

Equilibria of Dynamic Games with Many Players: Existence, Approximation, and Market Structure*

April 2011

Sachin Adlakha[†]

Ramesh Johari[‡]

Gabriel Y. Weintraub[§]

Abstract

In this paper we study stochastic dynamic games with many players that are relevant for a wide range of social, economic, and engineering applications. The standard solution concept for such games is *Markov perfect equilibrium* (MPE), but it is well known that MPE computation becomes intractable as the number of players increases. Further, MPE demands a perhaps implausible level of rationality on the part of players in large games. In this paper we instead consider *stationary equilibrium* (SE), where players optimize assuming the empirical distribution of others' states remains constant at its long run average. We make three main contributions that provide foundations for using SE. First, we provide exogenous conditions over model primitives to ensure stationary equilibria exist, in a general model with possibly unbounded state spaces. Second, we show that the same conditions that ensure existence of SE also ensure that SE is a good approximation to MPE in large finite games. Finally, we consider a series of applications, including dynamic oligopoly models, supply chain competition, and consumer learning. These examples highlight that our conditions amount to a dichotomy between “decreasing” and “increasing” returns to larger states; SE approximates MPE well in the former case in which the equilibrium market structure becomes fragmented in the limit. In the latter case, SE may not approximate MPE well.

1 Introduction

A common framework to study dynamic economic systems of interacting agents is a *stochastic game*, as pioneered by Shapley (1953). In a stochastic game agents' actions directly affect underlying state variables that influence their payoff. The state variables evolve according to a Markov process in discrete time, and players maximize their infinite horizon expected discounted payoff. Stochastic games provide a valuable

*The authors are grateful for helpful conversations with Vineet Abhishek, Lanier Benkard, Peter Glynn, Andrea Goldsmith, Ben Van Roy, and Assaf Zeevi, and seminar participants at the INFORMS Annual Meeting and the Behavioral and Quantitative Game Theory Workshop. This work was supported by DARPA under the ITMANET program, and by the National Science Foundation.

[†]S. Adlakha is with the Center for Mathematics of Information, California Institute of Technology, Pasadena, CA, 91125. adlakha@caltech.edu

[‡]R. Johari is with the Department of Management Science and Engineering, Stanford University, Stanford, CA, 94305. ramesh.johari@stanford.edu

[§]G. Y. Weintraub is with the Columbia Business School, Columbia University, New York, NY, 10027. gweintraub@columbia.edu

general framework for a range of economic settings, including dynamic oligopoly models, supply chain competition, and consumer learning.

The standard solution concept for stochastic games is *Markov perfect equilibrium* (MPE) (Maskin and Tirole 1988), where a player's equilibrium strategy depends on the current state of all players. However, MPE presents two significant obstacles as an analytical tool, particularly as the number of players grows large. First is *computability*: the state space expands in dimension with the number of players, and thus the "curse of dimensionality" kicks in, making computation of MPE infeasible in many problems of practical interest (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 all other agents.

To overcome these difficulties, previous research has considered an asymptotic regime in which the number of agents is infinite (Jovanovic and Rosenthal 1988, Hopenhayn 1992). In this case, individuals take a simpler view of the world: they postulate that fluctuations in the empirical distribution of other players' states have "averaged out" due to a law of large numbers, and thus they optimize holding the state distribution of other players fixed. Based on this insight, this approach considers an equilibrium concept where agents optimize only with respect to the long run average of the distribution of other players' states; Hopenhayn (1992) refers to this concept as *stationary equilibrium* (SE), and we adopt his terminology. SE are much simpler to compute and analyze than MPE, making this a useful approach across a wide range of applications (see for example, Luttmer 2007, Melitz 2003, and Klette and Kortum 2004 for applications in economics).

This paper is devoted to a unifying analysis of two key issues pertaining to SE: first, when does it exist? And second, if it exists, does it provide a good approximation to MPE in models with a finite number of agents? Our main results provide a parsimonious collection of exogenous conditions over model primitives that guarantee existence of a SE, and ensure that an appropriate approximation property holds. An important insight of our work is that the conditions we impose to ensure existence of SE are *identical* to those used to establish that SE approximates MPE well as the number of agents increases; thus approximation need not be verified separately. As we illustrate with examples, our results apply to a variety of dynamic games. In this way we provide theoretical foundations for using SE.

A key feature of our analysis is that we focus on models with *unbounded* state spaces. This is distinct from prior work on SE, which typically studies models with *compact* state spaces. Our analysis is informative in part because consideration of unbounded state spaces reveals natural tradeoffs between those models where SE is likely to serve as a useful approximate solution concept, and those where SE lacks predictive power. In particular, we show that our conditions on model primitives can be interpreted as enforcing "decreasing returns to higher states;" conversely, our analysis suggests that when these conditions are violated, the resulting models exhibit "increasing returns to higher states," and SE are not expected to provide accurate approximations, or may not even exist.

"Decreasing returns to higher states" ensure that the distribution of agents' states does not place excessive mass at larger states. Formally, our conditions on model primitives ensure that when any agent computes its optimal strategy given the long run average of other players' states, the resulting strategy induces state dynamics that have a sufficiently *light tail*; this is established using a Foster-Lyapunov argument. This result

is useful because (1) it establishes an appropriate compactness property to prove our existence theorem; and (2) it ensures that under our conditions, any SE induces a population state distribution with a light tail. As originally observed in Weintraub et al. (2008, 2010), the light tail condition is a key property needed to ensure SE approximate MPE well. In a light-tailed equilibrium, no single agent is “dominant;” without such a condition it is not possible for agents’ to rationally ignore the state fluctuations of their dominant competitors.

We illustrate the insights obtained through our analysis with a range of dynamic oligopoly models (including quality investments and learning-by-doing), a supply chain competition model, and a consumer learning model. The analysis of dynamic oligopoly models is relevant for the industrial organization literature pioneered by Sutton (1991), which aims to identify broad structural properties in industries that yield a fragmented or a concentrated market structure. Our results provide conditions for which a broad class of models that exhibit “decreasing returns to higher states” yield light-tailed SE, and hence, the market structure becomes fragmented in the limit. A novelty of our analysis is that it is done in a fully dynamic framework.

The remainder of the paper is organized as follows. Section 2 describes related literature. In Section 3 we introduce our stochastic game model, and we define both MPE and SE. Section 3.2 discusses extensions of our baseline model. We then preview our results and discuss the motivating examples above in detail in Section 4. In Section 5, we develop exogenous conditions over model primitives that ensure existence of light-tailed SE. In Section 6, we show that under our conditions any light-tailed SE approximates MPE asymptotically. We summarize our results in Section 7 and discuss some possible extensions. Due to space constraints, some proofs have been deferred to our companion technical report; see Adlakha et al. (2011). Key proofs have been provided in the appendix.

2 Related Literature

Stationary equilibrium (SE) is sometimes called *mean field equilibrium* because of its relationship to mean field models in physics, where large systems exhibit macroscopic behavior that is considerably more tractable than their microscopic description. In the context of stochastic games, SE and related approaches have been proposed under a variety of monikers across economics and engineering; see, e.g., studies of anonymous sequential games (Jovanovic and Rosenthal 1988), Nash certainty equivalent control (Huang et al. 2007), and mean field games (Lasry and Lions 2007). SE has also been studied in recent works on sensitivity analysis in aggregate games (Acemoglu and Jensen 2009), coupling of oscillators (Yin et al. 2010), scaling behavior of markets (Bodoh-Creed 2010), and stochastic games with complementarities (Adlakha and Johari 2010).

Existence of equilibrium in general stochastic games have been typically established only in restricted classes such as zero-sum games and games of identical interest; see Mertens et al. (1994) for background. Doraszelski and Satterthwaite (2010) and Escobar (2008) show existence of MPE for different classes of stochastic games under appropriate concavity assumptions. Our work is particularly related to Jovanovic and Rosenthal (1988), that establishes existence of SE for compact state spaces, and to Hopenhayn (1992) that studies existence of SE in a specific industry model. Adlakha and Johari (2010) provide an existence result

based on lattice theoretic methods in games with strategic complementarities. These papers study a different setting to ours and do not establish an approximation theorem. Several prior papers have considered various notions of approximation properties for SE in specific settings, either with bounded state spaces (Glynn 2004, Tembine et al. 2009, Bodoh-Creed 2010) or with an exogenous compactness assumption (Adlakha et al. 2010), or in linear-quadratic payoff models (Huang et al. 2007, Adlakha et al. 2008).

Our paper is closely related to Weintraub et al. (2010), who study a class of industry dynamic models. They show that if a SE satisfies an appropriate light-tail condition, then it approximates MPE well as the number of firms grows. Our paper provides several important contributions with respect to Weintraub et al. (2010). First, we consider a more general stochastic game model (though we do not consider entry and exit of firms as they do), and a stronger approximation property. Second, the light-tail condition used to prove the approximation result is a condition over *equilibrium outcomes*; by contrast, we provide conditions over *model primitives* that guarantee all SE are light-tailed and hence approximate MPE asymptotically. Finally, we provide a novel result pertaining to existence of SE, particularly over unbounded state spaces. Weintraub et al. (2010) also consider an analog of SE called “oblivious equilibrium” (OE) for models with finitely many agents (Weintraub et al. 2008). They study the relation between OE and SE by analyzing the hemicontinuity of the OE correspondence at the point where number of firms becomes infinite. Weintraub et al. (2008) also show that under a *uniform* light-tail condition over the sequence of OE, OE approximates MPE well as the number of agents grows.

3 Preliminaries and Definitions

In this section we define our general model of a stochastic game and then proceed to define Markov perfect equilibrium (MPE) and stationary equilibrium (SE). We then conclude by defining our approximation notion for SE, called the asymptotic Markov equilibrium property.

3.1 Stochastic Game Model

Compared to standard stochastic games in the literature (Shapley 1953), in our model, every player has an individual state. Players are coupled through their payoffs and state transitions. A stochastic game has the following elements:

Time. The game is played in discrete time. We index time periods by $t = 0, 1, 2, \dots$

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 $\mathcal{X} \subseteq \mathbb{Z}^d$ is a subset of the d -dimensional integer lattice. We use \mathbf{x}_t to denote the state of all players at time t and $\mathbf{x}_{-i,t}$ to denote the state of all players except player i at time t .¹

Action. The action taken by player i at time t is denoted by $a_{i,t} \in \mathcal{A}$, where $\mathcal{A} \subseteq \mathbb{R}^q$ is a subset of the q -dimensional Euclidean space. We use \mathbf{a}_t to denote the action of all players at time t .

¹For indication of how to proceed with compact but not necessarily discrete state spaces, we refer the reader to the recent independent work of Bodoh-Creed (2010).

Transition Probabilities. The state of a player evolves in a Markov fashion. Formally, let $h_t = \{\mathbf{x}_0, \mathbf{a}_0, \dots, \mathbf{x}_{t-1}, \mathbf{a}_{t-1}\}$ denote the *history* up to time t . Conditional on h_t , players' states at time t are *independent* of each other. This assumption implies that random shocks are idiosyncratic, ruling out aggregate random shocks that are common to all players. The assumption is important to derive our asymptotic results. Player i 's state $x_{i,t}$ at time t depends on the past history h_t only through the state of player i at time $t-1$, $x_{i,t-1}$; the states of other players at time $t-1$, $\mathbf{x}_{-i,t-1}$; and the action taken by player i at time $t-1$, $a_{i,t-1}$. We represent the distribution of the next state as a transition kernel \mathbf{P} , where:

$$\mathbf{P}(x'_i | x_i, a_i, \mathbf{x}_{-i}) = \text{Prob}(x_{i,t+1} = x'_i | x_{i,t} = x_i, a_{i,t} = a_i, \mathbf{x}_{-i,t} = \mathbf{x}_{-i}). \quad (1)$$

Payoff. In a given time period, if the state of player i is x_i , the state of other players is \mathbf{x}_{-i} , and the action taken by player i is a_i , then the single period payoff to player i is $\pi(x_i, a_i, \mathbf{x}_{-i}) \in \mathbb{R}$.

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: $\sum_{t=0}^{\infty} \beta^t \pi(x_{i,t}, a_{i,t}, \mathbf{x}_{-i,t})$.

In a variety of games, coupling between players is independent of the identity of the players. The notion of *anonymity* captures scenarios where the interaction between players is via aggregate information about the state (e.g., see Jovanovic and Rosenthal 1988). Let $\mathbf{f}_{-i,t}^{(m)}(y)$ denote the fraction of players (excluding player i) that have their state as y at time t , i.e.:

$$\mathbf{f}_{-i,t}^{(m)}(y) = \frac{1}{m-1} \sum_{j \neq i} \mathbf{1}_{\{x_{j,t}=y\}}, \quad (2)$$

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 $\mathbf{f}_{-i,t}^{(m)}$ as the *population state* at time t (from player i 's point of view).

Definition 1 (Anonymous Stochastic Game). *A stochastic game is called an anonymous stochastic game if the payoff function $\pi(x_{i,t}, a_{i,t}, \mathbf{x}_{-i,t})$ and transition kernel $\mathbf{P}(x'_{i,t} | x_{i,t}, a_{i,t}, \mathbf{x}_{-i,t})$ depend on $\mathbf{x}_{-i,t}$ only through $\mathbf{f}_{-i,t}^{(m)}$. In an abuse of notation, we write $\pi(x_{i,t}, a_{i,t}, \mathbf{f}_{-i,t}^{(m)})$ for the payoff to player i , and $\mathbf{P}(x'_{i,t} | x_{i,t}, a_{i,t}, \mathbf{f}_{-i,t}^{(m)})$ for the transition kernel for player i .*

For the remainder of the paper, we focus our attention on anonymous stochastic games. For ease of notation, we often drop the subscript i and t and denote a generic transition kernel by $\mathbf{P}(\cdot | x, a, f)$, and a generic payoff function by $\pi(x, a, f)$, where f represents the population state of players other than the player under consideration. Anonymity requires that a firm's single period payoff and transition kernel depend on the states of other firms via their empirical distribution over the state space, and not on their specific identify. The examples we discuss in the next section satisfy this assumption. Second, in an anonymous stochastic game the functional form of the payoff function is the same, regardless of the number of players m . In that sense, we often interpret the profit function $\pi(x, a, f)$ as representing a limiting regime in which the number of agents is infinite. In Section 4 we discuss how to derive this limiting profit function in different applications. Moreover, in Section 6 we briefly discuss how our results can be extended to include the case where there is a sequence of payoff functions that depends on the number of agents.

We introduce some additional useful notation. Let \mathfrak{F} be the set of all possible population states on \mathcal{X} :

$$\mathfrak{F} = \left\{ f : \mathcal{X} \rightarrow [0, 1] \mid f(x) \geq 0, \sum_{x \in \mathcal{X}} f(x) = 1 \right\}. \quad (3)$$

In addition, we let $\mathfrak{F}^{(m)}$ denote the set of all population states in \mathfrak{F} over $m - 1$ players, i.e.:

$$\mathfrak{F}^{(m)} = \left\{ f \in \mathfrak{F} : \text{there exists } \mathbf{x} \in \mathcal{X}^{m-1} \text{ with } f(y) = \frac{1}{m-1} \sum_j \mathbf{1}_{\{x_j=y\}}, \forall y \in \mathcal{X} \right\}.$$

3.2 Extensions to the Basic Model

We briefly mention two extensions for which all our results follow; the technical details are omitted, and the reader is referred to Adlakha et al. (2011).

First, note that players are ex-ante *homogeneous* in the model considered, in the sense that they share the same model primitives. This is not a particularly consequential choice, and is made primarily for notational convenience; indeed, by an appropriate redefinition of state we can model agent heterogeneity via types.

Second, note that in the game defined here, players are coupled through their states: both the transition kernel and the payoff depend on the current state of all players. All the results of this paper naturally extend to a setting where players may also be coupled through their *actions*, i.e., where the transition kernel and payoff may depend on the current actions of all players as well.

