Theoretical Economics 3 (2008), 431–458 1555-7561/20080431

Contagion through learning

Jakub Steiner
School of Economics, University of Edinburgh

Colin Stewart
Department of Economics, University of Toronto

We study learning in a large class of complete information normal form games. Players continually face new strategic situations and must form beliefs by extrapolation from similar past situations. We characterize the long-run outcomes of learning in terms of iterated dominance in a related incomplete information game with subjective priors. The use of extrapolations in learning may generate contagion of actions across games even if players learn only from games with payoffs very close to the current ones. Contagion may lead to unique long-run outcomes where multiplicity would occur if players learned through repeatedly playing the same game. The process of contagion through learning is formally related to contagion in global games, although the outcomes generally differ.

Keywords. Similarity, learning, contagion, case-based reasoning, global games.
JEL classification. C7, D8.

1. Introduction

In standard models of learning, players repeatedly interact in the *same* game, and use their experience from the history of play to decide which action to choose in each period. In many cases of interest, decision-makers are faced with *many different* strategic situations, and the number of possibilities is so vast that a particular situation is virtually never experienced twice. The history of play may nonetheless be informative when

Jakub Steiner: jakub.steiner@ed.ac.uk
Colin Stewart: colin.stewart@utoronto.ca

An earlier version of this paper was distributed under the title "Learning by Similarity in Coordination Problems." We are grateful to Eddie Dekel, Eduardo Faingold, Li Hao, Ed Hopkins, Philippe Jehiel, George Mailath, Ben Polak, Marzena Rostek, József Sákovics, Larry Samuelson, Santiago Sánchez-Pagés, the coeditor Bart Lipman, and several anonymous referees for their comments. We especially thank Stephen Morris and Avner Shaked for many valuable discussions throughout the project, and Hans Carlsson for generously providing simulation results. We also wish to thank the organizers of the VI Trento Summer School in Adaptive Economic Dynamics, and seminar participants at Cambridge, Edinburgh, MIT, NYU, Penn State, PSE, Queen's, Stanford, Toronto, UBC, Washington University in St. Louis, Western Ontario, Wisconsin, Yale, the Econometric Society meetings in Minneapolis and Vienna, and the Global Games workshop at Stony Brook. Jakub Steiner benefited from the grant "Stability of the Global Financial System: Regulation and Policy Response" during his research stay at LSE and from grant no. LC542 of the Ministry of Education of the Czech Republic.

	Invest	Not Invest
Invest	θ,θ	$\theta-1,0$
Not Invest	$0,\theta-1$	$0,0$

TABLE 1. Payoffs in the example of Section 2.

choosing an action, as previous situations, though different, may be similar to the current one. Thus, a tacit assumption of standard learning models is that players extrapolate their experience from previous interactions similar to the current one.

The central message of this paper is that such extrapolation has important effects: *similarity-based* learning can lead to contagion of behavior across very different strategic situations. Two situations that are not directly similar may be connected by a chain of intermediate situations, each of which is similar to the neighboring ones. One effect of this contagion is to select a unique long-run action in situations that would allow for multiple steady states if analyzed in isolation. For this to occur, the extrapolations at each step of the similarity-based learning process need not be large; in fact, the contagion effect remains even in the limit as extrapolation is based only on increasingly similar situations.

Our main application of similarity-based learning is to coordination games. Consider, as an example, the class of 2×2 games $\Gamma(\theta)$ in Table 1 parameterized by a state θ. The action *Invest* is strategically risky, as its payoff depends on the action of the opponent. The safe action, *Not Invest*, gives a constant payoff of 0. For extreme values of θ, the game $\Gamma(\theta)$ has a unique equilibrium as investing is dominant for $\theta>1$, and not investing is dominant for $\theta<0$. When θ lies in the interval $(0,1)$, the game has two strict pure strategy equilibria.

The contagion effect can be sketched without fully specifying the learning process, which we postpone to Section 3. Two myopic players interact in many periods in a game $\Gamma(\theta_t)$, with θ_t selected at random in each period. Roughly, we assume that, after observing the current state θ_t, players estimate their payoffs for each action on the basis of past experience with states similar to θ_t. Two games $\Gamma(\theta_t)$ and $\Gamma(\theta_s)$ are viewed by players as similar if the difference $|\theta_t-\theta_s|$ is small.

Since investing is dominant for all sufficiently high states, there is some $\overline{\theta}$ above which players eventually learn to invest. Once these high states have occurred sufficiently many times, the initial phase in which players may not have invested above $\overline{\theta}$ becomes negligible for the payoff estimates. Now suppose that a state $\theta_t=\overline{\theta}-\varepsilon$ just below $\overline{\theta}$ is drawn. At θ_t, investing may not be dominant, but players view some past games with values of θ above $\overline{\theta}$ as similar. Since the opponent has learned to invest in these similar games, strategic complementarities in payoffs increase the estimated gain from investing. When ε is small, this increase outweighs any loss that may have occurred from investing in games below $\overline{\theta}$ if the opponent did not also invest. Thus players eventually learn to invest in games with states below, but close to $\overline{\theta}$, giving a new threshold $\overline{\theta}'$ above which both players invest.

