti SK (2003) Pure strategy Markov equilibrium in stochastic games with a continuum of players. *J. Math. Econom.* 39(7):693–724.

Curtat L (1996) Markov equilibria in stochastic games with complementaries. *Games Econom. Behav.* 17:177–199.

Diamond P (1982) Aggregate demand management in search equilibrium. *J. Political Econom.* 90:881–894.

Doraszelski U, Pakes A (2007) A framework for applied dynamic analysis in IO. *Handbook of Industrial Organization*, Vol. 3 (Elsevier), 1887–1966.

Duffie D, Malamud S, Manso G (2009) Information percolation with equilibrium search dynamics. *Econometrica* 77(5):1513–1574.

Dutta PK, Majumdar MK, Sundaram RK (1994) Parametric continuity in dynamic programming problems. *J. Econom. Dynam. Control* 18(6):1069–1092.

Echenique F (2003) Mixed equilibria in games of strategic complementarities. *Econom. Theory* 22(1):33–44.

Filar J, Vrieze K (1996) *Competitive Markov Decision Processes* (Springer-Verlag, New York).

Friesz TL, Bernstein D, Smith TE, Tobin RL, Wie BW (1993) A variational inequality formulation of the dynamic network user equilibrium problem. *Oper. Res.* 41(1):179–191.

Fudenberg D, Tirole J (1991) *Game Theory* (MIT Press, Cambridge, MA).

Garcia CE, Prett DM, Morari M (1989) Model predictive control: Theory and practice—A survey. *Automatica* 25(3):335–348.

Glynn P (2004) Distributed algorithms for wireless networks. Presentation, Conf. Stochastic Networks, Montréal, Québec.

Hopenhayn HA (1992) Entry, exit, and firm dynamics in long run equilibrium. *Econometrica* 60(5):1127–1150.

Hopenhayn HA, Prescott EC (1992) Stochastic monotonicity and stationary distributions for dynamic economies. *Econometrica* 60(6):1387–1406.

Horst U (2005) Stationary equilibria in discounted stochastic games with weakly interacting players. *Games Econom. Behav.* 51(1):83–108.

Huang M, Caines PE, Malhamé RP (2007) Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ-Nash equilibria. *IEEE Trans. Automatic Control* 52(9):1560–1571.

Huang M, Malhamé RP, Caines PE (2005) Nash equilibria for large-population linear stochastic systems of weakly coupled agents. Boukas EK, Malhame RP, eds. *Analysis, Control and Optimization of Complex Dynamic Systems* (Springer), 215–252.

Huang M, Malhamé RP, Caines PE (2006) Large population stochastic dynamic games: Closed-loop Mckean-Vlasov systems and the Nash certainty equivalence principle. *Comm. Inform. Systems* 6(3):221–251.

Jovanovic B, Rosenthal RW (1988) Anonymous sequential games. *J. Math. Econom.* 17:77–87.

Kamae T, Krengel U, O'Brien GL (1977) Stochastic inequalities on partially ordered spaces. *Ann. Probab.* 5(6):899–912.

Kleene SC (1971) *Introduction to Metamathematics* (American Elsevier Publishing Co., New York).

Kunreuther H, Heal G (2003) Interdependent security. *J. Risk Uncertainty* 26(2):231–249.

Lasry JM, Lions PL (2007) Mean field games. *Japanese J. Math.* 2(1):229–260.

Maitra A (1968) Discounted dynamic programming on compact metric spaces. *Sankhya: The Indian J. Statist., Ser. A* 30(2):211–216.

Maskin E, Tirole J (1988) A theory of dynamic oligopoly, I and II. *Econometrica* 56(3):549–570.

Mézard M, Montanari A (2009) *Information, Physics, and Computation* (Oxford University Press, Oxford, UK).

Milgrom P, Roberts J (1990) Rationalizability, learning, and equilibrium in games with strategic complementarities. *Econometrica* 58:1255–1277.

Morris S (2000) Contagion. *Rev. Econom. Stud.* 67(1):57–78.

Nowak AS (2007) On stochastic games in economics. *Math. Methods Oper. Res.* 66(3):513–530.

Pakes A, McGuire P (2001) Stochastic algorithms, symmetric Markov perfect equilibrium, and the "curse" of dimensionality. *Econometrica* 69(5):1261–1281.

Shapley LS (1953) Stochastic games. *Proc. Natl. Acad. Sci.* 39:1095–1100.

Sleet C (2001) Markov perfect equilibria in industries with complementarities. *Econom. Theory* 17(2):371–397.

Smith JE, McCardle KF (2002) Structural properties of stochastic dynamic programs. *Oper. Res.* 50(5):796–809.

Stokey NL, Lucas RE Jr, Prescott EC (1989) *Recursive Methods in Economic Dynamics* (Harvard University Press, Cambridge, MA).

Tarski A (1955) A lattice-theoretical fixpoint theorem and its applications. *Pacific J. Math.* 5:285–309.

Tembine H, Le Boudec J-Y, El-Azouzi R, Altman E (2009) Mean field asymptotics of Markov decision evolutionary games and teams. *Proc. GameNets '09*, 140–150.

Topkis DM (1998) *Supermodularity and Complementarity* (Princeton University Press, Princeton, NJ).

Vives X (1990) Nash equilibrium with strategic complementarities. *J. Math. Econom.* 19:305–321.

Vives X (2009) Strategic complementarity in multi-stage games. *Econom. Theory* 40(1):151–171.

Weintraub GY, Benkard CL, Van Roy B (2008) Markov perfect industry dynamics with many firms. *Econometrica* 76(6):1375–1411.

Weintraub GY, Benkard CL, Van Roy B (2011) Industry dynamics: Foundations for models with an infinite number of firms. *J. Econom. Theory* 146(5):1965–1994.

Wiecek P, Altman E, Hayel Y (2011) Stochastic state dependent population games in wireless communication. *IEEE Trans. Automatic Control* 56(3):492–505.

Wunderlich KE, Kaufman DE, Smith RL (2000) Link travel time prediction for decentralized route guidance architectures. *IEEE Trans. Intelligent Transportation Systems* 1(1):4–14.

Yin H, Mehta PG, Meyn SP, Shanbhag UV (2010) Synchronization of coupled oscillators is a game. *IEEE Conf. Decision and Control* (IEEE, Piscataway, NJ).

Zhou L (1994) The set of Nash equilibria of a supermodular game is a complete lattice. *Games Econom. Behav.* 7:295–300.

Sachin Adlakha is a postdoctoral scholar in the Center for Mathematics of Information at the California Institute of Technology. His research interests lie at the intersection of operations/management science and engineering with applications to dynamic auctions, electricity grids, and related areas.

Ramesh Johari is an associate professor at Stanford University and the Cisco Faculty Scholar in the School of Engineering, with a full-time appointment in the Department of Management Science and Engineering (MS&E), and courtesy appointments in the Departments of Computer Science and Electrical Engineering (EE). He is a member of the operations research group in MS&E, the Information Systems Laboratory in EE, and the Institute for Computational and Mathematical Engineering.