To model a game where players are coupled through actions, we now assume that f is a distribution over both states and actions. We refer to f as the *population state-action profile* (to distinguish it from just the population state, which is the marginal distribution of f over \mathcal{X}). For simplicity, since our basic model assumes state spaces are discrete, whenever players are coupled through actions we restrict attention to games with a *finite* action space $S \subset \mathbb{Z}^k$. Thus the population state-action profile is a distribution over $\mathcal{X} \times S$. Because in this setting we restrict attention to finite action spaces, we assume that players maximize payoffs with respect to randomized strategies over S .² (See Section 5.2.1 for further details on games with finite action spaces.)

We again let $x_{i,t} \in \mathcal{X}$ be the state of player i at time t , where $\mathcal{X} \subseteq \mathbb{Z}^d$. We let $s_{i,t} \in S$ be the action taken by player i at time t . Let $\mathbf{f}_{-i,t}^{(m)}$ denote the empirical population state-action profile at time t in an m -player game; in other words, $f_{i,t}^{(m)}(x, s)$ is the fraction of players other than i at state x who play s at time t . With these definitions, $x_{i,t}$ evolves according to the transition kernel \mathbf{P} as before, i.e., $x_{i,t+1} \sim \mathbf{P}(\cdot | x_{i,t}, a_{i,t}, \mathbf{f}_{-i,t}^{(m)})$. A player acts to maximize his expected discounted payoff, as before. Note that a player's time t payoff and transition kernel depend on the actions of his competitors, which are chosen *simultaneously* with his own action. Thus to evaluate the time t expected payoffs and transition kernel, a player must take an expectation with respect to the randomized strategies employed by his competitors. With these definitions, all the analysis and results of this paper go through for a game where agents are coupled through actions with modest additional technical work.

We conclude by commenting on the restriction imposed when players are coupled through actions that

²This is done to ensure existence of equilibrium.

the action space must be finite. From a computational standpoint this is not very restrictive, since in many applications discretization is required or can be used efficiently. From a theoretical standpoint, we can analyze games with general compact Euclidean action spaces using techniques similar to this paper, at the expense of additional measure-theoretic complexity, since now the population state-action profile is a measure over a continuous extended state space.

3.3 Markov Perfect Equilibrium

In studying stochastic games, attention is typically focused on *Markov* strategies (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 them as *cognizant* strategies.

Definition 2. Let \mathfrak{M} be the set of cognizant strategies available to a player. That is, $\mathfrak{M} = \{\mu \mid \mu : \mathcal{X} \times \mathfrak{F} \rightarrow \mathcal{A}\}$.

Consider an m -player anonymous stochastic game and let $\mu_i \in \mathfrak{M}$ denote the cognizant strategy used by player i , i.e., we have $a_{i,t} = \mu_i(x_{i,t}, \mathbf{f}_{-i,t}^{(m)})$. The next state of player i is randomly drawn according to the kernel \mathbf{P} :

$$x_{i,t+1} \sim \mathbf{P} \left(\cdot \mid x_{i,t}, \mu_i(x_{i,t}, \mathbf{f}_{-i,t}^{(m)}), \mathbf{f}_{-i,t}^{(m)} \right). \quad (4)$$

We let $\boldsymbol{\mu}^{(m)}$ denote the strategy vector where every player has chosen strategy μ . Define $V^{(m)}(x, f \mid \mu', \boldsymbol{\mu}^{(m-1)})$ to be the expected net present value for a player with initial state x , and with initial population state $f \in \mathfrak{F}^{(m)}$, given that the player follows a strategy μ' and every other player follows the strategy μ . In particular, we have

$$V^{(m)}(x, f \mid \mu', \boldsymbol{\mu}^{(m-1)}) \triangleq \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \pi(x_{i,t}, a_{i,t}, \mathbf{f}_{-i,t}^{(m)}) \mid x_{i,0} = x, \mathbf{f}_{-i,0}^{(m)} = f; \mu_i = \mu', \boldsymbol{\mu}_{-i} = \boldsymbol{\mu}^{(m-1)} \right], \quad (5)$$

where $\boldsymbol{\mu}_{-i}$ denotes the strategies employed by every player except i . Note that state sequence $x_{i,t}$ and population state sequence $\mathbf{f}_{-i,t}^{(m)}$ evolve according to the transition dynamics (4).

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 $V^{(m)}(x, f \mid \boldsymbol{\mu}^{(m)})$ to refer to the expected discounted value as given in equation (5) when every player follows the same cognizant strategy μ .

Definition 3 (Markov Perfect Equilibrium). *The vector of cognizant strategies $\boldsymbol{\mu}^{(m)} \in \mathfrak{M}$ is a symmetric Markov perfect equilibrium (MPE) if for all initial states $x \in \mathcal{X}$ and population states $f \in \mathfrak{F}^{(m)}$ we have $\sup_{\mu' \in \mathfrak{M}} V^{(m)}(x, f \mid \mu', \boldsymbol{\mu}^{(m-1)}) = V^{(m)}(x, f \mid \boldsymbol{\mu}^{(m)})$.*

Thus, a Markov perfect equilibrium is a profile of cognizant strategies that simultaneously maximize

the expected discounted payoff for every player, given the strategies of other players.³ It is a well known fact that computing a Markov perfect equilibrium for a stochastic game is computationally challenging in general (Doraszelski and Pakes 2007). The state space $\mathfrak{F}^{(m)}$ grows too quickly as the number of agents m and/or the number of individual states \mathcal{X} becomes large. Hence, computing MPE is only feasible for models with few agents and few individual states, severely restricting the set of problems for which MPE can be used. The concept of stationary equilibrium alleviates these difficulties.

3.4 Stationary Equilibrium

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. Based on this intuition, related schemes for approximating MPE have been proposed in different application domains via a solution concept we call *stationary equilibrium* or SE (see Sections 1 and 2 for references on SE and related work).

We consider a limiting model with an infinite number of agents in which a law of large numbers holds exactly. In an SE of this model, each player optimizes its payoff assuming the population state is fixed at its long-run average. Hence, a single player's immediate action depends only on his own current state. We call such players *oblivious*, and refer to their strategies as *oblivious strategies*. (This terminology is due to Weintraub et al. 2008.) Formally, we let \mathfrak{M}_O denote the set of (stationary, nonrandomized) oblivious strategies, defined as follows.

Definition 4. Let \mathfrak{M}_O be the set of oblivious strategies available to a player. That is, $\mathfrak{M}_O = \{\mu \mid \mu : \mathcal{X} \rightarrow \mathcal{A}\}$.

Given a strategy $\mu \in \mathfrak{M}_O$, the next state of an oblivious player is randomly distributed according to the transition kernel \mathbf{P} :

$$x_{i,t+1} \sim \mathbf{P}(\cdot \mid x_{i,t}, a_{i,t}, f) \text{ where } a_{i,t} \sim \mu(x_{i,t}). \quad (6)$$

Note that an oblivious player conjectures the population state to be fixed at f and hence its state evolves according to a transition kernel with fixed population state f .

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

$$\tilde{V}(x \mid \mu, f) \triangleq \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \pi(x_{i,t}, a_{i,t}, f) \mid x_{i,0} = x; \mu \right]. \quad (7)$$

Note that the state sequence $x_{i,t}$ is determined by the strategy μ according to the dynamics (6), where the population state is fixed at f . We define the *optimal oblivious value function* $\tilde{V}^*(x \mid f)$ as $\tilde{V}^*(x \mid f) = \sup_{\mu \in \mathfrak{M}_O} \tilde{V}(x \mid \mu, f)$. Note that because an oblivious player does not track the evolution of the population state and its state evolution depends only on the population state f , if an optimal stationary nonrandomized

³Under the assumptions we make later in this paper, it can be shown that for any vector of cognizant strategies of players other than i , an optimal cognizant strategy always exists for player i .

strategy exists, it will only be a function of the player’s current state—i.e., it must be oblivious even if optimizing over cognizant strategies. We capture this optimization step via the correspondence \mathcal{P} defined next.

Definition 5. *The correspondence $\mathcal{P} : \mathfrak{F} \rightarrow \mathfrak{M}_O$ maps a distribution $f \in \mathfrak{F}$ to the set of optimal oblivious strategies for a player. That is, $\mu \in \mathcal{P}(f)$ if and only if $\tilde{V}(x | \mu, f) = \tilde{V}^*(x | f)$ for all x .*

Note that \mathcal{P} maps a distribution to a *stationary, nonrandomized* oblivious strategy. This is typically without loss of generality, since in most models of interest there always exists such an optimal strategy. We later establish that under our assumptions $\mathcal{P}(f)$ is nonempty.

Now suppose that the population state is f , and *all* players are oblivious and play using a stationary strategy μ . Because of averaging effects, we expect that if the number of agents is large, then the long run population state should in fact be an invariant distribution of the Markov process on \mathcal{X} that describes the evolution of an individual agent, with transition kernel (6). We capture this relationship via the correspondence \mathcal{D} , defined next.

Definition 6. *The correspondence $\mathcal{D} : \mathfrak{M}_O \times \mathfrak{F} \rightarrow \mathfrak{F}$ maps the oblivious strategy μ and population state f to the set of invariant distributions $\mathcal{D}(\mu, f)$ associated with the dynamics (6).*

Note that the image of the correspondence \mathcal{D} is empty if the strategy does not result in an invariant distribution. We later establish conditions under which $\mathcal{D}(\mu, f)$ is nonempty.

Assume that every agent conjectures the long run population state to be f and plays an optimal oblivious strategy μ . Stationary equilibrium requires that the equilibrium population state f must in fact be an invariant distribution of the dynamics (6) under the strategy μ and the initially conjectured population state f . The consistency of players’ conjectures is captured in the following definition.

Definition 7 (Stationary Equilibrium). *An oblivious strategy $\mu \in \mathfrak{M}_O$ and a distribution $f \in \mathfrak{F}$ constitute a stationary equilibrium (SE) if $\mu \in \mathcal{P}(f)$ and $f \in \mathcal{D}(\mu, f)$.*

In the event that there exist multiple optimal strategies given f or that the chain induced by μ and f has multiple invariant distributions, the agents must all choose to play the same optimal strategy and they must all have the same conjecture about the equilibrium population state. In many models of interest (such as the examples presented in Section 4), both \mathcal{P} and \mathcal{D} are singletons, so such problems do not arise. For later reference, we define the correspondence $\Phi : \mathfrak{F} \rightarrow \mathfrak{F}$ as follows:

$$\Phi(f) = \mathcal{D}(\mathcal{P}(f), f). \tag{8}$$

Observe that with this definition, a pair (μ, f) is an SE if and only if f is a fixed point of Φ , $f \in \Phi(f)$, such that $\mu \in \mathcal{P}(f)$ and $f \in \mathcal{D}(\mu, f)$

3.5 Approximation

A central goal of this paper is to determine conditions under which SE provides a good approximation to MPE as the number of players grows large. Here we formalize the approximation property of interest,

referred to as the asymptotic Markov equilibrium (AME) property. Intuitively, this property requires that a stationary equilibrium strategy is approximately optimal even when compared against Markov strategies, as the number of players grows large.

Definition 8 (Asymptotic Markov Equilibrium). *A stationary equilibrium (μ, f) possesses the asymptotic Markov equilibrium (AME) property if for all states x and sequences of cognizant strategies $\mu_m \in \mathfrak{M}$, we have:*

$$\limsup_{m \rightarrow \infty} V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)}) - V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)}) \leq 0, \quad (9)$$

almost surely, where the initial population state $f^{(m)}$ is derived by sampling each other player's initial state independently from the probability mass function f .

AME requires that the error when using the SE strategy approaches zero almost surely with respect to the randomness in the initial population state. Hence, AME requires that the SE strategy becomes approximately optimal as the number of agents grows, with respect to population states that have nonzero probability of occurrence when sampling individual states according to the invariant distribution.⁴

4 Preview of Results and Motivating Examples

This paper is concerned with two central questions: First, when do SE exist? And second, when are SE a good approximation to MPE? Our main results provide exogenous conditions over model primitives that ensure that SE exist, and the AME property holds, in a broad class of stochastic games over unbounded state spaces.

Here we briefly preview our main results. First, we establish the existence of SE by applying Kakutani's fixed point theorem to show that the map Φ (cf. (8)) possesses a fixed point. A key requirement in application of Kakutani's theorem is to ensure that the image of Φ , $\Phi(\mathfrak{F})$, is compact (under a conveniently defined norm). Because state spaces are unbounded, the last step requires conditions that ensure agents' optimal oblivious strategies do not yield state dynamics that induce invariant distributions with overly heavy tails, i.e., with excessive mass at larger states. The exogenous conditions we develop address this hurdle by ensuring that the image of Φ has *uniformly light tails* in a sense that we make precise; this also implies that all SE are light-tailed (see Section 5 for details of our existence result).

This approach to existence yields a significant benefit. We later show that an SE has the AME property provided the SE under consideration has a sufficiently light tail, and the model primitives are appropriately continuous. The first property ensures that population states in the system with finitely many players converge to the fixed population state in the limit model. The second property translates convergence of population states to convergence of payoffs, as required by the AME property. More broadly, the light-tail condition ensures that no single dominant agent emerges; in this case players can ignore the fluctuations of

⁴This definition can be shown to be stronger than the definition considered by Weintraub et al. (2008), where AME is defined only in expectation with respect to randomness in the initial population state. Moreover, as noted earlier, under the assumptions we make an optimal cognizant strategy can be shown to exist, for any vector of cognizant strategies of the opponents. Therefore the AME property can be equivalently stated as the requirement that for all x : $\lim_{m \rightarrow \infty} \left(\sup_{\mu_m \in \mathfrak{M}} V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)}) - V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)}) \right) = 0$, almost surely.

their competitors’ states when the number of agents is large. Observe that our existence result ensures existence of light-tailed SE, and is derived under continuity conditions on the model primitives; taken together, these conditions are also sufficient to ensure the AME property holds for any SE (see Section 6 for details).

As is apparent in the preceding discussion, the conditions that ensure light tails in SE play a key role in our analysis. Informally, these conditions enforce a form of “decreasing returns to higher states” in the optimization problem faced by an individual agent. An important goal of this section is to illustrate the dichotomy between models that, broadly, exhibit “decreasing” or “increasing returns to higher states.” As we point out, when the examples below violate the assumptions we require—in particular, in models that exhibit increasing returns to higher states—we also expect that SE will *not* satisfy the AME property, and indeed, may not exist. Thus despite the fact that we only discuss *sufficient* conditions for existence and approximation in this paper, the examples suggest that perhaps these sufficient conditions identify a reasonable boundary between those models that admit analysis via SE, and those that do not. In particular, the first three examples focus on industry models, and provide conditions that ensure “decreasing returns to higher states” so that light-tailed SE exist, and the market structure becomes fragmented in the limit. When these conditions fail, we expect the industry under consideration to become concentrated in the limit—precisely where SE should not provide useful approximations.

In each of the following examples, we provide sufficient conditions for the existence of SE, and to ensure the AME property holds. Each of these propositions are corollaries of the general results derived in Sections 5 and 6. We discuss this relationship in Section 7; technical details of this verification are provided in Adlakha et al. (2011) due to space constraints. For the rest of this section, we consider stochastic games with m players in which the state of a player takes values on \mathbb{Z}_+ .

4.1 Dynamic Oligopoly Models

Dynamic oligopoly models have received significant attention in the recent industrial organization literature (see Doraszelski and Pakes 2007 for a survey). These models are characterized by the following features.

States. Each firm has a state variable that captures its competitive advantage; for example, the state could represent the firm’s product quality, its current productivity level, or its capacity. For concreteness, here we consider the quality ladder model of Pakes and McGuire (1994), where the state $x_{i,t} \in \mathbb{Z}_+$ represents the quality of the product produced by firm i at time t .

Actions. Investments improve the state variable over time. At each time t , firm i invests $a_{i,t} \in [0, \bar{a}]$ to improve the quality of its product. The action changes the state of the firm in a stochastic fashion as described below.