Repeating the argument with $\overline{\theta}$ replaced by $\overline{\theta}'$, investment continues to spread to games with smaller states, even though these are not directly similar to games in the

dominance region. The process continues until a threshold state θ^* is reached at which, on average, the gain from investment by the opponent above θ^* is exactly balanced by the loss that would occur if the opponent did not invest below θ^*. Not investing spreads contagiously beginning from low states by a symmetric process. These processes meet at the same threshold, giving rise to a unique long-run outcome provided that players place enough weight on states very close to the current one when forming their payoff estimates.

Contagion effects have previously been studied in local interaction and incomplete information games. In local interaction models, actions may spread contagiously across members of a population because each has an incentive to coordinate with her neighbors in a social network (e.g. Morris 2000). In incomplete information games with strategic complementarities (global games), actions may spread contagiously across types because private information gives rise to uncertainty about the actions of other players (Carlsson and Damme 1993). Unlike these models, contagion through learning depends neither on any network structure nor on high orders of reasoning about the beliefs of other players. The contagion is driven solely by a natural solution to the problem of learning the payoffs to one's actions when the strategic situation is continually changing. This problem is familiar from econometrics, where one often wishes to estimate a function of a continuous variable using only a finite data set.[1] The similarity-based payoff estimates used by players in our model have a direct parallel in the use of kernel estimators by econometricians. Moreover, Gilboa and Schmeidler (2001) provide axiomatic foundations for choice according to similarity-weighted payoff estimation in a single-agent context. Our learning model applies case-based decision making to strategic environments.

The main tool for understanding the result of contagion through learning is a formal parallel to equilibrium play in a modified version of the game. One may view the original family of games as a single game of complete information with a move by Nature. Taking this perspective, the modified game differs from the original game only in the prior beliefs: we show that players eventually behave *as if* they incorrectly believe their own observation of the state to be noisy, while correctly believing that other players perfectly observe the true state. More precisely, players learn not to play strategies that would be serially dominated in the modified version of the game (see Theorem 1). The distribution of the noise in the modified game corresponds directly to the similarity functions used in the learning process. Thus, in the long-run, the use of extrapolations across states in learning plays a role analogous to subjective uncertainty across states in static equilibrium.

The relationship between the long-run outcomes of similarity-based learning and serially undominated strategies in the modified game is quite robust. The result holds for a broad class of games and a large class of learning processes that vary in the knowledge players have of the environment. In addition, very little structure is imposed on the similarity functions used by the players in the learning process. Roughly speaking, the

[1] An important difference between the usual econometric problem and our setting is that players' actions can affect their future estimates by influencing other players' action choices.

modified game result holds as long as payoffs and similarity are sufficiently continuous in the state.

In Section 5, we apply the theory to learning in coordination problems. By solving the modified game, we identify the long-run learning outcomes in a class of binary-action coordination games closely related to the global games of Carlsson and Damme (1993) (see Morris and Shin 2003 for a survey). The original game has a continuum of equilibria, but contagion leads to a unique history-independent learning outcome when similarity is concentrated on nearby states. As in global games, this outcome involves symmetric strategies characterized by a single threshold state at which players switch actions. However, the value of the threshold depends on the shape of the similarity function. The similarity between the learning outcome and the global game equilibrium selection arises because the modified game shares much of the structure of global games. However, the learning outcome generally differs from the global game equilibrium. In terms of the modified game characterization, this difference results from the heterogeneity of the priors in the modified game as opposed to the common priors used in global games.

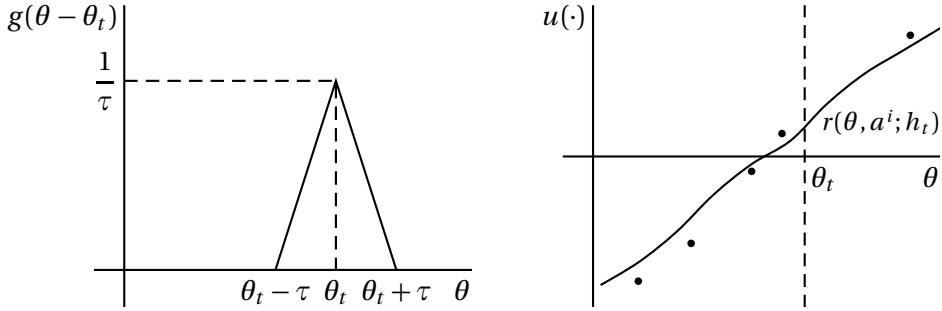
2. Example

Before introducing the general model in Section 3, we elaborate on the example from the previous section to illustrate in more detail the process of contagion through learning. The underlying family of coordination problems consists of the 2-player games in Table 1. We denote by $U(\theta, a^i, a^{-i})$ the payoff to action a^i in state θ when the opponent chooses action a^{-i}. To simplify notation, we refer to investing as action 1 and not investing as action 0.

The game is played repeatedly in periods $t \in \mathbb{N}$, with the state θ_t drawn independently across periods according to a uniform distribution on an interval $[-b, 1+b]$, where $b > 0$. Each realization θ_t is perfectly observed by both players, who play a myopic best response to their beliefs in each period. Beliefs are based on players' previous experience, but since θ is drawn from a continuous distribution, players (almost surely) have no past experience with the current game $\Gamma(\theta_t)$, and must extrapolate from their experience playing different games. In each period, players estimate their payoffs as a weighted average of historical payoffs in which the weights are determined by the similarity between the current and past states. In forming these estimates, players treat the past actions of their opponents as given. Thus following any history $h_t = \{\theta_s, a_s^1, a_s^2\}_{s<t}$, the estimated payoff to player i from choosing action a^i given the state θ_t is

$$r(\theta_t, a^i; h_t) = \frac{\sum_{s<t} g(\theta_s - \theta_t) U(\theta_s, a^i, a_s^{-i})}{\sum_{s<t} g(\theta_s - \theta_t)}, \tag{1}$$

where $g \geq 0$ is the *similarity function* determining the relative weight assigned to past cases. Each player chooses the action giving the highest estimated payoff. Estimates may be chosen arbitrarily if the history contains no state similar to θ_t, that is, if $\sum_{s<t} g(\theta_s - \theta_t) = 0$.



(a) Example similarity function g.

(b) Example payoff estimates. Dots represent observed past payoffs.

FIGURE 1. Example similarity function with corresponding payoff estimates according to (1).

For this example, suppose that g is the piecewise-linear function illustrated in Figure 1(a). Figure 1(b) illustrates the estimated payoffs from choosing action 1 as a function of θ for a particular history of observed payoffs using this similarity function.

The learning process is stochastic, but suppose that the empirical distribution of realized cases may be approximated by the true distribution over θ (this idea is formalized in Section 3). If the opponent plays according to a fixed strategy s^{-i}, player i's expected estimated return to investing in state $\theta \in [-b+\tau, 1+b-\tau]$ is given by

$$\int_\Theta U(\theta', 1, s^{-i}(\theta'))g(\theta'-\theta)\,d\theta'. \tag{2}$$

This expression is formally equivalent to the conditional expectation

$$E[U(\theta', 1, s^{-i}(\theta'))\,|\,\theta]$$

when θ is an imprecise signal of θ', with $\theta'-\theta$ distributed according to the density g. Thus, in the long-run, the similarity-based learner behaves *as if* she observes only a noisy signal of the true state. Theorem 1 makes this connection precise by showing that players learn to play strategies that would be serially undominated in a *modified* game of incomplete information in which each holds these (subjective) beliefs about the information structure.

The long-run outcome of the learning process may be identified by solving this modified game. Suppose that both players follow cut-off strategies with threshold θ^*; that is, both choose action 0 at signals below θ^* and action 1 at signals above θ^*. Each player assigns probability $\frac{1}{2}$ to the true state being greater than her own signal. Since each believes that the other player observes the true state, a player who receives exactly the threshold signal θ^* believes that the other player chooses action 1 with probability $\frac{1}{2}$. In order for this strategy profile to be a Bayesian Nash equilibrium, players must be indifferent between their two actions at the threshold signal. Therefore, θ^* is the unique solution to

$$\tfrac{1}{2}U(\theta, 1, 0)+\tfrac{1}{2}U(\theta, 1, 1)=0. \tag{3}$$

This equilibrium turns out to be the unique serially undominated strategy profile in the modified game, and therefore the unique long-run outcome of the learning process (in the original family of games).

The particular form of condition (3) depends on the symmetry of the similarity function. For asymmetric similarity functions, the probability a player assigns to the true state being greater than her signal can differ from $\frac{1}{2}$. Under general conditions, there is again a unique long-run outcome of learning, but the coefficients in (3) depend on the similarity function g.

The process of contagion through learning has a flavor similar to contagion due to incomplete information in global games. For example, the static game of Table 1 becomes a global game if players observe θ with some private noise. When the noise is small, this game has a unique equilibrium in which players follow the threshold strategy characterized by (3) independently of the shape of the noise (see Carlsson and Damme 1993). Therefore, with asymmetric similarity functions, the learning outcome generally differs from the global game selection. This difference arises through the heterogeneity of the priors in the modified game. In global games with common priors, each player assigns probability $\frac{1}{2}$ to her opponent's expectation of the state being higher than her own, and therefore the threshold type believes that her opponent invests with probability $\frac{1}{2}$ (regardless of the shape of the noise). In the modified game, this belief depends on the similarity function. Moreover, in games with more than two players, a difference in outcomes can arise even with symmetric similarity. We leave a detailed discussion of this comparison to Section 5.

3. The learning process

3.1 *The model*

We begin with a formal description of the stochastic learning process. A fixed set of $I \geq 2$ players interact in periods $t = 0, 1, \ldots$ as follows.