Payoffs. We consider a payoff function derived from price competition under a classic logit demand system. In such a model, there are n consumers in the market. In period t , consumer j receives utility u_{ijt} from consuming the good produced by firm i given by: $u_{ijt} = \theta_1 \ln(x_{it} + 1) + \theta_2 \ln(Y - p_{it}) + \nu_{ijt}$, where $\theta_1, \theta_2 > 0$, Y is the consumer’s income, and p_{it} is the price of the good produced by firm i . Here ν_{ijt} are i.i.d. Gumbel random variables that represent unobserved characteristics for each consumer-good pair.

We assume that there are m firms that set prices in the spot market. For a constant marginal production cost c , there is a unique Nash equilibrium in pure strategies of the pricing game, denoted p_t^* (Caplin and

Nalebuff 1991). For our limit profit function, we consider an asymptotic regime in which the market size n and the number of firms m grow to infinity at the same rate. The limiting profit function corresponds to a logit model of monopolistic competition (Besanko et al. 1990) and is given by $\pi(x, a, f) = \frac{\tilde{c}(x+1)^{\theta_1}}{\sum_y f(y)(y+1)^{\theta_1}} - da$, where \tilde{c} is a constant that depends on the limit equilibrium price, c , θ_2 , and Y . Here the second term is the cost of investment, where $d > 0$ is the marginal cost per unit investment.

Transition dynamics. We use dynamics similar to those in Pakes and McGuire (1994) that have been widely used in dynamic oligopoly models. Compared to that paper, we assume random shocks are idiosyncratic. At each time period, a firm's investment of a is successful with probability $\frac{\alpha a}{1+\alpha a}$ for some $\alpha > 0$, in which case the quality level of its product increases by one level. The parameter α represents the effectiveness of the investment. The firm's product depreciates one quality level with probability $\delta \in (0, 1)$ independently at each time period. Thus a firm's state decreases by one with probability $\frac{\delta}{1+\alpha a}$; it increases by one with probability $\frac{(1-\delta)\alpha a}{1+\alpha a}$ and stays at the same level with probability $\frac{1-\delta+\delta\alpha a}{1+\alpha a}$.

Discussion. Our main result for this model is the following proposition.

Proposition 1. *Suppose that $\theta_1 < 1$. Then there exists an SE for the dynamic oligopoly model, and all SE possess the AME property.*

The preceding result has a natural interpretation in terms of increasing and decreasing returns to higher states. Recall that θ_1 represents how much consumers value the quality of the products, and hence if $\theta_1 < 1$, firms have strictly *decreasing* marginal returns in their payoff from increasing their own state. This implies that as their state grows, firms have less incentives to invest in improving their own state and ensures that, in equilibrium, the distribution of firms over the state space has a light tail. On the other hand, if $\theta_1 \geq 1$, then firms have an *increasing* marginal gain in their payoff from increasing their own state. Because the marginal cost of investment is constant, firms may continue to invest large amounts to improve their state even at very large states. Thus, a single firm optimization problem may not even induce a stable Markov process, and hence an SE may not exist (and the AME property may fail).

This result matches our intuition for exactly those regimes where SE work well as approximations to equilibria in finite models. In industries with decreasing returns, we expect to see a fragmented structure in the limit. By contrast, in industries with increasing returns, market concentration would likely result in the limit, i.e., a few firms capture most of the demand in the market. This is precisely where the AME property ceases to hold.

4.2 Dynamic Oligopoly Models with Positive Spillovers

Spillovers are commonly observed in industry data and could arise, for example, due to laggard firms imitating leaders' R&D activities (Griliches 1998). To introduce spillovers, we extend the previous model by modifying the transition kernel as below; we keep the state space, action space and the payoff identical to the previous model.

Transition dynamics. We follow the model of Xu (2008), in which transition dynamics depend not only on the action of the firm, but also on the state of its competitors. Formally, let $s_{-i,t}^{(m)}$ be the *spillover effect* of the population state on player i at time t , where: $s_{-i,t}^{(m)} = \sum_{y \in \mathcal{X}} \mathbf{f}_{-i,t}^{(m)}(y) h_{i,t}(y)$. Here $h_{i,t}(y)$ is a weight

function that distinguishes the effect of different states. For this example, we use $h_{i,t}(y) = \zeta(y)\mathbf{1}_{\{y > x_{i,t}\}}$ for some uniformly bounded function $\zeta(y)$. In this case, a firm is affected with spillovers only from firms that have a better state than its own, which seems natural. We define the *effective investment* of player i at time t by: $a_{i,t} + \gamma s_{-i,t}^{(m)} \triangleq e_{i,t}$. The constant γ is a spillover coefficient and it captures the effect of industry state on the state transition. A higher value of γ means a higher spillover effect. With an effective investment of e , similar to Section 4.1, a firm's state increases by one level with probability $\frac{\alpha e}{1+\alpha e}$. Finally, as before, the firm's product depreciates in quality by one level with probability $\delta \in (0, 1)$ independently at each time period.

Discussion. Since the kernel now depends on the population state f , even if $\theta_1 < 1$, the population state of an agent may grow due to large competitor states. This may lead to a scenario where the image of Φ is unbounded, because firms may exhibit unbounded growth. The following proposition provides a simple condition for existence of SE.

Proposition 2. *Suppose that $\theta_1 < 1$, and:*

$$\gamma < \frac{\delta}{(1 - \delta)\alpha \sup_y \zeta(y)} \quad (10)$$

Then there exists an SE for the dynamic oligopoly model with spillovers, and all SE possess the AME property.

Condition (10) admits a simple interpretation. This condition enforces a form of decreasing returns in the spillover effect. If the spillover effect is too large relative to depreciation—i.e., if (10) fails—then the state of a given firm has positive drift whenever other firms have large states; and in this case we expect that, for some f , the spillover effect can lead to optimal oblivious strategies that yield unbounded growth. On the other hand, when (10) holds, then this effect is controlled, and despite the presence of positive spillovers the state distribution has a light tail in equilibrium and the industry becomes fragmented in the limit.

4.3 Learning-By-Doing

Another example that commonly arises in oligopoly setting is *learning-by-doing*, where firms become more efficient by producing goods. In this section we study a model inspired by Fudenberg and Tirole (1983).

States. We let the state $x_{i,t}$ represent the cumulative experience level of a firm at time t ; this represents the knowledge accumulated through past production.

Actions. The action $a_{i,t}$ represents the firm's output (i.e., goods produced) at time t . We consider a model in which firms compete on quantity; thus firms are coupled to each other through their *actions*. As discussed in Section 3.2, such an extension can be accommodated within our framework by restricting pure actions to lie on a finite subset $S = \{0, 1, \dots, s_{\max}\}$ of the integers.⁵

⁵This amounts to discretizing the action space of production quantities. In this case, we allow for mixed strategies to ensure existence of SE (see Proposition 7). However, note that in many models of interest, under the appropriate concavity assumptions, this is not very restrictive as firms will mix between two adjacent pure actions in equilibrium.

Payoffs. At each time period, firms produce goods and compete in a market with n consumers. Let $P_n(\cdot) \geq 0$ be the inverse demand function for a market size of n . For state x , pure action s , and population state-action profile f , we can write the payoff function as $\pi_n(x, s, f, m) = sP_n\left(s + (m-1) \sum_{x', s'} s' f(x', s')\right) - C(x, s)$, where the argument of P_n is the aggregate output (from m firms) in the market. Note that f is a distribution over state-action pairs. Here, $C(x, s)$ denotes the cost of producing quantity s when the firm's experience level is x . We assume that C is nonnegative, decreasing, and convex in x ; is increasing and convex in s ; and has decreasing differences between x and s . Consider a limiting case where both the number of firms m and the market size n become large at the same rate. We assume that there exists a limiting decreasing continuous demand function P such that the limit profit function is given by $\pi(x, s, f) = sP\left(\sum_{x', s'} s' f(x', s')\right) - C(x, s)$. Note that the limiting case represents perfect competition as firms become price takers.

Transition dynamics. A firm's cumulative experience is improved as it produces more goods since it learns from the production process. On the other hand, experience capital depreciates over time due to "organizational forgetting." We assume that a firm's experience evolves independent of the experience level or the output of other firms in the market. For concreteness, we assume the transition dynamics are the same as those described in Section 4.1.

Discussion. Let $\lim_{x \rightarrow \infty} C(x, s) = \underline{C}(s)$, that is, $\underline{C}(s)$ is the cost of producing quantity s for a firm with infinite experience. Our main result for this model is the following proposition.

Proposition 3. *Let s^* be the production level that maximizes $sP(0) - \underline{C}(s)$. Suppose that for all sufficiently large x and all actions $s \in [0, s^*]$, we have $\sum_{x'} x' \mathbf{P}(x'|x, s) < x$; i.e., the state has negative drift at all such pairs (x, s) . Then there exists an SE for the learning-by-doing model, and all SE possess the AME property.*

Observe that $sp - C(x, s)$ is the single period profit to a firm when the market price is p , the firm produces quantity s , and its experience level is x . Generally speaking, because of learning, firms at low experience levels face strong incentives to increase their experience, leading them to produce beyond the single period optimal quantity. However, for firms at high experience levels, the choice of optimal quantity is driven primarily by maximization of single period profit (because $C(x, s)$ is decreasing and convex in x). The quantity s^* is an upper bound on the maximizer of single period profits, so the drift condition in the proposition ensures that at high experience levels, firms' maximization of single period profit does not continue to yield unbounded growth in the experience level.⁶ The condition requires that the transition kernel must exhibit sufficiently strong decreasing returns to scale; as long as the possible productivity gains induced by learning-by-doing are reduced at larger states, light-tailed SE will exist. However, if there are not diminishing returns to learning-by-doing, then a firm's experience level will grow without bound and hence a light-tailed SE may not exist. This is consistent with prior observations: an industry for which learning-by-doing is prevalent may naturally become concentrated over time (Dasgupta and Stiglitz 1988).

⁶For example, consider $C(x, s) = s/x$. Then s^* is the largest allowable pure action, hence, the condition requires that all actions have negative drift for sufficiently large experience levels. For a less restrictive case, consider $C(x, s) = s^2/x + s^2/c$. Then, $s^* = cP(0)/2$, so the condition requires that all actions less than or equal to $cP(0)/2$ eventually have negative drift.

4.4 Supply Chain Competition

We now consider an example of supply chain competition among firms (Cachon and Lariviere 1999), where the firms use a common resource that is sold by a single supplier. The firms only interact with each other in the sourcing stage as the goods produced are assumed to be sold in independent markets.

States. We let the state $x_{i,t}$ be the inventory of goods held by firm i at time t .

Actions. At each time period, the supplier runs an auction to sell the goods. Each firm i places a bid $a_{i,t}$ at time t ; for example, $a_{i,t}$ may denote the willingness-to-pay of the supplier, or it may be a two-dimensional bid consisting of desired payment and quantity. Since the interaction between firms is via their action profiles we again assume that the action taken by a firm lies in a finite subset S of the integer lattice.

Transition dynamics. Suppose that each firm i sees demand $d_{i,t}$ at time t ; we assume $d_{i,t}$ are i.i.d. and independent across firms, with bounded nonnegative support and positive expected value. Further, suppose that when a firm bids s and the population state-action profile is f , the firm receives an allocation $\xi(s, f)$. Then the state evolution for a firm i is given by $x_{i,t+1} = \max\{x_{i,t} - d_{i,t}, 0\} + \xi(s_{i,t}, \mathbf{f}_{-i,t}^{(m)})$. Note that ξ depends on $\mathbf{f}_{-i,t}^{(m)}$ only through the marginal distribution over actions. We make the natural assumptions that $\xi(s, f)$ is increasing in s and decreasing in f (where the set of distributions is ordered in the first order stochastic dominance sense). Thus the transition kernel captures inventory evolution in the usual way: demand consumes inventory, and procurement restocks inventory. The amount of resource procured by a firm and the price it pays depends on its own bid, as well as bids of other firms competing for the resource.

As one example of how ξ might arise, suppose that the supplier uses a *proportional allocation mechanism* (Kelly 1997). In such a mechanism, the bid s denotes the total amount a firm pays. Further, suppose the total quantity Q_m of the resource available scales with the number of firms, i.e., $Q_m = mQ$. Let $k(s|f) = \sum_x f(x, s)$ denote the fraction of agents bidding s in population state-action profile f . As $m \rightarrow \infty$, and introducing R as a small “reserve” bid that ensures the denominator is always nonzero, we obtain the following limiting proportional allocation function: $\xi(s, f) = sQ / \left(R + \sum_{s'} s' k(s'|f) \right)$. Note that this expression is increasing in s and decreasing in f .

Payoffs. A firm earns revenue for demand served, and incurs a cost both for holding inventory, as well as for procuring additional goods via restocking. We assume every firm faces an exogenous retail price ϕ . (Heterogeneity in the retail price could be captured via the description in Section 3.2.) Let h be the unit cost of holding inventory for one period and let $\Omega(s, f)$ be the procurement payment made by a firm with bid s , when the population state-action profile is f ; of course, Ω also depends on f only through $k(\cdot|f)$. In general we assume that Ω is increasing in f for each fixed s . In the proportional allocation mechanism described above, we simply have $\Omega(s, f) = s$. Since the demand is i.i.d., the single period payoff for a firm is given by the expected payoff it receives, where the expectation is over the demand uncertainty; i.e. $\pi(x, s, f) = \phi \mathbb{E}[\min\{d, x\}] - hx - \Omega(s, f)$.

Discussion. We have the following proposition.

Proposition 4. *Suppose that d has positive expected value. Then there exists an SE for the supply chain competition model with the proportional allocation mechanism, and all SE possess the AME property.*

More generally, for other choices of allocation mechanism, it can be shown that the same result holds if

d has positive expected value and the following conditions hold: (1) if ξ and Ω are uniformly bounded and appropriately continuous in f for each pure action s ; (2) $0 \in S$ and $\xi(0, f) = 0$ for all f ; and (3) bidding zero maximizes a firm's single period payoff, and this induces negative drift in the inventory level.

In this model, decreasing returns to higher states are naturally enforced because the payoff function becomes *decreasing* in the state as the state grows. Simply because holding inventory is costly, firms prefer not to become arbitrarily large. Thus in this model light tails in the population state can be guaranteed under fairly weak assumptions on the model primitives.

4.5 Consumer Learning

In this section, we analyze a model of social learning. Imagine a scenario where a group of individuals decide to consume a product (e.g., visiting a restaurant). These individuals learn from each other's experience, perhaps through product reviews or word-of-mouth (see, for example, Ching 2010).

States. We let $x_{i,t}$ be the experience level of an individual at time t .

Actions. At each time period t , an individual invests an "effort" $a_{i,t} \in [0, \bar{a}]$ in searching for a new product.

Payoffs. At each time period, an individual selects a product to consume. The quality of the product is a normally distributed random variable Q with a distribution given by $Q \sim \mathcal{N}(\gamma a, \omega(x, f))$, where $\gamma > 0$ is a constant. Thus, the average quality of the product is proportional to the amount of effort made. Furthermore, the variance of the product is dependent on both individual and population experience levels.

We assume that $\omega(x, f)$ is continuous in the population state f (in an appropriate norm, cf. Section 5). We make the natural assumption that $\omega(x, f)$ is a nonincreasing function of f and strictly decreasing in x (where the set of distributions is ordered in the first order stochastic dominance sense). This is natural as we expect that as an individual's experience increases or if she can learn from highly expert people, the randomness in choosing a product will decrease. We also assume that there exists constants σ_L, σ_H , such that $\sigma_L^2 \leq \omega(x, f) \leq \sigma_H^2$.

The individual receives a utility $U(Q)$, where $U(\cdot)$ is a nondecreasing concave function of the quality. For concreteness, we let $U(Q) = 1 - e^{-Q}$. Since at each time, the individual selects the product or the restaurant in an i.i.d. manner, the single period payoff is given by $\pi(x, a, f) = E[U(Q) \mid Q \sim \mathcal{N}(\gamma a, \omega(x, f))] - da = 1 - e^{-\gamma a + \frac{1}{2}\omega(x, f)} - da$, where d is the marginal cost of effort.