1. At the beginning of each period t, Nature draws a state $\theta_t \in \Theta$ according to a continuous distribution Φ with support on a compact, convex set $\Theta \subset \mathbb{R}^N$ and continuous, positive density ϕ.[2] Draws are independent across periods.[3]

2. All players perfectly observe the realized state θ_t, and then simultaneously choose actions $a^i \in A^i$ according to rules described below. The action set A^i available to each player i is the same across all states in Θ and all periods. Each set A^i is finite. As usual, we write $A = \times_{i=1}^{I} A^i$ for the set of action profiles.

3. At the end of the period, each player i observes signals $v^i(\theta_t, a^i, a_t^{-i})$ for each $a^i \in A^i$, where $v^i : \Theta \times A \longrightarrow V^i$ is the signal function mapping to an arbitrary

[2]The modified game result, Theorem 1, holds also for discrete distributions over Θ; in fact, the proof for the discrete case is much simpler. Our main focus, however, is on the continuous case, which better captures the idea that players cannot learn from repeated interaction in the same game.

[3]All of our results continue to hold as long as the process generating states θ_t is strictly stationary and ergodic, with stationary distribution Φ.

set V^i. To simplify notation, we denote by $v_t^i(a^i)$ the realized signal $v^i(\theta_t, a^i, a_t^{-i})$ for action a^i in period t given θ_t and a_t^{-i}. We write v_t^i for the vector of signals $(v_t^i(a^i))_{a^i \in A^i}$. Note that player i observes counterfactual signals $v_t^i(a^i)$ for actions $a^i \neq a_t^i$ that she did not play at t. We discuss the assumption of counterfactual observations in Section 3.2.

Action choices in each period t are determined as follows. After observing θ_t, each player i forms beliefs over the possible signals $v_t^i(a^i)$, and chooses a_t^i to maximize the expected payoff $u^i(\theta_t, a^i, v_t^i(a^i))$, for some fixed $u^i : \Theta \times A^i \times V^i \longrightarrow \mathbb{R}$. Defining the payoff functions

$$U^i(\theta, a) \equiv u^i(\theta, a^i, v^i(\theta, a^i, a^{-i})),$$

the process describes learning in a family of simultaneous-move games $\Gamma(\theta)$ with payoffs $U^i(\theta, a)$. Beliefs over $v_t^i(a^i)$ are formed based on the realized values of $v_s^i(a^i)$ in past periods $s < t$ in which the realized state θ_s was similar to θ_t.

Similarity is measured according to a fixed *similarity function* $g^i : \Theta \times \Theta \longrightarrow \mathbb{R}_+$ for each player i, where, for each θ, $g^i(\cdot, \theta)$ is integrable. The value $g^i(\theta_s, \theta_t)$ represents the weight assigned to the past state θ_s given that the current state is θ_t.

Following a private history $(\theta_0, v_0^i, \ldots, \theta_{t-1}, v_{t-1}^i, \theta_t)$, player i forms her beliefs as follows. If $\sum_{s<t} g^i(\theta_s, \theta_t) = 0$, then player i forms arbitrary beliefs. The interpretation of this case is that player i perceives all past data to be irrelevant to the problem in state θ_t, and hence ignores it. All of our results are independent of the initial beliefs of players in the learning process.[4]

If, on the other hand, $\sum_{s<t} g^i(\theta_s, \theta_t) > 0$, then for each a^i, player i forms the following belief concerning the distribution of $v_t^i(a^i)$:

$$\Pr\left(v_t^i(a^i) = v\right) = \frac{\sum_{s<t} g^i(\theta_s, \theta_t) \mathbf{1}_v\left(v_s^i(a^i)\right)}{\sum_{s<t} g^i(\theta_s, \theta_t)}$$

for each $v \in V^i$, where $\mathbf{1}_v$ denotes the indicator function for the set $\{v\}$. That is, player i believes that $v_t^i(a^i)$ is distributed according to the past frequency of signals $v_s^i(a^i)$, weighted according to the degree of similarity between θ_s and θ_t. These are precisely the beliefs that would arise if players used kernel estimators to estimate the distribution of $v_t^i(a^i)$ conditional on the state θ_t using the kernel function $g^i(\cdot, \theta_t)$.

After forming her beliefs about $v_t^i(a^i)$, player i chooses her action a_t^i to maximize her expected payoff in period t. That is, she chooses action a_t^i according to

$$a_t^i \in \arg\max_{a^i \in A^i} \frac{\sum_{s<t} u^i(\theta_t, a^i, v_s^i(a^i)) g^i(\theta_s, \theta_t)}{\sum_{s<t} g^i(\theta_s, \theta_t)}. \tag{4}$$

In case there is more than one optimal action, the choice among them is according to an arbitrary fixed rule.

[4] All of our results hold without modification if, instead of initial beliefs, players begin with an arbitrary finite history of play that is sufficiently rich to prevent the case $\sum_{s<t} g^i(\theta_s, \theta_t) = 0$ from occurring.