Transition dynamics. An individual's experience level is improved as she expends effort because she learns more about the quality of products. However, this experience level also depreciates over time; this depreciation is assumed to be player-specific and comes about because an individual's tastes may change over time. Thus, an individual's experience evolves (independently of the experience of others or their investments) in a stochastic manner. Several specifications for the transition kernel satisfying our assumptions can be used; for concreteness we assume that the dynamics are the same as those described in Section 4.1.

Discussion. Our main result is the following proposition.

Proposition 5. *Suppose that:*

$$d \geq \gamma e^{-\gamma c_0 + \frac{1}{2}\sigma_H^2}, \quad (11)$$

where $c_0 = \delta/(\alpha(1 - \delta))$. Then there exists an SE for the consumer learning model, and all SE possess the AME property.

Recall that $\delta \in (0, 1)$ is the probability that the experience depreciates and $\alpha > 0$ controls the probability that a player is successful in improving the experience. The right hand side is an upper bound to the marginal gain in utility due to effort, at effort level c_0 ; while the left hand side is the marginal cost of effort. Thus the condition (11) can be interpreted as a requirement that the marginal cost of effort should be sufficiently large relative to the marginal gain in utility due to effort. Otherwise, an individual's effort level when her experience is high will cause her state to continue to increase, so a light-tailed SE may not exist. Hence we see the same dichotomy as before: decreasing returns to higher states yield existence of SE and the AME property, while increasing returns may not.

5 Theory: Existence

In this section, we study the existence of *light-tailed stationary equilibria*. We recall that (μ, f) is a stationary equilibrium if and only if f is a fixed point of $\Phi(f) = \mathcal{D}(\mathcal{P}(f), f)$, such that $\mu \in \mathcal{P}(f)$ and $f \in \mathcal{D}(\mu, f)$. Thus our approach is to find conditions under which the correspondence Φ has a fixed point; in particular, we aim to apply Kakutani's fixed point theorem to Φ to find an SE.

Kakutani's fixed point theorem requires three essential pieces: (1) *compactness* of the range of Φ ; (2) *convexity* of both the domain of Φ , as well as $\Phi(f)$ for each f ; and (3) appropriate *continuity* properties of the operator Φ . It is clear, therefore, that our analysis requires topologies on both the set of possible strategies and the set of population states. For the set of oblivious strategies \mathfrak{M}_O , we use the topology of pointwise convergence.

For the set of population states, we recall that a key concept in our analysis is that of "light-tailed" population states. To formalize this notion, for the set of population states we consider a topology induced by the $1-p$ norm. Given $p > 0$, the $1-p$ -norm of a function $f : \mathcal{X} \rightarrow \mathbb{R}$ is given by $\|f\|_{1-p} = \sum_{x \in \mathcal{X}} \|x\|_p^p |f(x)|$, where $\|x\|_p$ is the usual p -norm of a vector. Let \mathfrak{F}_p be the set of all possible population states on \mathcal{X} with finite $1-p$ norm, i.e., $\mathfrak{F}_p = \{f \in \mathfrak{F} : \|f\|_{1-p} < \infty\}$. The requirement $f \in \mathfrak{F}_p$ imposes a *light-tail condition* over the population state f . The exponent p controls the weight in the tail of the population state: distributions with finite $1-p$ -norms for larger p have lighter tails. The condition essentially requires that larger states must have a small probability of occurrence under f .

We start with the following restatement of Kakutani's theorem.

Theorem 1 (Kakutani-Fan-Glicksberg). *Suppose there exists a set $\mathfrak{C} \subseteq \mathfrak{F}_p$ such that (1) \mathfrak{C} is convex and compact (in the $1-p$ norm), with $\Phi(\mathfrak{C}) \subset \mathfrak{C}$; (2) $\Phi(f)$ is convex and nonempty for every $f \in \mathfrak{C}$; and (3) Φ has a closed graph on \mathfrak{C} . Then there exists a stationary equilibrium (μ, f) with $f \in \mathfrak{C}$.*

In the remainder of this section, we find exogenous conditions on model primitives to ensure these requirements are met. We tackle them in reverse order. We first show that under an appropriate continuity condition, Φ has a closed graph. Next, we study conditions under which $\Phi(f)$ can be guaranteed to be convex. Finally, we provide conditions on model primitives under which there exists a compact, convex set

\mathcal{C} with $\Phi(\mathfrak{F}) \subset \mathcal{C}$. The conditions we provide suffice to guarantee that $\Phi(f)$ is nonempty for all $f \in \mathfrak{F}$. Taken together our conditions ensure existence of SE, as well as an additional stronger characterization: *all* SE are light-tailed, i.e., they have finite $1-p$ norm. This fact will allow us to show that every SE satisfies the AME property in the next section.

5.1 Closed Graph

In this section we develop conditions to ensure the model is appropriately “continuous.” Before stating the desired assumption, we introduce one more piece of notation. Without loss of generality, we can view the state Markov process in terms of the increments from the current state. In particular, we can write $x_{i,t+1} = x_{i,t} + \xi_{i,t}$, where $\xi_{i,t}$ is a random increment distributed according to the probability mass function $\mathbf{Q}(\cdot | x, a, f)$ defined by $\mathbf{Q}(z' | x, a, f) = \mathbf{P}(x + z' | x, a, f)$. Note that $\mathbf{Q}(z' | x, a, f)$ is positive for only those z' such that $x + z' \in \mathcal{X}$. We make the following assumptions over model primitives.

- Assumption 1** (Continuity). *1. Compact action set. The set of feasible actions for a player, denoted by \mathcal{A} , is compact.*
- 2. Bounded increments. There exists $M \geq 0$ such that, for all z with $\|z\|_\infty > M$, $\mathbf{Q}(z | x, a, f) = 0$, for all $x \in \mathcal{X}$, $a \in \mathcal{A}$, and $f \in \mathfrak{F}$.*
- 3. Growth rate bound. There exist constants K and $n \in \mathbb{Z}_+$ such that $\sup_{a \in \mathcal{A}, f \in \mathfrak{F}} |\pi(x, a, f)| \leq K(1 + \|x\|_\infty)^n$ for every $x \in \mathcal{X}$, where $\|\cdot\|_\infty$ is the sup norm.*
- 4. Payoff and kernel continuity. For each fixed $x, x' \in \mathcal{X}$ and $f \in \mathfrak{F}$, the payoff $\pi(x, a, f)$ and the kernel $\mathbf{P}(x' | x, a, f)$ are continuous in $a \in \mathcal{A}$.*
- In addition, for each fixed $x, x' \in \mathcal{X}$, the payoff $\pi(x, a, f)$ and the kernel $\mathbf{P}(x' | x, a, f)$ are jointly continuous in $a \in \mathcal{A}$ and $f \in \mathfrak{F}_p$ (where \mathfrak{F}_p is endowed with the $1-p$ norm).⁷*

The assumptions are fairly mild and are satisfied in a variety of models of interest. For example, all models in Section 4 satisfy it. The first assumption is standard. We also place a finite (but possibly large) bound on how much an agent’s state can change in one period (Assumption 1.2), an assumption that is reasonably weak. The polynomial growth rate bound on the payoff is quite weak, and serves to exclude the possibility of strategies that yield infinite expected discounted payoff.

Finally, Assumption 1.4 ensures that the impact of action on payoff and transitions is continuous. It also imposes that the payoff function and transition kernel are “smooth” functions of the population state under an appropriate norm. We note that when \mathcal{X} is finite, then $\|f\|_{1-p}$ induces the same topology as the standard Euclidean norm. However, when \mathcal{X} is infinite, the $1-p$ -norm weights larger states higher than smaller states. In many applications, other players at larger states have a greater impact on the payoff; in such settings, continuity of the payoff in f in the $1-p$ -norm naturally controls for this effect. Given a particular model, the exponent p should be chosen to ensure continuity of the payoff and transition kernel.⁸ The following

⁷Here we view $\mathbf{P}(x' | x, a, f)$ as a real valued function of a and f , for fixed x, x' ; note that since we have also assumed bounded increments, this notion of continuity is equivalent to assuming that $\mathbf{P}(\cdot | x, a, f)$ is jointly continuous in a and f , for fixed x , with respect to the topology of weak convergence on distributions over \mathcal{X} .

⁸See Section 4 and its more detailed analysis in Adlakha et al. (2011) for concrete examples. For example, in subsection 4.1 the payoff function depends on the distribution f via its θ_1 moment so it is natural to endow the set of distributions with the $1-p$ norm with $p = \theta_1$.

proposition establishes that the continuity assumptions embodied in Assumption 1 suffice to ensure that Φ has a closed graph.

Proposition 6. *Suppose that Assumption 1 holds. Then Φ has a closed graph on \mathfrak{F}_p ; i.e., the set $\{(f, g) : g \in \Phi(f)\} \subset \mathfrak{F}_p \times \mathfrak{F}_p$ is closed (where \mathfrak{F}_p is endowed with the 1-p norm).*

5.2 Convexity

Next, we develop conditions to ensure that $\Phi(f)$ is convex. We first provide a result for mixed strategies and then a result for pure strategies.

5.2.1 Mixed Strategies

We start by considering a simple model, where the action set \mathcal{A} is the simplex of *randomized actions* on a base set of finite pure actions. This setting is particularly useful when we assume players are coupled through actions (see Section 3.2). Formally, we have the following definition.

Definition 9. *An anonymous stochastic game has a finite action space if there exists a finite set S such that the following three conditions hold:*

1. \mathcal{A} consists of all probability distributions over S : $\mathcal{A} = \{a \geq 0 : \sum_s a(s) = 1\}$.
2. $\pi(x, a, f) = \sum_s a(s)\pi(x, s, f)$, where $\pi(x, s, f)$ is the payoff evaluated at state x , population state f , and pure action s .
3. $\mathbf{P}(x' | x, a, f) = \sum_s a(s)\mathbf{P}(x' | x, s, f)$, where $\mathbf{P}(x' | x, s, f)$ is the kernel evaluated at states x' and x , population state f , and pure action s .

Essentially, the preceding definition allows inclusion of *randomized* strategies in our search for SE. This model inherits Nash's original approach to establishing existence of an equilibrium for static games, where randomization induces convexity on the strategy space. We show next that in any game with finite action spaces, the set $\Phi(f)$ is always convex.

Proposition 7. *Suppose Assumption 1 holds. In any anonymous stochastic game with a finite action space, $\Phi(f)$ is convex for all $f \in \mathfrak{F}$.*

The preceding result ensures that if randomization is allowed over a set of finite actions, then the map Φ is convex-valued. We conclude by noting that another simplification is possible when working with a finite action space. In particular, it is straightforward to show that *if Assumption 1 holds for the payoff and transition kernel over all pure actions, then it also holds for the payoff and transition kernel over all mixed actions*; Proposition 6 follows similarly. The proof follows in an easy manner using the linearity of the payoff and transition kernel. This is a valuable insight, since in applications it simplifies the complexity of checking the model assumptions necessary to guarantee existence of an equilibrium. We discuss a similar point in Section 5.3.

5.2.2 Pure Strategies

In contrast to the preceding section, many relevant applications typically require existence of equilibria in pure strategies. For such examples, we employ an approach based on the following proposition.

Proposition 8. *Suppose that $\mathcal{P}(f)$ is a singleton for all $f \in \mathfrak{F}$. Then $\Phi(f)$ is convex for all $f \in \mathfrak{F}$.*

The proof is straightforward: $\mathcal{D}(\mu, f)$ is convex-valued for any fixed μ and f , since the set of invariant distributions for the kernel defined by μ and f are identified by a collection of linear equations. Thus if $\mathcal{P}(f)$ is a singleton, then $\Phi(f) = \mathcal{D}(\mathcal{P}(f), f)$ will be convex.

We now provide two different assumptions over model primitives that guarantee that $\mathcal{P}(f)$ is a singleton, for all $f \in \mathfrak{F}$. The first assumption is a condition introduced by Doraszelski and Satterthwaite (2010) and is described in detail there. The assumption has found wide application in dynamic oligopoly models.

Assumption 2. 1. *The state space is scalar, i.e., $\mathcal{X} \subseteq \mathbb{Z}_+$, and the action space \mathcal{A} is a compact interval of the real numbers.*

2. *The payoff $\pi(x, a, f)$ is strictly decreasing and concave in a for fixed x and f .*

3. *For all $f \in \mathfrak{F}$, the transition kernel \mathbf{P} is unique investment choice (UIC) admissible: there exist functions g_1, g_2, g_3 such that $\mathbf{P}(x' | x, a, f) = g_1(x, a, f)g_2(x', x, f) + g_3(x', x, f), \forall x', x, a, f$, where $g_1(x, a, f)$ is strictly increasing and strictly concave in a .*

The preceding conditions ensure that for all population states f and initial states x , and all continuation value functions, the maximization problem in the right hand side of Bellman's equation (cf. (12) in the Appendix) is strictly concave, or that the unique maximizer is a corner solution.

The previous assumption requires a single-dimensional state space and action space. Our next assumption imposes a different set of conditions over the payoff and the transition kernel, and allows for multi-dimensional state and action spaces. Before providing our second condition, we require some additional terminology. Let $S \subset \mathbb{R}^n$. We say that a function $g : S \rightarrow \mathbb{R}$ is *nondecreasing* if $g(x') \geq g(x)$ whenever $x' \geq x$ (where we write $x' \geq x$ if x' is at least as large as x in every component). We say g is *strictly increasing* if the inequality is strict. Let \mathbf{P}_θ be a family of probability distributions on \mathcal{X} indexed by $\theta \in S$. Given a nondecreasing function $u : \mathcal{X} \rightarrow \mathbb{R}$, define $\mathbb{E}_\theta[u] = \sum_x u(x)\mathbf{P}_\theta(x)$. We say that \mathbf{P}_θ is *stochastically nondecreasing* in the parameter θ , if $\mathbb{E}_\theta[u]$ is nondecreasing in θ for every nondecreasing function u . Similarly, we say that \mathbf{P}_θ is *stochastically concave* in the parameter θ if $\mathbb{E}_\theta[u]$ is a concave function of θ for every nondecreasing function u . We say that \mathbf{P}_θ is *strictly stochastically concave* if, in addition, $\mathbb{E}_\theta[u]$ is strictly concave for every strictly increasing function u . We have the following assumption.

Assumption 3. 1. *The action set \mathcal{A} is convex.*

2. *The payoff $\pi(x, a, f)$ is strictly increasing in x for fixed a and f , and the kernel $\mathbf{P}(\cdot | x, a, f)$ is stochastically nondecreasing in x for fixed a and f .*

3. *The payoff is concave in a for fixed x and f , and the kernel is stochastically concave in a for fixed x and f , with at least one of the two strictly concave in a .*

The following result shows the preceding conditions on model primitives ensure the optimal oblivious strategy is unique.

Proposition 9. *Suppose Assumption 1 holds, and that at least one of Assumptions 2 or 3 holds. Then $\mathcal{P}(f)$ is a singleton, and thus $\Phi(f)$ is convex for all $f \in \mathfrak{F}$.*

5.3 Compactness

In this section, we provide conditions under which we can guarantee the existence of a compact, convex, nonempty set \mathcal{C} such that $\Phi(\mathfrak{F}) \subset \mathcal{C}$. The assumptions we make are closely related to those needed to ensure that $\Phi(f)$ is nonempty. To see the relationship between these results, observe that in Lemma 2 in the Appendix, we show that under Assumption 1 an optimal oblivious strategy always exists for any $f \in \mathfrak{F}$. Thus to ensure that $\Phi(f)$ is nonempty, it suffices to show that there exists at least one strategy that possesses an invariant distribution. Our approach to demonstrating existence of an invariant distribution is to use a Foster-Lyapunov argument. Intuitively, this criterion checks whether the process that describes the evolution of an agent eventually has “negative” drift and in this way controls for the growth of the agent’s state. This same argument also allows us to bound the moments of the invariant distribution—precisely what is needed to find the desired set \mathcal{C} that is compact in the $1-p$ norm.

One simple condition under which $\Phi(f)$ is nonempty is that the state space is finite; any Markov chain on a finite state space possesses at least one positive recurrent class. In this case the entire set \mathfrak{F} is compact in the $1-p$ norm. Thus we have the following result.