Following any history $h_t = (\theta_s, a_s)_{s=0}^{t-1}$, the learning process defines a pure strategy $s_t^i : \Theta \longrightarrow A^i$ for each player describing the action that would be chosen in period t at each possible state θ if the realized state were $\theta_t = \theta$. The process therefore gives rise to a probability distribution over sequences of strategy profiles $(s_0, s_1, \ldots)$. All our probabilistic results are with respect to this distribution.

Having formally described the stochastic process we now elaborate on its interpretation as a model of learning. First, note that players form beliefs over signals $v_t^i(a^i)$ directly and do not make further inferences about the action profiles that generate these signals. Our interpretation is that players do not know the functional form of the signal-generating process v^i, and are therefore unable to "back out" any further information from the signals they receive. This formulation is without loss of generality because cases in which players make inferences based on received signals can be captured by choosing the signal function v^i to include all information inferred by player i. Under this interpretation, although player i knows the function u^i, she may not know the game payoffs U^i since she does not know the signal function v^i.

For a given family of games with payoffs $U^i(\theta, a)$, there are many different learning processes corresponding to different ways of decomposing $U^i(\theta, a) = u^i(\theta, a^i, v^i(\theta, a))$ into functions u^i and v^i. The various processes differ in the informational feedback players receive. Two natural cases arise when players observe opponents' actions and when they observe only their own payoffs.

Strategy-based learning Observed signals consist precisely of opponents' action profiles, so that $V^i = A^{-i}$ and $v^i(\theta, a^i, a^{-i}) \equiv a^{-i}$. In particular, signals are independent of a^i. In this case, the payoff functions U^i and u^i are identical: $U^i(\theta, a) \equiv u^i(\theta, a^i, v) = u^i(\theta, a^i, a^{-i})$.

Payoff-based learning Observed signals consist only of the player's own payoffs, so that $V^i = \mathbb{R}$ and $u^i(\theta, a^i, v) \equiv v$. In this case, the functions U^i and v^i are identical: $U^i(\theta, a) \equiv v^i(\theta, a)$.

Strategy-based learning requires that each player i knows her own payoff function $U^i(\theta, a^i, a^{-i})$ and needs to estimate only her opponents' action profile a_t^{-i} at θ_t. Before choosing the action a_t^i, player i forms beliefs about her opponents' actions a_t^{-i} according to their similarity weighed frequency in past periods $s < t$. That is, player i chooses her action a_t^i to maximize

$$\frac{\sum_{s<t} U^i(\theta_t, a_t^i, a_s^{-i}) g^i(\theta_s, \theta_t)}{\sum_{s<t} g^i(\theta_s, \theta_t)}.$$

The informational feedback required for this process is minimal: each player observes only her opponents' actions a_s^{-i} at the end of each period s, and no counterfactual observations are needed.

Payoff-based learning places no requirements on players' knowledge of the payoff functions U^i. At the end of each period s, player i observes the payoffs $U^i(\theta_s, a^i, a_s^{-i})$ that she received or would have received for each action $a^i \in A^i$. Before choosing an

action a_t^i in period t, each player forms beliefs about the payoff to each action according to a similarity-weighted average of the performance of that action in past states θ_s. That is, player i chooses her action a_t^i to maximize

$$\frac{\sum_{s<t} U^i(\theta_s, a_t^i, a_s^{-i}) g^i(\theta_s, \theta_t)}{\sum_{s<t} g^i(\theta_s, \theta_t)}.$$

Players in this learning process are strategically naïve in the sense that they do not reason about the actions of other players; indeed, they treat the problem simply as a single-person decision problem with unknown payoffs and they may not be aware that they are interacting with other players.

In addition to these two processes, the general model encompasses many other processes with varying degrees of informational feedback. Since player i knows the function u^i, these processes also differ in the knowledge of the payoff function U^i that player i must possess.

We impose the following technical assumptions on the learning process.

A1 (Bounded payoffs) There exist upper and lower bounds on $u^i(\theta, a^i, v)$ uniformly over all $(\theta, a^i, v) \in \Theta \times A^i \times V^i$.

A2 Each similarity function g^i is bounded.

The following assumption ensures that players eventually obtain relevant data for every state.

A3 For each θ, $\int_\Theta \phi(\theta') g^i(\theta', \theta)\, d\theta' > 0$.

We require the following continuity assumption.

A4 For every a^i and a^{-i}, the expression $u^i(\theta, a^i, v^i(\theta', a^i, a^{-i})) g^i(\theta', \theta)$ is continuous in θ uniformly over all θ'.[5]

Note that the continuity in Assumption A4 is uniform over $(\theta, \theta', a^i, a^{-i}) \in \Theta \times \Theta \times A^i \times A^{-i}$ because Θ is compact and the action sets are finite. Also note that in the case of payoff-based learning described above, Assumption A4 holds if $g^i(\theta', \theta)$ is continuous.

3.2 *Discussion of the model*

In order to form the payoff estimates in (4), player i must observe only values of $v^i(a^i, \theta_s, a_s^{-i})$ at the end of each period s given the particular state θ_s and given the particular actions a_s^{-i} chosen by the opponents in that period. However, player i must observe the value of $v^i(\theta_s, a^i, a_s^{-i})$ for *every* action $a^i \in A^i$, regardless of the action she actually chose in period s. In some instances of the learning process, such as strategy-based learning, the value of $v^i(\theta_s, a^i, a_s^{-i})$ does not depend on a^i, and hence each

[5] Formally, for each a^i and a^{-i}, given any $\varepsilon > 0$, there exists some $\delta > 0$ (independent of θ and θ') such that $|u^i(\theta, a^i, v^i(\theta', a^i, a^{-i})) g^i(\theta', \theta) - u^i(\theta'', a^i, v^i(\theta', a^i, a^{-i})) g^i(\theta', \theta'')| < \varepsilon$ whenever $|\theta - \theta''| < \delta$.

player needs to observe $v^i(\theta_s, a_s^i, a_s^{-i})$ only for the action a_s^i she actually chose. In other cases, however, players must observe certain counterfactual values of v^i. The observation of these counterfactuals may be viewed as an approximation to a model in which, in each period, players choose according to the preceding rules with high probability, but experiment with some small independent probability by choosing a random action from A^i.

The role of the counterfactual observations is to prevent uninteresting cases in which players fail to learn simply because they never take a particular action. If players observed only the signal for the action actually chosen in each period, one could modify the learning model by basing belief formation for each action only on signals in past periods in which that action was chosen. In the application to coordination problems in Section 5, all of our results hold in the model without counterfactuals if we suppose that each player always plays her dominant action in an open set of states in each dominance region. The choice of dominant actions suffices to begin the process of contagion.

Our notion of similarity $g^i(\theta_t, \theta_s)$ is not directly linked to strategic considerations. Players may view games in states θ_t and θ_s as similar even if they differ in terms of strategic structure. For instance, in the example of Section 2, players may treat the game in state $\theta_t = .99$ as similar to the game in state $\theta_s = 1.01$ even though investing is dominant only in the latter state. While a similarity function that distinguishes sharply between strategically different games may be reasonable for players with precise knowledge of payoffs, the connection must be weaker if players do not know exactly where the divisions lie. For example, if a player who is trying to estimate her opponent's actions does not know her opponent's payoffs, she may expect her opponent to behave similarly even in situations that her opponent views as strategically different.

We rule out time-dependent similarity functions in order to simplify the analysis. More generally, one could suppose that observations are discounted over time according to a nonincreasing sequence $\delta(\tau) \in (0, 1]$ by modifying equation (4) to include an additional factor of $\delta(t - s)$ in both sums. In the undiscounted model, the convergence results presented below rely on the property that changes in payoff estimates in a single period become negligible once players have accumulated enough experience. Since this property continues to hold as long as the series $\sum_{\tau=0}^{\infty} \delta(\tau)$ diverges, we conjecture that all of our results hold in this more general setting. If, on the other hand, this sum converges, then the situation becomes more complicated, as the learning process does not converge in general. It is therefore not possible for the long-run behavior to agree with that of the undiscounted process in every period. However, as long as memory is "sufficiently long," we expect this agreement to occur in a large fraction of periods. For example, if memory is discounted exponentially, so that $\delta(\tau) = \rho^\tau$ for some $\rho \in (0, 1)$, then we expect play to be consistent with our results most of the time when ρ is close to 1. Simulations run by Carlsson (personal communication) lend some support to this conjecture. In an environment similar to the example of Section 2, Carlsson simulates a learning model akin to strategy-based learning with a fixed finite memory. In these simulations, strategies converge to the long-run outcomes of our model except in a small measure of states around the threshold, where behavior oscillates.

4. Long-run characterization

In this section, we characterize the long-run outcomes of the learning process from Section 3 in terms of the equilibria of a particular game, which we call the modified game. We begin by informally outlining an observation that lies at the core of this connection. The informal outline is based on a heuristic application of the Law of Large Numbers treating the strategies as stationary; Theorem 1 below formalizes the connection allowing for strategies to change over time.

Suppose that the learning process converges to a time-invariant strategy profile $s(\theta)$. By the Law of Large Numbers, player i's long-run estimated payoff for action a^i in state θ_t against the profile s^{-i} approaches

$$\frac{\int_\Theta u^i(\theta_t, a^i, v^i(\theta, a^i, s^{-i}(\theta)))\phi(\theta)g^i(\theta,\theta_t)\,d\theta}{\int_\Theta \phi(\theta')g^i(\theta',\theta_t)\,d\theta'} = \int_\Theta u^i(\theta_t, a^i, v^i(\theta, a^i, s^{-i}(\theta)))q^i(\theta \mid \theta_t)\,d\theta,$$

where

$$q^i(\theta \mid \theta_t) = \frac{\phi(\theta)g^i(\theta,\theta_t)}{\int_\Theta \phi(\theta')g^i(\theta',\theta_t)\,d\theta'}.$$

That is, player i's expected estimated payoff at state θ_t against the strategy profile s^{-i} coincides with the expected payoff against the same strategy profile of a player with payoffs $u^i(\theta_t, a^i, v^i(\theta, a^i, s^{-i}(\theta)))$ and beliefs $q^i(\theta \mid \theta_t)$ over the state θ.