Proposition 10. *Suppose Assumption 1 holds, and that the state space \mathcal{X} is finite. Then $\Phi(f)$ is nonempty for all $f \in \mathfrak{F}$, and \mathfrak{F} is compact in the $1-p$ norm.*

We now turn our attention to the setting where the state space may be unbounded; for notational simplicity, in the remainder of the section we assume $\mathcal{X} = \mathbb{Z}_+^d$. In this case, we must make additional assumptions to control for the agent’s growth; these assumptions ensure the optimal strategy does not allow the state to become transient, and also allows us to bound moments of the invariant distribution of any optimal oblivious strategy.

In the sequel we restrict attention to multiplicatively separable transition kernels, as defined below.

Definition 10. *The transition kernel is multiplicatively separable if there exist transition kernels $\mathbf{P}_1, \dots, \mathbf{P}_d$ such that for all $x, x' \in \mathcal{X}, a \in \mathcal{A}, f \in \mathfrak{F}$, there holds $\mathbf{P}(x'|x, a, f) = \prod_{\ell=1}^d \mathbf{P}_\ell(x'_\ell|x, a, f)$. In this case we let $\mathbf{Q}_1, \dots, \mathbf{Q}_\ell$ be the coordinatewise increment transition kernels; i.e., $\mathbf{Q}_\ell(z_\ell|x, a, f) = \mathbf{P}_\ell(x_\ell + z_\ell|x, a, f)$, for z such that $x + z \in \mathcal{X}$.*

This is a natural class of dynamics in models with multidimensional state spaces. We note that if \mathcal{X} is one-dimensional, the definition is vacuous. We introduce the following assumption.

Assumption 4.

1. For all $\Delta \in \mathbb{Z}_+^d$, there holds $\limsup_{\|x\|_\infty \rightarrow \infty} \sup_{a \in \mathcal{A}, f \in \mathfrak{F}} \pi(x + \Delta, a, f) - \pi(x, a, f) \leq 0$.
2. The transition kernel \mathbf{P} is multiplicatively separable.
3. For $\ell = 1, \dots, d$, $\mathbf{P}_\ell(\cdot|x, a, f)$ is stochastically nondecreasing in $x \in \mathcal{X}$ and $a \in \mathcal{A}$ for fixed $f \in \mathfrak{F}$.
4. For $\ell = 1, \dots, d$, and for each $a \in \mathcal{A}$ and $f \in \mathfrak{F}$, $\mathbf{Q}_\ell(\cdot|x, a, f)$ is stochastically nonincreasing in $x \in \mathcal{X}$. Further, for all $x \in \mathcal{X}$, $\sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell|x, a, f)$ is continuous in a .

5. There exists a compact set $\mathcal{A}' \subset \mathcal{A}$, a constant K' , and a continuous, strictly increasing function $\kappa : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\kappa(0) = 0$, such that:

- (a) For all $x \in \mathcal{X}$, $f \in \mathfrak{F}$, $a \notin \mathcal{A}'$, there exists $a' \in \mathcal{A}'$ with $a' \leq a$, such that $\pi(x, a', f) - \pi(x, a, f) \geq \kappa(\|a' - a\|_\infty)$.
- (b) For all ℓ , and all x' such that $x'_\ell \geq K'$, $\sup_{a' \in \mathcal{A}'} \sup_{f \in \mathfrak{F}} \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x', a', f) < 0$.

Some of the previous conditions are natural, while others impose a type of “decreasing returns to higher states.” First, we discuss the former. Multiplicative separability (Assumption 4.2) is natural. The first part of Assumption 4.3 is also fairly weak. The transition kernel is stochastically nondecreasing in state in models for which the state is persistent, in the sense that a larger state today increases the chances of being at a larger state tomorrow. The transition kernel is stochastically nondecreasing in action in models where larger actions take agents to larger states.

Assumption 4.1, 4.4, and 4.5 impose a form of “decreasing returns to higher states” in the model. In particular, Assumption 4.1 ensures the marginal gain in payoff by increasing one’s state becomes nonpositive as the state grows large. This assumption is used to show that for large enough states agents effectively become myopic; increasing the state further does not provide additional gains. Assumption 4.5 then implies that as the state grows large, optimal actions produce negative drift inducing a “light-tail” on any invariant distribution of the resulting optimal oblivious strategy. The set \mathcal{A}' can be understood as (essentially) the set of actions that maximize the single period payoff function. Assumption 4.5 is often natural because in many models of interest increasing the state beyond a certain point is costly and requires dynamic incentives; agents will take larger actions that induce positive drift only if they consider the future benefits of doing so.

The first part of Assumption 4.4 imposes a form of decreasing returns in the transition kernel. The second part of Assumption 4.4 will hold if, for example, the transition kernel is coordinatewise stochastically nonincreasing in $f \in \mathfrak{F}$ (with respect to the first order stochastic dominance ordering) and Assumption 1 holds. In this case $\sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x, a, f) = \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x, a, \underline{f})$, where \underline{f} is the distribution that places all its mass at state 0.

Much of the difficulty in the proof of the result lies in ensuring that the tail of any invariant distribution obtained from an optimal oblivious strategy is uniformly light over the image of Φ . The fact that Assumptions 4.1, 4.4, and 4.5 are uniform over f are crucial for this purpose.

Under the preceding assumptions we have the following result.

Proposition 11. *Suppose $\mathcal{X} = \mathbb{Z}_+^d$, and Assumptions 1 and 4 hold. Then $\Phi(f)$ is nonempty for all $f \in \mathfrak{F}$, and there exists a compact, convex, nonempty set \mathcal{C} such that $\Phi(\mathfrak{F}) \subset \mathcal{C}$.*

Note that the preceding result ensures $\Phi(f) \subset \mathcal{C}$ for all $f \in \mathfrak{F}$.

We conclude this section with a brief comment regarding finite action spaces, cf. Definition 9. The key observation we make is that *if Assumption 4 holds with respect to the pure actions—i.e., with \mathcal{A} replaced by S —then the same result as Proposition 11 holds for mixed actions.* A nearly identical argument applies to establish the result.

6 Theory: Approximation

In this section we show that under the assumptions of the preceding section, any SE (μ, f) possesses the AME property. We emphasize that the AME property is essentially a continuity property in the population state f . Under reasonable assumptions, we show that the time t population state in the system with m players, $f_{-i,t}^{(m)}$, approaches the deterministic population state f in an appropriate sense almost surely for all t as $m \rightarrow \infty$; in particular, this type of uniform law of large numbers will hold as long as f has tails that are sufficiently light. If $f_{-i,t}^{(m)}$ approaches f almost surely, then informally, if the payoff satisfies an appropriate continuity property in f , we should expect the AME property to hold. The remainder of the section is devoted to formalizing this argument.

Theorem 2 (AME). *Suppose Assumption 1 holds. Let (μ, f) be a stationary equilibrium with $f \in \mathfrak{F}_p$. Then the AME property holds for (μ, f) .*

Observe that Assumption 1 is also required for the existence of SE that satisfy $f \in \mathfrak{F}_p$. In this sense, under our assumptions, *the AME property is a direct consequence of existence*. This relationship between existence and the AME property is a significant insight of our work.

The proof of the AME property exploits the fact that the 1- p -norm of f must be finite (since $f \in \mathfrak{F}_p$) to show that $\|f_{-i,t}^{(m)} - f\|_{1-p} \rightarrow 0$ almost surely as $m \rightarrow \infty$; i.e., the population state of other players approaches f almost surely under an appropriate norm. Continuity of the payoff π in f , together with the growth rate bounds in Assumption 1, yields the desired result.

In practice, the light-tail condition—i.e., the requirement that $f \in \mathfrak{F}_p$ —ensures that an agent’s state rarely becomes too large under the invariant distribution f associated with the dynamics (6). Weintraub et al. (2010) provide a similar result in a dynamic industry model with entry and exit. Our result, on the other hand, is more general in terms of the definition of the AME property, as well as the payoff functions and transition kernels considered. In particular, we allow for dependence of the transition kernel on the population state. This necessitates a significantly different proof technique, since agents’ states do not evolve independently in general. We note that the light-tail condition is consequential, as it is possible to construct examples for which stationary equilibria exist, but $f \notin \mathfrak{F}_p$ and the AME property does not hold (Weintraub et al. 2010).

We conclude by noting that in many models of interest it is more reasonable to assume that the payoff function explicitly depends on the number of agents. To study these environments, we consider a sequence of payoff functions indexed by the number of agents, $\pi_m(x, a, f)$. Here, the profit function π is a *limit*: $\lim_{m \rightarrow \infty} \pi_m(x, a, f) = \pi(x, a, f)$. (See Section 4 for concrete examples.) In this case, if the number of players is m , the actual expected net present value is defined with π_m ; hence, the payoff function in the AME property depends on m . Adlakha et al. (2011) shows that under a strengthening of Assumption 1, Theorem 2 can be generalized to this setting.⁹

⁹However, note that the SE (μ, f) is still fixed and is computed with the limit payoff function π . Alternatively, it is possible to define an “oblivious equilibrium” (OE) for each finite model like in Weintraub et al. (2008). We conjecture that a variant of the assumptions that guarantee existence of SE in Section 5, but that applies uniformly over all finite models, would guarantee that the sequence of OE satisfies the AME property. For clarity of presentation, we chose to work with the SE of the limit model directly. Moreover, we believe that the existence result for the limit model that we provide is important, because even though OE might exist

7 Discussion

This paper considered stationary equilibria of dynamic games with many players. Our main results provide a parsimonious set of assumptions on the model primitives which ensure that SE exist in a large variety of games. We also showed that the same set of assumptions ensure that SE is a good approximation to MPE in large finite games.

Our results can be succinctly summarized by the following corollary that imposes conditions over model primitives to guarantee the existence of light-tailed SE and the AME property.

Corollary 1. *Suppose that (1) Assumption 1 holds; (2) either the game has a finite action space, or Assumption 2 holds, or Assumption 3 holds; and (3) either the state space \mathcal{X} is finite, or $\mathcal{X} = \mathbb{Z}_+^d$ and Assumption 4 holds. Then, there exists a SE, and every SE (μ, f) has $f \in \mathfrak{F}_p$. Furthermore, the AME property holds for all SE.*

Our theoretical analysis provides a unifying thread: *all* the propositions for the examples in Section 4 are consequences of Corollary 1. In the Corollary, condition (1) is used to guarantee continuity properties; condition (2) is used to ensure the convexity of the image of Φ ; and condition (3) is used to ensure the existence of a compact subset $\mathcal{C} \subset \mathfrak{F}_p$ such that $\Phi(\mathfrak{F}) \subset \mathcal{C}$.

We now briefly sketch how these conditions are verified for the propositions in Section 4.

Continuity properties. In all the examples, verifying Assumption 1 is fairly straightforward, through an appropriate choice of 1- p -norm. For example, for the dynamic oligopoly model in Section 4.1, note that the payoff function depends on the distribution f via its θ_1 moment, and hence we endow the set of distributions with the topology induced by the 1- p norm with $p = \theta_1$. Since the payoff is continuous and nonincreasing in the θ_1 moment of f , and the transition kernel is independent of f , it is straightforward to check that Assumption 1 holds. Verification for the other examples proceeds in a similar manner.

Convexity of the image of Φ . This is straightforward to verify in each example; e.g., the examples in Sections 4.1 and 4.2 satisfy Assumption 2; the example in Section 4.5 satisfies Assumption 3; and the examples in Sections 4.3 and 4.4 have finite action spaces.

Compactness and “decreasing returns.” The main technical difficulty arises in ensuring existence of a compact set that contains the image of Φ . Indeed, the conditions in the propositions of Section 4 are in fact used to ensure Assumption 4 holds. As discussed in Section 4, these conditions impose a form of “decreasing returns to higher states.”

To illustrate this point consider the dynamic oligopoly model in Section 4.1. It is straightforward to check that Assumptions 4.2 to 4.4 hold; we omit the details. Assumption 4.5 holds because investment is costly and there is depreciation; in particular, it suffices to set $\mathcal{A}' = \{0\}$. Thus the central condition to check in this model is Assumption 4.1. This assumption holds *if and only if* $\theta_1 < 1$: in this case, $\sup_{a,f} \pi(x + \Delta, a, f) - \pi(x, a, f) \rightarrow 0$ as $x \rightarrow \infty$ for all $\Delta > 0$. Thus the result in Proposition 1 follows.

The conditions in the other propositions exhibit similar connections to Assumption 4. The analysis in the model with spillovers (Section 4.2) is similar to the previous paragraph, but the additional condition in Proposition 2 ensures that for zero investment, the drift eventually becomes negative even in the presence of

under mild conditions for each finite model, SE in the limit model may fail to exist (see Section 4).

spillovers—thus guaranteeing Assumption 4.5. Analogously, the condition in Proposition 3 for the learning-by-doing model (Section 4.3) implies Assumption 4.5. We refer the reader to Adlakha et al. (2011) for further details.

References

- Acemoglu, D., M. K. Jensen. 2009. Aggregate comparative statics SSRN Working Paper.
- Adlakha, S., R. Johari. 2010. Mean field equilibrium in dynamic games with complementarities Submitted.
- Adlakha, S., R. Johari, G. Weintraub, A. Goldsmith. 2008. Oblivious equilibrium for large-scale stochastic games with unbounded costs. *Proceedings of the IEEE Conference on Decision and Control*. 5531–5538.
- Adlakha, S., R. Johari, G.Y. Weintraub. 2011. Equilibria of dynamic games with many players: Existence, approximation, and market structure. Tech. rep.
- Adlakha, S., R. Johari, G.Y. Weintraub, A. Goldsmith. 2010. On oblivious equilibrium in large population stochastic games. *Proceedings of the IEEE Conference on Decision and Control*. 3117–3124.
- Besanko, D., M. K. Perry, R. H. Spady. 1990. The logit model of monopolistic competition: Brand diversity. *The Journal of Industrial Economics* **38**(4) 397 – 415.
- Bodoh-Creed, A. 2010. The simple behavior of large mechanisms. *Submitted* .
- Cachon, Gérard P., M. A. Lariviere. 1999. Capacity choice and allocation: Strategic behavior and supply chain performance. *Management Science* **45**(8) 1091–1108.
- Caplin, A., B. Nalebuff. 1991. Aggregation and imperfect competition - on the existence of equilibrium. *Econometrica* **59**(1) 25 – 59.
- Ching, A.T. 2010. Consumer learning and heterogeneity: Dynamics of demand for prescription drugs after patent expiration. *International Journal of Industrial Organization* .
- Dasgupta, P., J. Stiglitz. 1988. Learning-by-doing, market structure and industrial and trade policies. *Oxford Economic Papers* **40** 246–268.
- Doraszelski, U., A. Pakes. 2007. A framework for applied dynamic analysis in IO. *Handbook of Industrial Organization, Volume 3* .
- Doraszelski, U., M. Satterthwaite. 2010. Computable Markov-perfect industry dynamics. *RAND Journal of Economics* **41**(2) 215 – 243.
- Escobar, J. 2008. Existence of pure and behavior strategy stationary markov equilibrium in dynamic stochastic games. Working paper, University of Chile.
- Fudenberg, D., J. Tirole. 1983. Learning-by-doing and market performance. *Bell Journal of Economics* **14**(14) 552–530.
- Fudenberg, D., J. Tirole. 1991. *Game Theory*. The MIT Press.
- Glynn, P. 2004. Distributed algorithms for wireless networks Presented at the Conference on Stochastic Networks, Montréal, Québec, Canada.
- Glynn, P.W., A. Zeevi. 2006. Bounding Stationary Expectations of Markov Processes. *et al., Markov processes and related topics: A Festschrift for Thomas G. Kurtz. Selected papers of the conference, Madison, WI, USA, July*. 10–13.
- Griliches, Z. 1998. *R&D and Productivity: The Econometric Evidence*. The University of Chicago Press.