The virtual conditional belief $q^i(\theta \mid \theta_t)$ has a convenient interpretation. Suppose that the state θ is drawn according to the distribution Φ, and player i observes only a noisy signal θ_t of θ, where the signal is conditionally distributed according to the density $g^i(\theta,\cdot)/\int_\Theta g^i(\theta,\theta')\,d\theta'$. Then $q^i(\theta \mid \theta_t)$ is precisely the density describing player i's posterior beliefs over θ after observing the signal θ_t. Thus a player with beliefs $q^i(\theta \mid \theta_t)$ can be seen as viewing her observation of θ to be noisy, corresponding to the use of different past values of θ in the learning process. This interpretation motivates the following definition.

Definition 1. The *modified game* is a Bayesian game with heterogeneous priors. The players $i \in \{1,\dots,I\}$ simultaneously choose actions $a^i \in A^i$. The state space is given by $\Omega = \Theta^{I+1}$, with typical member $(\theta, \theta^1, \dots, \theta^I)$, where each θ^i denotes the type of player i, and θ is a common payoff parameter. Each player i has payoff function $u^i(\theta^i, a^i, v^i(\theta, a^i, a^{-i}))$. Player i assigns probability 1 to the event that $\theta^j = \theta$ for all $j \neq i$, and has prior beliefs over (θ, θ^i) given by the density $\phi(\theta)g^i(\theta,\theta^i)/\int_\Theta g^i(\theta,\theta')\,d\theta'$.

Whereas the family of games $\Gamma(\theta)$ describes the actual environment in which the players interact, the modified game describes a virtual setting. The beliefs in the modified game do not describe what players literally believe in the learning process. Rather, the modified game merely provides a useful tool for studying the learning outcomes because the learning process converges, in a sense that is made precise below, to the set of strategies that are serially undominated in the modified game.

In order to describe this connection formally, note first that for any game with subjective priors, we may define (interim) dominated strategies in the same way as for Bayesian games with common priors.[6] In fact, we also use a stronger form of dominance in which the payoff difference exceeds some fixed $\pi \geq 0$. Consider any function $\mathbf{a}^i : \Theta \longrightarrow 2^{A^i}$. We interpret $\mathbf{a}^i(\theta)$ as the set of admissible actions for player i at type θ. The profiles $(\mathbf{a}^i)_i$ and $(\mathbf{a}^j)_{j \neq i}$ are denoted, as usual, by $\mathbf{a}$ and $\mathbf{a}^{-i}$ respectively.

Definition 2. A strategy s^i is *consistent with* $\mathbf{a}^i$ if $s^i(\theta) \in \mathbf{a}^i(\theta)$ for all $\theta \in \Theta$. A strategy profile s^{-i} is consistent with $\mathbf{a}^{-i}$ if each component of s^{-i} is consistent with the corresponding component of $\mathbf{a}^{-i}$.

For any profile $\mathbf{a}$, action $a^i \in \mathbf{a}^i(\theta)$ is said to be *π-dominated*[7] at θ under the profile $\mathbf{a}$ if there exists $a^{i\prime} \in \mathbf{a}^i(\theta)$ such that for all s^{-i} consistent with $\mathbf{a}^{-i}$,

$$E_{q(\theta' \mid \theta)}\left[u^i(\theta, a^{i\prime}, v^i(\theta', a^{i\prime}, s^{-i}(\theta'))) - u^i(\theta, a^i, v^i(\theta', a^i, s^{-i}(\theta'))) \mid \theta\right] > \pi.$$

We define iterated elimination of π-dominated strategies in the usual way. For each i and $\pi > 0$, let $\mathbf{a}^i_{0,\pi}(\theta) \equiv A^i$. For $k = 1, 2, \ldots$, define $\mathbf{a}^i_{k,\pi}(\theta)$ to be the set of actions that are not π-dominated for type θ of player i under the profile $\mathbf{a}_{k-1,\pi}$. The set of *serially π-undominated* actions for type θ of player i is given by $\mathbf{a}^i_{\infty,\pi}(\theta) = \bigcap_k \mathbf{a}^i_{k,\pi}(\theta)$. Since π-domination agrees with the usual notion of strict dominance when $\pi = 0$, we drop the prefix π in that case.

The need to consider π-domination instead of ordinary strict domination arises because of the difference between estimated payoffs following finite histories and their long-run expectations. In the proof of Theorem 1 below, we show that for any $\pi > 0$, estimated payoffs under the learning process almost surely eventually lie within π of the corresponding expected payoffs in the modified game. It follows that actions that are serially π-dominated in the modified game will (almost surely eventually) not be played under the learning process. The following lemma, proved in the Appendix, shows that considering serial π-domination for arbitrary $\pi > 0$ suffices to prove the result for $\pi = 0$, that is, for serial strict domination. The lemma is trivial for a single round of elimination, but not for multiple rounds since differences between π-domination and strict domination generally become compounded as the iterative elimination proceeds.