- Hajek, B. 1982. Hitting-time and occupation-time bounds implied by drift analysis with applications. *Advances in Applied probability* **14**(3) 502–525.
- Hopenhayn, H. A. 1992. Entry, exit, and firm dynamics in long run equilibrium. *Econometrica* **60**(5) 1127–1150.
- Huang, M., P. E. Caines, R. P. Malhamé. 2007. Large-population cost-coupled LQG problems with nonuniform agents: Individual-mass behavior and decentralized ϵ -Nash equilibria. *IEEE Transactions on Automatic Control* **52**(9) 1560–1571.
- Jovanovic, B., R. W. Rosenthal. 1988. Anonymous sequential games. *Journal of Mathematical Economics* **17** 77–87.
- Kelly, Frank P. 1997. Charging and rate control for elastic traffic. *European Transactions on Telecommunications* **8** 33–37.
- Klette, T. J., S. Kortum. 2004. Innovating firms and aggregate innovation. *Journal of Political Economy* **112**(5) 986 – 1018.
- Lasry, J. M., P. L. Lions. 2007. Mean field games. *Japanese Journal of Mathematics* **2** 229–260.
- Luttmer, E. G. J. 2007. Selection, growth, and the size distribution of firms. *Quarterly Journal of Economics* **122**(3) 1103 – 1144.
- Maskin, E., J. Tirole. 1988. A theory of dynamic oligopoly, I and II. *Econometrica* **56**(3) 549 – 570.
- Melitz, M. J. 2003. The impact of trade on intra-industry reallocations and aggregate industry productivity. *Econometrica* **71**(6) 1695 – 1725.
- Mertens, J.-F., S. Sorin, R. Zamir. 1994. Repeated games CORE Discussion Papers 9420, 9421, 9422.
- Meyn, S. P., R. L. Tweedie. 1993. *Markov Chains and Stochastic Stability*. Springer-Verlag.
- Pakes, A., P. McGuire. 1994. Computing Markov-perfect Nash equilibria: Numerical implications of a dynamic differentiated product model. *RAND Journal of Economics* **25**(4) 555–589.
- Shapley, L. S. 1953. Stochastic games. *Proceeding of the National Academy of Sciences* **39** 1095–1100.
- Sutton, J. 1991. *Sunk Costs and Market Structure*. 1st ed. MIT Press.
- Sznitman, A.S. 1991. Topics in propagation of chaos. *Ecole d'Eté de Probabilités de Saint-Flour XIX1989* 165–251.
- Tembine, H., J.-Y. Le Boudec, R. El-Azouzi, E. Altman. 2009. Mean field asymptotics of Markov decision evolutionary games and teams. *Proceedings of GameNets '09*. 140–150.
- Weintraub, G. Y., C. L. Benkard, B. Van Roy. 2008. Markov perfect industry dynamics with many firms. *Econometrica* **76**(6) 1375–1411.
- Weintraub, G. Y., C. L. Benkard, B. Van Roy. 2010. Industry dynamics: Foundations for models with an infinite number of firms, Submitted.
- Xu, Y. 2008. A structural empirical model of R&D, firm heterogeneity, and industry evolution. Working paper, NYU University.
- Yin, H., P. G. Mehta, S. P. Meyn, U. V. Shanbhag. 2010. Synchronization of coupled oscillators is a game. *IEEE CDC*.

A Existence and AME: Preliminary Lemmas

We begin with the following lemma, which follows from the growth rate bound and bounded increments in Assumption 1; see Adlakha et al. (2011).

Lemma 1. *Suppose Assumption 1 holds. Let $x_0 = x$. Let $a_t \in \mathcal{A}$ be any sequence of (possibly history dependent) actions, and let $f_t \in \mathfrak{F}$ be any sequence of (possibly history dependent) population states. Let x_t be the state sequence generated, i.e., $x_t \sim \mathbf{P}(\cdot \mid x_{t-1}, a_{t-1}, f_{t-1})$. Then for all $T \geq 0$, there exists $C(x, T) < \infty$ such that $\mathbb{E} \left[\sum_{t=T}^{\infty} \beta^t |\pi(x_t, a_t, f_t)| \mid x_0 = x \right] \leq C(x, T)$. Further, $C(x, T) \rightarrow 0$ as $T \rightarrow \infty$.*

We now show that the Bellman equation holds for the dynamic program solved by a single agent given a population state f . Given our unbounded state space, our proof involves the use of a *weighted sup norm*, defined as follows. For each $x \in \mathcal{X}$, let $W(x) = (1 + \|x\|_{\infty})^n$. For a function $F : \mathcal{X} \rightarrow \mathbb{R}$, define:

$$\|F\|_{W-\infty} = \sup_{x \in \mathcal{X}} \left| \frac{F(x)}{W(x)} \right|.$$

This is the weighted sup norm with weight function W . We let $B(\mathcal{X})$ denote the set of all functions $F : \mathcal{X} \rightarrow \mathbb{R}$ such that $\|F\|_{W-\infty} < \infty$.

Let T_f denote the dynamic programming operator with population state f : given a function $F : \mathcal{X} \rightarrow \mathbb{R}$, we have $(T_f F)(x) = \sup_{a \in \mathcal{A}} \left\{ \pi(x, a, f) + \beta \sum_{x' \in \mathcal{X}} F(x') \mathbf{P}(x' \mid x, a, f) \right\}$. We define T_f^k to be the composition of the mapping T_f with itself k times. The following lemma applies standard dynamic programming arguments; see Adlakha et al. (2011) for details.

Lemma 2. *Suppose Assumption 1 holds. For all $f \in \mathfrak{F}$, if $F \in B(\mathcal{X})$ then $T_f F \in B(\mathcal{X})$. Further, there exist k, ρ independent of f with $0 < \rho < 1$ such that T_f is a k -stage ρ -contraction on $B(\mathcal{X})$; i.e., if $F, F' \in B(\mathcal{X})$, then for all f : $\left\| T_f^k F - T_f^k F' \right\|_{W-\infty} \leq \rho \|F - F'\|_{W-\infty}$.*

In particular, value iteration converges to $\tilde{V}^(\cdot \mid f) \in B(\mathcal{X})$ from any initial value function in $B(\mathcal{X})$, and for all $f \in \mathfrak{F}$ and $x \in \mathcal{X}$, the Bellman equation holds:*

$$\tilde{V}^*(x \mid f) = \sup_{a \in \mathcal{A}} \left\{ \pi(x, a, f) + \beta \sum_{x' \in \mathcal{X}} \tilde{V}^*(x' \mid f) \mathbf{P}(x' \mid x, a, f) \right\}. \quad (12)$$

Further, $\tilde{V}^(x \mid f)$ is continuous in $f \in \mathfrak{F}_p$.*

Finally, there exists at least one optimal oblivious strategy among all (possibly history-dependent, possibly randomized) strategies; i.e., $\mathcal{P}(f)$ is nonempty for all $f \in \mathfrak{F}$. An oblivious strategy $\mu \in \mathfrak{M}_O$ is optimal given f if and only if $\mu(x)$ achieves the maximum on the right hand side of (12) for every $x \in \mathcal{X}$.

B Existence: Proof

The proof of Proposition 6 involves applying the continuity properties in Assumption 1 and Lemma 2 to establish Φ has a closed graph. The proofs of Proposition 7 and Proposition 9 use straightforward convexity arguments. We omit the details of these arguments and instead refer the reader to Adlakha et al. (2011).

In this section we focus, therefore, on compactness of the image of Φ . Throughout this section we suppose $\mathcal{X} = \mathbb{Z}_+^d$ and that Assumptions 1 and 4 are in effect.

Lemma 3. Given $x' \geq x$, $x, x' \in \mathcal{X}$, $a \in \mathcal{A}$, and $f \in \mathfrak{F}$, there exists a probability space with random variables $\xi' \sim \mathbf{Q}(\cdot | x', a, f)$, $\xi \sim \mathbf{Q}(\cdot | x, a, f)$, such that $\xi' \leq \xi$ almost surely, and $x' + \xi' \geq x + \xi$ almost surely.

Proof. The proof uses a coupling argument. Let U be a uniform random variable on $[0, 1]$. Let F_ℓ (resp., F'_ℓ) be the cumulative distribution function of $\mathbf{Q}_\ell(\cdot | x, a, f)$ (resp., $\mathbf{Q}_\ell(\cdot | x', a, f)$), and let G_ℓ (resp., G'_ℓ) be the cumulative distribution function of $\mathbf{P}_\ell(\cdot | x, a, f)$ (resp., $\mathbf{P}_\ell(\cdot | x', a, f)$). By Assumption 4, $\mathbf{P}_\ell(\cdot | x, a, f)$ is stochastically nondecreasing in x , and $\mathbf{Q}_\ell(\cdot | x, a, f)$ is stochastically nonincreasing in x . Thus for all z , $F_\ell(z) \leq F'_\ell(z)$, but for all y , $G_\ell(y) \geq G'_\ell(y)$; further, $G_\ell(y) = F_\ell(y - x_\ell)$ (and $G'_\ell(y) = F'_\ell(y - x'_\ell)$). Let $\xi_\ell = \inf\{z_\ell : F_\ell(z_\ell) \geq U\}$, and let $\xi'_\ell = \inf\{z_\ell : F'_\ell(z_\ell) \geq U\}$. Then $\xi_\ell \geq \xi'_\ell$ for all ℓ , i.e., $\xi \geq \xi'$. Rewriting the definitions, we also have $x_\ell + \xi_\ell = \inf\{y_\ell : F_\ell(y_\ell - x_\ell) \geq U\}$, and $x'_\ell + \xi'_\ell = \inf\{y_\ell : F'_\ell(y_\ell - x'_\ell) \geq U\}$, i.e., $x_\ell + \xi_\ell = \inf\{y_\ell : G_\ell(y_\ell) \geq U\}$, and $x'_\ell + \xi'_\ell = \inf\{y_\ell : G'_\ell(y_\ell) \geq U\}$. Thus $x_\ell + \xi_\ell \leq x'_\ell + \xi'_\ell$ for all ℓ , i.e., $x' + \xi' \geq x + \xi$, as required. \blacksquare

Given a set S define $\rho_\infty(x, S) = \inf_{y \in S} \|x - y\|_\infty$. Thus ρ_∞ gives the ∞ -norm distance to a set. We have the following lemma.

Lemma 4. As $\|x\|_\infty \rightarrow \infty$, $\sup_{f \in \mathfrak{F}} \sup_{\mu \in \mathcal{P}(f)} \rho_\infty(\mu(x), \mathcal{A}') \rightarrow 0$.

Proof. Suppose the statement of the lemma fails; then there exists $r > 0$ and a sequence $f_n \in \mathfrak{F}$, $\mu_n \in \mathcal{P}(f_n)$, and x_n (where $\|x_n\|_\infty \rightarrow \infty$) such that $\rho_\infty(\mu_n(x_n), \mathcal{A}') \geq r$ for all n . We use this fact to construct a profitable deviation from the policy μ_n , for sufficiently large n .

Observe that by Assumption 4, there must exist $a'_n \in \mathcal{A}'$ with $a'_n \leq \mu_n(x_n)$, such that:

$$\pi(x_n, a'_n, f_n) - \pi(x_n, \mu_n(x_n), f_n) \geq \kappa(\|a'_n - \mu_n(x_n)\|_\infty) \geq \kappa(r) > 0,$$

where the last inequality follows since κ is strictly increasing with $\kappa(0) = 0$. Importantly, note the bound on the right hand side is a constant, independent of n .

Let $x_{0,n} = x_n$, and let $x_{t,n}$ and $a_{t,n}$ denote the state and action sequence realized under μ_n , starting from $x_{0,n}$, under the kernel $\mathbf{P}(\cdot | x, a, f_n)$. We consider a deviation from μ_n , where at time 0, instead of playing $a_{0,n} = \mu_n(x_n)$, the agent plays $a'_{0,n} = a'_n$; and then at all times in the future, the agent follows the same actions as the original sequence, i.e., $a'_{t,n} = a_{t,n}$. Let $x'_{t,n}$ denote the resulting state sequence.

Since the kernel is stochastically nondecreasing in a , and $a'_n \leq a_n$, it follows that there exists a common probability space together with increments $\xi_{0,n}, \xi'_{0,n}$, such that $\xi_{0,n} \sim \mathbf{Q}(\cdot | x_n, a_n, f_n)$, $\xi'_{0,n} \sim \mathbf{Q}(\cdot | x_n, a'_n, f_n)$, and $\xi'_{0,n} \leq \xi_{0,n}$ almost surely. Thus we can couple together $x_{1,n}$ and $x'_{1,n}$, by letting $x_{1,n} = x_n + \xi_{0,n}$, and $x'_{1,n} = x_n + \xi'_{0,n}$. In particular, observe that with these definitions we have $x_{1,n} \geq x'_{1,n}$. Let $\Delta_n = \xi_{0,n} - \xi'_{0,n} \geq 0$. Note that $\|\Delta_n\|_\infty \leq 2M$, by Assumption 1 (bounded increments).

Next, it follows from Lemma 3 that there exists a probability space with random variables $\xi_{1,n}, \xi'_{1,n}$ such that $\xi_{1,n} \sim \mathbf{Q}(\cdot | x_{1,n}, a_{1,n}, f_n)$ and $\xi'_{1,n} \sim \mathbf{Q}(\cdot | x'_{1,n}, a_{1,n}, f_n)$, $\xi_{1,n} \leq \xi'_{1,n}$ almost surely, and yet $x_{1,n} + \xi_{1,n} \geq x'_{1,n} + \xi'_{1,n}$ almost surely. Thus we can couple together $x_{2,n}$ and $x'_{2,n}$, by letting $x_{2,n} = x_{1,n} + \xi_{1,n}$, and let $x'_{2,n} = x'_{1,n} + \xi'_{1,n}$. Proceeding inductively, it can be shown that there exists a joint probability measure under which $0 \leq x_{t,n} - x'_{t,n} \leq \Delta_n$, almost surely, for all $t \geq 1$ (where the inequalities are interpreted coordinatewise); this follows by a standard application of the Kolmogorov extension theorem.

We now compare the payoffs obtained under these two sequences. We have:

$$\begin{aligned} \mathbb{E} \left[\sum_t \beta^t (\pi(x_{t,n}, a_{t,n}, f_n) - \pi(x'_{t,n}, a'_{t,n}, f_n)) \right] &= \pi(x_n, \mu_n(x_n), f_n) - \pi(x_n, a'_n, f_n) \\ &\quad + \mathbb{E} \left[\sum_{t \geq 1} \beta^t (\pi(x_{t,n}, a_{t,n}, f_n) - \pi(x'_{t,n}, a_{t,n}, f_n)) \right] \\ &\leq -\kappa(r) + \mathbb{E} \left[\sum_{t \geq 1} \beta^t \sup_{\delta \geq 0: \|\delta\|_\infty \leq 2M} \sup_{a, f} (\pi(x_{t,n}, a, f) - \pi(x_{t,n} - \delta, a, f)) \right]. \end{aligned}$$

Since increments are bounded (Assumption 4), in time t , the maximum distance the state could have moved in each coordinate from the initial state x is bounded by tM . Thus if $x_{0,n} = x_n$, then:

$$\sup_{\delta \geq 0: \|\delta\|_\infty \leq 2M} \sup_{a, f} (\pi(x_{t,n}, a, f) - \pi(x_{t,n} - \delta, a, f)) \leq \sup_{\substack{\delta \geq 0, \epsilon: \|\delta\|_\infty \leq 2M, \\ \|\epsilon\|_\infty \leq tM}} \sup_{a, f} (\pi(x_n + \epsilon, a, f) - \pi(x_n + \epsilon - \delta, a, f)).$$

Let $A_{t,n}$ denote the right hand side of the preceding equation; note that this is a deterministic quantity, and that the supremum is over a finite set. Thus from Assumption 4, we have $\limsup_{n \rightarrow \infty} A_{t,n} \leq 0$.

Finally, observe that since $\limsup_{\|x\|_\infty \rightarrow \infty} \sup_{a, f} (\pi(x + \delta, a, f) - \pi(x, a, f)) \leq 0$, it follows that:

$$\sup_{y \in \mathbb{Z}_+^d, \delta \geq 0: \|\delta\|_\infty \leq 2M} \sup_{a, f} (\pi(y, a, f) - \pi(y - \delta, a, f)) < \infty.$$

We denote the left hand side of the preceding inequality by D . Note that this is a constant independent of n .

Combining our arguments, we have that for all sufficiently large n , there holds:

$$\mathbb{E} \left[\sum_t \beta^t (\pi(x_{t,n}, a_{t,n}, f_n) - \pi(x'_{t,n}, a'_{t,n}, f_n)) \right] \leq -\kappa(r) + \sum_{t=1}^T \beta^t A_{t,n} + \frac{\beta^T D}{1 - \beta}.$$

By taking T sufficiently large, we can ensure that the last term on the right hand side is strictly less than $\kappa(r)/2$; and by then taking n sufficiently large, we can ensure that the second term on the right hand side is also strictly less than $\kappa(r)/2$. Thus for sufficiently large n , we conclude that the left hand side is negative—contradicting optimality of μ_n . The lemma follows. \blacksquare

Lemma 5. *There exists $\bar{\epsilon} > 0$ and \bar{K} such that for all ℓ and all x with $x_\ell \geq \bar{K}$,*

$$\sup_f \sup_{\mu \in \mathcal{P}(f)} \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x, \mu(x), f) < -\bar{\epsilon}.$$

Proof. Fix $\epsilon > 0$ so that for all ℓ and all x' with $x'_\ell \geq K'$, $\sup_{a' \in \mathcal{A}'} \sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x', a', f) < -\epsilon$; such a constant exists by the last part of Assumption 4. Observe that since \mathcal{A} is compact and $\sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x, a, f)$ is continuous in a (Assumption 4), it follows that $\sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x, a, f)$ is in fact *uniformly* continuous in $a \in \mathcal{A}$. Let $e^{(\ell)}$ denote the ℓ 'th standard basis vector (i.e., $e_{\ell'}^{(\ell)} = 0$ for $\ell' \neq \ell$, and $e_\ell^{(\ell)} = 1$). By uniform

continuity, we can conclude there must exist a $\delta_\ell > 0$ such that if $\|a - a'\|_\infty < \delta_\ell$, then:

$$\left| \sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | K' e^{(\ell)}, a, f) - \sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | K' e^\ell, a', f) \right| < \epsilon/2.$$

Note in particular, if $\rho_\infty(a, \mathcal{A}') < \delta_\ell$, then there exists $a' \in \mathcal{A}'$ with $\|a - a'\|_\infty < \delta_\ell$. By our choice of ϵ we have $\sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | K' e^{(\ell)}, a, f) < -\frac{\epsilon}{2}$. Now let $\delta = \min\{\delta_1, \dots, \delta_d\}$. Since the increment kernel is stochastically nonincreasing in x , it follows that if $\rho_\infty(a, \mathcal{A}') < \delta$ and $x_\ell \geq K'$, then $\sup_f \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x, a, f) < -\frac{\epsilon}{2}$. Since $\sup_f \sup_{\mu \in \mathcal{P}(f)} \rho_\infty(\mu(x), \mathcal{A}') \rightarrow 0$ as $\|x\|_\infty \rightarrow \infty$, the result follows if we let $\bar{\epsilon} = \epsilon/2$. \blacksquare

Lemma 6. *For every $f \in \mathfrak{F}$, $\Phi(f)$ is nonempty.*

Proof. As described in the discussion of Section 5.3, it suffices to show that the state Markov chain induced by an optimal oblivious strategy possesses at least one invariant distribution—i.e., that $\mathcal{D}(\mu, f)$ is nonempty, where μ is an optimal oblivious strategy given f .

We first show that for every f and every $\mu \in \mathcal{P}(f)$, the Markov chain on \mathcal{X} induced by μ and f has at least one closed class. Let $S = \{x : \|x\|_\infty \leq \bar{K} + M\}$. By Lemma 5, if $x \notin S$, then there exists some state x' with $\mathbf{P}(x'|x, \mu(x), f) > 0$ such that $x'_\ell \leq x_\ell - \bar{\epsilon}$ for all ℓ where $x_\ell > \bar{K}$. On the other hand, since increments are bounded, for any ℓ where $x_\ell \leq \bar{K}$, we have $x'_\ell \leq \bar{K} + M$. Applying this fact inductively, we find that for any $x \notin S$, there must exist a positive probability sequence of states from x to S ; i.e., a sequence $y_0, y_1, y_2, \dots, y_\tau$ such that $y_0 = x, y_\tau \in S$, and $\mathbf{P}(y_t|y_{t-1}, \mu(y_{t-1}), f) > 0$ for all t . We say that S is *reachable* from x .

So now suppose the chain induced by μ and f has no closed class. Fix $x_0 \in S$. Since the class containing x_0 is not closed, there must exist a state x' reachable from x_0 with positive probability, such that the chain never returns to x_0 starting from x' . If $x' \in S$, let $x_1 = x'$. If $x' \notin S$, then using the argument in the preceding paragraph, there must exist a state $x_1 \in S$ reachable from x' . Arguing inductively, we can construct a sequence of states x_0, x_1, x_2, \dots where $x_t \in S$ for all t , and yet x_0, \dots, x_{t-1} are not reachable from x_t . But S is finite, so at least one state must repeat in this sequence—contradicting the construction. We conclude that the chain must have at least one closed class.

To complete the proof, we use a Foster-Lyapunov argument. Let $U(x) = \sum_\ell x_\ell^2$. Then $\{x \in \mathcal{X} : U(x) \leq R\}$ is finite for all R . So now let $\omega = (2d\bar{K}M + dM^2 + 1)/(2\bar{\epsilon})$, and suppose $\|x\|_\infty > \max\{\omega, \bar{K}\}$. We reason as follows:

$$\begin{aligned} \sum_{x'} U(x') \mathbf{P}(x'|x, \mu(x), f) &= U(x) + 2 \sum_\ell x_\ell \sum_{z_\ell} z_\ell \mathbf{Q}_\ell(z_\ell | x, \mu(x), f) + \sum_\ell \sum_{z_\ell} z_\ell^2 \mathbf{Q}_\ell(z_\ell | x, \mu(x), f) \\ &\leq U(x) + 2 \sum_{\ell: x_\ell \leq \bar{K}} M x_\ell - 2 \sum_{\ell: x_\ell > \bar{K}} \bar{\epsilon} x_\ell + dM^2 \leq U(x) - 1. \end{aligned}$$

The first equality follows by definition of \mathbf{Q} and U , and multiplicative separability of \mathbf{Q} . The next step follows since increments are bounded (Assumption 4), and by applying Lemma 5 for $x_\ell > \bar{K}$. The last inequality follows from the fact that the state space is d -dimensional, $\|x\|_\infty > \max\{\bar{K}, \omega\}$, and by definition

of ω . Since increments are bounded, it is trivial that for every R :

$$\sup_{x: \|x\|_\infty \leq R} \left(\sum_{x'} U(x') \mathbf{P}(x' | x, \mu(x), f) - U(x) \right) < \infty.$$

It follows by the Foster-Lyapunov criterion that every closed class of the Markov chain induced by μ is positive recurrent, as required (Hajek 1982, Meyn and Tweedie 1993, Glynn and Zeevi 2006). ■

The following lemma uses a Foster-Lyapunov argument as well, in a spirit similar to the proof of the preceding lemma; we omit the details (see Adlakha et al. 2011).

Lemma 7. *For every $\eta \in \mathbb{Z}_+$, $\sup_f \sup_{\phi \in \Phi(f)} \sum_x \|x\|_\eta^\eta \phi(x) < \infty$.*

Proof of Proposition 11. We have already established that $\Phi(f)$ is nonempty in Lemma 6. Define $B = \sup_f \sup_{\phi \in \Phi(f)} \sum_x \|x\|_{p+1}^{p+1} \phi(x) < \infty$, where the inequality is the result of Lemma 7.

We define the set $\mathcal{C} = \left\{ f \in \mathfrak{F} : \sum_x \|x\|_{p+1}^{p+1} f(x) \leq B \right\}$. By the preceding observation, $\Phi(\mathcal{C}) \subset \mathcal{C}$. It is clear that \mathcal{C} is nonempty and convex. It remains to be shown that \mathcal{C} is compact in the 1- p -norm. It is straightforward to check that \mathcal{C} is complete; in Adlakha et al. (2011) we show that \mathcal{C} is totally bounded, thus establishing compactness. ■

C AME: Proof

Throughout this section we suppose Assumption 1 holds. We begin by defining the following sets.

Definition 11. *For every $x \in \mathcal{X}$, define*

$$\mathcal{X}_x = \left\{ z \in \mathcal{X} \mid \mathbf{P}(x | z, a, f) > 0 \text{ for some } a \in \mathcal{A} \text{ and for some } f \in \mathfrak{F}_p \right\}. \quad (13)$$

Also define $\mathcal{X}_{x,t}$ as

$$\mathcal{X}_{x,t} = \left\{ z \in \mathcal{X} \mid \|z\|_\infty \leq \|x\|_\infty + tM \right\}. \quad (14)$$

Thus, \mathcal{X}_x is the set of all initial states that can result in the final state as x . Since the increments are bounded (Assumption 1), for every $x \in \mathcal{X}$, the set \mathcal{X}_x is finite. The set $\mathcal{X}_{x,t}$ is a superset of all possible states that can be reached at time t starting from state x (since the increments are uniformly bounded over action a and distribution f); note that $\mathcal{X}_{x,t}$ is finite as well.

The following key lemma establishes that as the number of players grows large, the population empirical distribution in a game with finitely many players approaches the limiting SE population. Due to space constraints we do not provide the proof here. The result is similar in spirit to related results on mean field limits of interacting particle systems, cf. Sznitman (1991); there the main insight is that, under appropriate conditions, the stochastic evolution of a finite-dimensional interacting particle system approaches the deterministic mean field limit over finite time horizons. Our model introduces two sources of complexity. First, agents' state transitions are coupled, so the population state Markov process is not simply the aggregation

of independent agent state dynamics. Second, our state space is unbounded, so additional care is required to ensure the tail of the population state distribution is controlled in games with a large but finite number of players. This is where the light tail condition plays a key role. Our proof proceeds by induction over time periods; see Adlakha et al. (2011) for details.

Lemma 8. *Let (μ, f) be a stationary equilibrium with $f \in \mathfrak{F}_p$. Consider an m -player game. Let $x_{i,0}^{(m)} = x_0$ and suppose the initial state of every player (other than player i) is independently sampled from the distribution f . That is, suppose $x_{j,0}^{(m)} \sim f$ for all $j \neq i$; let $f^{(m)} \in \mathfrak{F}^{(m)}$ denote the initial population state. Let $a_{i,t}^{(m)}$ be any sequence of (possibly random, possibly history dependent) actions. Suppose players' states evolve as $x_{i,t+1}^{(m)} \sim \mathbf{P}(\cdot \mid x_{i,t}^{(m)}, a_{i,t}^{(m)}, f_{-i,t}^{(m)})$ and for all $j \neq i$, as $x_{j,t+1}^{(m)} \sim \mathbf{P}(\cdot \mid x_{j,t}^{(m)}, \mu(x_{j,t}^{(m)}), f_{-j,t}^{(m)})$. Then, for every initial state x_0 , for all times t , $\left\| f_{-i,t}^{(m)} - f \right\|_{1-p} \rightarrow 0$ almost surely¹⁰ as $m \rightarrow \infty$.*

Before we prove the AME property, we need some additional notation. Let (μ, f) be a stationary equilibrium. Consider again an m player game and focus on player i . Let $x_{i,0}^{(m)} = x_0$ and assume that player i uses a cognizant strategy μ_m . The initial state of every other player $j \neq i$ is independently drawn from the distribution f , that is, $x_{j,0}^{(m)} \sim f$. Denote the initial distribution of all $m - 1$ players (excluding player i) by $f^{(m)} \in \mathfrak{F}^{(m)}$. The state evolution of player i is given by

$$x_{i,t+1}^{(m)} \sim \mathbf{P} \left(\cdot \mid x_{i,t}^{(m)}, a_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)} \right), \quad (15)$$

where $a_{i,t}^{(m)} = \mu_m(x_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)})$ and $\mathbf{f}_{-i,t}^{(m)}$ is the actual population distribution. Here the superscript m on the state variable represents the fact that we are considering an m player stochastic game. Let every other player j use the oblivious strategy μ and thus their state evolution is given by

$$x_{j,t+1}^{(m)} \sim \mathbf{P} \left(\cdot \mid x_{j,t}^{(m)}, \mu(x_{j,t}^{(m)}), f_{-j,t}^{(m)} \right). \quad (16)$$

Define $V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)})$ to be the actual value function of player i , with its initial state x , the initial distribution of the rest of the population as $f^{(m)} \in \mathfrak{F}^{(m)}$, when the player uses a cognizant strategy μ_m and every other player uses an oblivious strategy μ . We have

$$V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)}) = \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \pi(x_{i,t}, a_{i,t}, \mathbf{f}_{-i,t}^{(m)}) \mid x_{i,0} = x, f_{-i,0}^{(m)} = f^{(m)}; \right. \\ \left. \mu_i = \mu_m, \boldsymbol{\mu}_{-i} = \boldsymbol{\mu}^{(m-1)} \right]. \quad (17)$$

We define a new player that is coupled to player i in the m player stochastic games defined above. We call this player the *coupled* player. Let $\hat{x}_{i,t}^{(m)}$ be the state of this coupled player at time t . The subscript i and the superscript m reflect the fact that this player is coupled to player i in an m player stochastic game. We assume that the state evolution of this player is given by:

$$\hat{x}_{i,t+1}^{(m)} \sim \mathbf{P} \left(\cdot \mid \hat{x}_{i,t}^{(m)}, \hat{a}_{i,t}^{(m)}, f \right), \quad (18)$$

¹⁰Note that the convergence is almost surely in the randomness associated with the initial population state.

where $\hat{a}_{i,t}^{(m)} = a_{i,t}^{(m)} = \mu_m(x_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)})$. In other words, this coupled player takes the same action as player i at every time t and this action depends on the *actual* population state of $m - 1$ players. However, note that the state evolution is dependent only on the mean field population state f . Let us define

$$\hat{V}^{(m)}(x \mid f; \mu_m, \boldsymbol{\mu}^{(m-1)}) = \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \pi(\hat{x}_{i,t}^{(m)}, \hat{a}_{i,t}^{(m)}, f) \mid \hat{x}_{i,0}^{(m)} = x_0, \hat{a}_{i,t}^{(m)} = \mu_m(x_{i,t}, \mathbf{f}_{-i,t}^{(m)}); \boldsymbol{\mu}^{(m-1)} \right]. \quad (19)$$

Thus, $\hat{V}^{(m)}(x \mid f; \mu_m, \boldsymbol{\mu})$ is the expected net present value of this coupled player, when the player's initial state is x , the long run average population state is f , and the initial population state is $f_{-i,0}^{(m)} = f^{(m)}$. Observe that

$$\begin{aligned} \hat{V}^{(m)}(x \mid f; \mu_m, \boldsymbol{\mu}^{(m-1)}) &\leq \sup_{\mu' \in \mathfrak{M}} \hat{V}^{(m)}(x \mid f; \mu', \boldsymbol{\mu}^{(m-1)}) = \sup_{\mu' \in \mathfrak{M}_O} \hat{V}^{(m)}(x \mid f; \mu', \boldsymbol{\mu}^{(m-1)}) \\ &= \tilde{V}^*(x \mid f) = \tilde{V}(x \mid \mu, f). \end{aligned} \quad (20)$$

Here, the first equality follows from Lemma 2, which implies that the supremum over all cognizant strategies is the same as the supremum over oblivious strategies (since the state evolution of other players does not affect the payoff of this coupled player), and the last equality follows since $\mu \in \mathcal{P}(f)$.

Lemma 9. *Let (μ, f) be a stationary equilibrium and consider an m player game. Let the initial state of player i be $x_{i,0}^{(m)} = x$, and let $f^{(m)} \in \mathfrak{F}^{(m)}$ be the initial population state of $m - 1$ players whose initial state is sampled independently from the distribution f . Assume that player i uses a cognizant strategy μ_m and every other player uses the oblivious strategy μ . Their state evolutions are given by equation (15) and (16). Also define a coupled player with initial state $\hat{x}_{i,0}^{(m)} = x$ and let its state evolution be given by equation (18). Then, for all times t , and for every $y \in \mathcal{X}$, we have $\left| \text{Prob}(\hat{x}_{i,t}^{(m)} = y) - \text{Prob}(x_{i,t}^{(m)} = y) \right| \rightarrow 0$, almost surely¹¹ as $m \rightarrow \infty$.*

Proof. The lemma is trivially true for $t = 0$. Let us assume that it holds for all times $\tau = 0, 1, \dots, t - 1$. Then, we have

$$\begin{aligned} \text{Prob}(x_{i,t}^{(m)} = y) &= \sum_{z \in \mathcal{X}_y} \text{Prob}(x_{i,t-1}^{(m)} = z) \mathbf{P}(y \mid z, \mu_m(z, f_{-i,t-1}^{(m)}), f_{-i,t-1}^{(m)}) \\ \text{Prob}(\hat{x}_{i,t}^{(m)} = y) &= \sum_{z \in \mathcal{X}_y} \text{Prob}(\hat{x}_{i,t-1}^{(m)} = z) \mathbf{P}(y \mid z, \mu_m(z, f_{-i,t-1}^{(m)}), f). \end{aligned}$$

Here we use the fact that the coupled player uses the same action as player i and the state evolution of the coupled player is given by equation (18). Note that the summation is over all states in the finite set \mathcal{X}_y , where \mathcal{X}_y is defined as in equation (13).

From Lemma 8, we know that for all times t , $\left\| f_{-i,t}^{(m)} - f \right\|_{1-p} \rightarrow 0$ almost surely as $m \rightarrow \infty$. From Assumption 1, we know that the transition kernel is jointly continuous in the action a and distribution f (where

¹¹The almost sure convergence of the probabilities is in the randomness associated with the initial population state.

the set of distributions \mathfrak{F}_p is endowed with 1- p norm). Since the action set is compact, this implies that for all $y, z \in \mathcal{X}$, $\lim_{m \rightarrow \infty} \sup_{a \in \mathcal{A}} \left| \mathbf{P} \left(y \mid z, a, f_{-i,t-1}^{(m)} \right) - \mathbf{P} \left(y \mid z, a, f \right) \right| = 0$ almost surely. It follows that for every $y, z \in \mathcal{X}$, $\lim_{m \rightarrow \infty} \left| \mathbf{P} \left(y \mid z, \mu_m(z, f_{-i,t-1}^{(m)}), f_{-i,t-1}^{(m)} \right) - \mathbf{P} \left(y \mid z, \mu_m(z, f_{-i,t-1}^{(m)}), f \right) \right| = 0$ almost surely. From the induction hypothesis, we know that for every $z \in \mathcal{X}$, $\left| \text{Prob} \left(\hat{x}_{i,t-1}^{(m)} = z \right) - \text{Prob} \left(x_{i,t-1}^{(m)} = z \right) \right| \rightarrow 0$ almost surely as $m \rightarrow \infty$. This along with the finiteness of the set \mathcal{X}_y , gives that for every $y \in \mathcal{X}$ $\left| \text{Prob} \left(\hat{x}_{i,t}^{(m)} = y \right) - \text{Prob} \left(x_{i,t}^{(m)} = y \right) \right| \rightarrow 0$ almost surely as $m \rightarrow \infty$. This proves the lemma. \blacksquare

Lemma 10. *Let (μ, f) be a stationary equilibrium and consider an m player game. Let the initial state of player i be $x_{i,0}^{(m)} = x$, and let $f^{(m)} \in \mathfrak{F}^{(m)}$ be the initial population state of $m - 1$ players whose initial state is sampled independently from the distribution f . Assume that player i uses a cognizant strategy μ_m and every other player uses the oblivious strategy μ . Their state evolutions are given by equation (15) and (16). Also define a coupled player with initial state $\hat{x}_{i,0}^{(m)} = x$ and let its state evolution be given by equation (18). Then, for all times t , we have $\limsup_{m \rightarrow \infty} \mathbb{E} \left[\pi \left(x_{i,t}^{(m)}, \mu_m(x_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)}) - \pi \left(\hat{x}_{i,t}^{(m)}, \mu_m(x_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)}), f \right) \right] \leq 0$, almost surely¹²*

Proof. Let us write $a_{i,t}^{(m)} = \mu_m(x_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)})$. We have

$$\begin{aligned} \Delta_{i,t}^{(m)} &= \mathbb{E} \left[\pi \left(x_{i,t}^{(m)}, a_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)} \right) - \pi \left(\hat{x}_{i,t}^{(m)}, a_{i,t}^{(m)}, f \right) \right] \\ &= \mathbb{E} \left[\pi \left(x_{i,t}^{(m)}, a_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)} \right) - \pi \left(x_{i,t}^{(m)}, a_{i,t}^{(m)}, f \right) \right] + \mathbb{E} \left[\pi \left(x_{i,t}^{(m)}, a_{i,t}^{(m)}, f \right) - \pi \left(\hat{x}_{i,t}^{(m)}, a_{i,t}^{(m)}, f \right) \right] \\ &\triangleq T_{1,t}^{(m)} + T_{2,t}^{(m)}. \end{aligned}$$

Consider the first term. We have

$$\begin{aligned} T_{1,t}^{(m)} &\leq \sum_{y \in \mathcal{X}} \text{Prob} \left(x_{i,t}^{(m)} = y \right) \sup_{a \in \mathcal{A}} \left| \pi \left(y, a, \mathbf{f}_{-i,t}^{(m)} \right) - \pi \left(y, a, f \right) \right| \\ &= \sum_{y \in \mathcal{X}_{x,t}} \text{Prob} \left(x_{i,t}^{(m)} = y \right) \sup_{a \in \mathcal{A}} \left| \pi \left(y, a, \mathbf{f}_{-i,t}^{(m)} \right) - \pi \left(y, a, f \right) \right|, \end{aligned}$$

where the last equality follows from the fact that $x_{i,0}^{(m)} = x$ and from equation (14). From Assumption 1, we know that the payoff is jointly continuous in action a and distribution f (with the set of distributions \mathfrak{F}_p endowed with 1- p norm) and the set \mathcal{A} is compact. Thus, for every $y \in \mathcal{X}$, we have $\sup_{a \in \mathcal{A}} \left| \pi \left(y, a, \mathbf{f}_{-i,t}^{(m)} \right) - \pi \left(y, a, f \right) \right| \rightarrow 0$, almost surely as $m \rightarrow \infty$. This along with the fact that $\mathcal{X}_{x,t}$ is finite shows that $\limsup_{m \rightarrow \infty} T_{1,t}^{(m)} \leq 0$ almost surely.

Now consider the second term. We have

¹²The almost sure convergence of the expected value of the payoff is in the randomness associated with the initial population state.

$$\begin{aligned}
T_{2,t}^{(m)} &= \mathbb{E} \left[\pi \left(x_{i,t}^{(m)}, a_{i,t}^{(m)}, f \right) - \pi \left(\hat{x}_{i,t}^{(m)}, a_{i,t}^{(m)}, f \right) \right] \\
&\leq \sum_{y \in \mathcal{X}} \left| \text{Prob} \left(x_{i,t}^{(m)} = y \right) - \text{Prob} \left(\hat{x}_{i,t}^{(m)} = y \right) \right| \sup_{a \in \mathcal{A}} |\pi(y, a, f)| \\
&= \sum_{y \in \mathcal{X}_{x,t}} \left| \text{Prob} \left(x_{i,t}^{(m)} = y \right) - \text{Prob} \left(\hat{x}_{i,t}^{(m)} = y \right) \right| \sup_{a \in \mathcal{A}} |\pi(y, a, f)|,
\end{aligned}$$

where the last equality follows from the fact that $x_{i,0}^{(m)} = \hat{x}_{i,0}^{(m)} = x$ and from Definition 11. From Lemma 9, we know that for every $y \in \mathcal{X}$, $\left| \text{Prob} \left(x_{i,t}^{(m)} = y \right) - \text{Prob} \left(\hat{x}_{i,t}^{(m)} = y \right) \right| \rightarrow 0$ almost surely $m \rightarrow \infty$. Since $\mathcal{X}_{x,t}$ is finite for every fixed $x \in \mathcal{X}$ and every time t , this implies that $\limsup_{m \rightarrow \infty} T_{2,t}^{(m)} \leq 0$ almost surely. This proves the lemma. \blacksquare

Before we proceed further, we need one additional piece of notation. Once again let (μ, f) be a stationary equilibrium and consider an oblivious player. Let \tilde{x}_t be the state of this oblivious player at time t . We assume that $\tilde{x}_0 = x$ and since the player used the oblivious strategy μ , the state evolution of this player is given by

$$\tilde{x}_{t+1} \sim \mathbf{P} \left(\cdot \mid \tilde{x}_t, \tilde{a}_t, f \right) \quad (21)$$

where $\tilde{a}_t = \mu(\tilde{x}_t)$. We let $\tilde{V}(x \mid \mu, f)$ (as defined in equation (7)) to be the oblivious value function for this player starting from state x .

Also, consider an m player game and focus on player i . We represent the state of player i at time t by $\check{x}_{i,t}^{(m)}$. As before, the superscript m on the state variable represents the fact that we are considering an m player stochastic game. Let $\check{x}_{i,0}^{(m)} = x$ and let player i also use the oblivious strategy μ . The initial state of every other player $j \neq i$ is drawn independently from the distribution f , that is, $\check{x}_{j,0}^{(m)} \sim f$. Denote the initial distribution of all $m - 1$ players (excluding player i) by $f^{(m)} \in \mathfrak{F}^{(m)}$. The state evolution of player i is then given by

$$\check{x}_{i,t+1}^{(m)} \sim \mathbf{P} \left(\cdot \mid \check{x}_{i,t}^{(m)}, \check{a}_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)} \right), \quad (22)$$

where $\check{a}_{i,t}^{(m)} = \mu(\check{x}_{i,t}^{(m)})$. Note that even though the player uses an oblivious strategy, its state evolution is affected by the *actual* population state. Let every other player j also use the oblivious strategy μ and let their state evolution be given by

$$\check{x}_{j,t+1}^{(m)} \sim \mathbf{P} \left(\cdot \mid \check{x}_{j,t}^{(m)}, \mu(\check{x}_{j,t}^{(m)}), \mathbf{f}_{-j,t}^{(m)} \right). \quad (23)$$

Define $V^{(m)}(x, f^{(m)} \mid \mu^{(m)})$ to be the actual value function of the player, when the initial state of the player is x , the initial population distribution is $f^{(m)}$ and every player uses the oblivious strategy μ . That is,

$$V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)}) = \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \pi(\tilde{x}_{i,t}, \tilde{a}_{i,t}, \mathbf{f}_{-i,t}^{(m)}) \mid \tilde{x}_{i,0} = x, f_{-i,0}^{(m)} = f^{(m)}; \mu_i = \mu, \boldsymbol{\mu}_{-i} = \boldsymbol{\mu}^{(m)} \right]. \quad (24)$$

The proofs of the next two lemmas have similar arguments to Lemmas 9 and 10; we omit the details and refer the reader to Adlakha et al. (2011).

Lemma 11. *Let (μ, f) be a stationary equilibrium and consider an m player stochastic game. Let $\tilde{x}_{i,0}^{(m)} = x$, and let $f^{(m)} \in \mathfrak{F}^{(m)}$ be the initial population state of $m - 1$ players whose initial state is sampled independently from f . Assume that every player uses the oblivious strategy μ and their state evolutions are given by equations (22) and (23). Also, consider an oblivious player with $\tilde{x}_0 = x$ and let its state evolution be given by equation (21). Then, for every time t and for all $y \in \mathcal{X}$, we have $\left| \text{Prob}(\tilde{x}_t = y) - \text{Prob}(\tilde{x}_{i,t}^{(m)} = y) \right| \rightarrow 0$, almost surely as $m \rightarrow \infty$.*

Lemma 12. *Let (μ, f) be a stationary equilibrium and consider an m player stochastic game. Let $\tilde{x}_{i,0}^{(m)} = x$, and let $f^{(m)} \in \mathfrak{F}^{(m)}$ be the initial population state of $m - 1$ players whose initial state is sampled independently from f . Assume that every player uses the oblivious strategy μ and their state evolutions are given by equations (22) and (23). Also, consider an oblivious player with $\tilde{x}_0 = x$ and let its state evolution be given by equation (21). Then for all times t , we have $\mathbb{E} \left[\pi(\tilde{x}_t, \mu(\tilde{x}_t), f) - \pi(\tilde{x}_{i,t}^{(m)}, \mu(\tilde{x}_{i,t}^{(m)}), \mathbf{f}_{-i,t}^{(m)}) \right] \rightarrow 0$, almost surely as $m \rightarrow \infty$.*

Proof of Theorem 2. Let us define

$$\Delta V^{(m)}(x, f^{(m)}) \triangleq V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)}) - V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)}).$$

Then we need to show that for all x , $\limsup_{m \rightarrow \infty} \Delta V^{(m)}(x, f^{(m)}) \leq 0$ almost surely. We can write

$$\begin{aligned} \Delta V^{(m)}(x, f^{(m)}) &= V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)}) - \tilde{V}(x \mid \mu, f) + \tilde{V}(x \mid \mu, f) - V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)}) \\ &\leq V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)}) - \hat{V}^{(m)}(x \mid f; \mu_m, \boldsymbol{\mu}^{(m-1)}) + \tilde{V}(x \mid \mu, f) - V^{(m)}(x, f^{(m)} \mid \boldsymbol{\mu}^{(m)}) \\ &\triangleq T_1^{(m)} + T_2^{(m)}. \end{aligned}$$

Here the inequality follows from equation (20). Consider the term $T_1^{(m)}$. We have

$$\begin{aligned} T_1^{(m)} &= V^{(m)}(x, f^{(m)} \mid \mu_m, \boldsymbol{\mu}^{(m-1)}) - \hat{V}^{(m)}(x \mid f; \mu_m, \boldsymbol{\mu}^{(m-1)}) \\ &= \mathbb{E} \left[\sum_{t=0}^{\infty} \beta^t \left(\pi(x_{i,t}^{(m)}, a_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)}) - \pi(\hat{x}_{i,t}^{(m)}, \hat{a}_{i,t}^{(m)}, f) \right) \right], \end{aligned}$$

where the last equality follows from equations (17) and (19). Note that $x_{i,0} = \hat{x}_{i,0} = x$ and $a_{i,t} = \hat{a}_{i,t} = \mu_m(x_{i,t}, \mathbf{f}_{-i,t}^{(m)})$ and the state transitions of players are given by equations (15), (16), and (18). From

Lemma 10, we have $\limsup_{m \rightarrow \infty} \mathbb{E} \left[\sum_{t=0}^{T-1} \beta^t \left(\pi(x_{i,t}^{(m)}, a_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)}) - \pi(\hat{x}_{i,t}^{(m)}, \hat{a}_{i,t}^{(m)}, f) \right) \right] \leq 0$, almost surely for any finite time T . From Lemma 1, we have, almost surely

$$\mathbb{E} \left[\sum_{t=T}^{\infty} \beta^t \left(\pi(x_{i,t}^{(m)}, \hat{a}_{i,t}^{(m)}, \mathbf{f}_{-i,t}^{(m)}) - \pi(\hat{x}_{i,t}^{(m)}, a_{i,t}^{(m)}, f) \right) \right] \leq 2C(x, T),$$

which goes to zero as $T \rightarrow \infty$. This proves that $\limsup_{m \rightarrow \infty} T_1^{(m)} \leq 0$ almost surely. Similar analysis (with an application of Lemma 12) shows that $\limsup_{m \rightarrow \infty} T_2^{(m)} \leq 0$ almost surely, yielding the result. ■