Lemma 1. *Fix any type θ of player i in the modified game and any $k \in \mathbb{N}$. If $a^i \notin \mathbf{a}^i_{k,0}(\theta)$, then there exists some $\pi > 0$ such that $a^i \notin \mathbf{a}^i_{k,\pi}(\theta)$.*

The main result of this section, given in the following theorem (which is proved in the Appendix), shows that, in the long-run, players do not play strategies that are serially dominated in the modified game. Note that strategies in each period of the learning process are defined over the set of states Θ, which is identical to the type space for each

[6] Since no other notion of domination is employed here, we henceforth drop the term "interim" and refer simply to "dominated strategies."

[7] The notion of π-domination should not be confused with the unrelated concept of p-dominance that has appeared in the literature on higher-order beliefs.

player in the modified game. Strategies s_t^i under the learning process may therefore be identified with strategies s^i in the modified game; to keep the notation simple, we do not distinguish between the two.

THEOREM 1. *(i) For any $k \in \mathbb{N}$ and any $\pi > 0$, the strategy profiles s_t under the learning process are almost surely eventually consistent with $\mathbf{a}_{k,\pi}$.*[8]

(ii) The probability that the action profile in period t under the learning dynamics is consistent with the set of serially undominated actions at θ_t in the modified game approaches one as time tends to infinity. That is,

$$\Pr(s_t^i(\theta_t) \text{ is consistent with } \mathbf{a}_{\infty,0}^i(\theta_t)\ \forall i) \to 1$$

as $t \to \infty$.

Using the Strong Law of Large Numbers, it is relatively straightforward to show that in a given state against a fixed strategy, the long-run payoff estimate in the learning process approaches the expected payoff in the modified game. The main difficulty in the proof of the preceding theorem arises because, in order for the analogue of iterated elimination of dominated strategies to occur under the learning dynamics, players must learn not to play dominated actions in finite time at an *uncountable* set of states. Accordingly, the proof demonstrates that it is possible to reduce the problem to one involving a finite state space while introducing only an arbitrarily small error in the payoff estimates.

Theorem 1 characterizes a set of strategy profiles to which the learning process converges with probability 1. We focus below on cases in which this set consists of a single element. More generally, if the set is not a singleton, a natural question is whether one can identify a smaller set of outcomes to which the learning process must converge. While a full characterization is beyond the scope of this paper, it is possible to suggest the form that such restrictions might take. Note that when similarity varies continuously in θ, payoff estimates in the learning process must also be continuous in θ after any history. It follows that, unless payoff estimates happen to be identical for two actions across an interval of states, strategies under learning must not be highly discontinuous in θ. Thus, for instance, strategies with a dense set of discontinuities generally do not appear even if they are serially undominated in the modified game. Carlsson (2004) proposes a related restriction on strategies that simplifies the characterization of monotone equilibria in a large class of games.

A different approach is to consider alternative equilibrium concepts in the modified game. Thus, for example, one might ask whether the learning process must converge to the set of Bayesian Nash equilibria of the modified game. Roughly speaking, if the strategies in the learning process converge with positive probability, then, conditional on convergence, convergence is almost surely to Bayesian Nash equilibrium. Otherwise, the long-run payoff estimates must differ from the expected payoffs in the modified game, which is a zero probability event.

[8] Recall that a property holds *eventually* if there exists some T such the property holds for all $t \geq T$.

5. Contagion in coordination games

We now focus on learning by similarity in a class of symmetric binary-action coordination games $\Gamma(\theta)$, where the distribution $\Phi(\theta)$ has support $\Theta = [\underline{\theta}, \overline{\theta}]$.[9] Each of I players chooses between two actions, 0 and 1. We normalize the payoff from action 0 to be 0 in every state θ against every action profile. We denote by $U(\theta, l)$ the payoff from choosing action 1 in state θ when $l \in \{0, \ldots, I-1\}$ opponents choose action 1.

The similarity function is identical across players, and depends only on the difference $\theta' - \theta$ between states and a scaling parameter $\tau > 0$ according to

$$g^i(\theta', \theta) \equiv \frac{1}{\tau} g\left(\frac{\theta' - \theta}{\tau}\right),$$

where $g : \mathbb{R} \longrightarrow \mathbb{R}_+$. We normalize g to be a probability density function. While additional restrictions on the similarity function seem natural—for example, that g be decreasing in $|\theta' - \theta|$—our main results hold whether or not we impose such restrictions.

We focus on outcomes in the limit as τ tends to 0, where similarity is narrowly concentrated on nearby states. Away from this limit, similarity-based learning generally leads to inconsistent estimates (in the statistical sense) even if the opponents' strategies are fixed. This inconsistency arises because nonlinearities in payoffs or asymmetries in similarity can push the similarity-weighted average of payoffs around θ away from the payoff at θ. When τ is small, if play converges, these inconsistencies become small except possibly in states close to discontinuities in payoffs.

As before, the learning process may take different forms, such as payoff-based or strategy-based learning, depending on the feedback players receive over time. To capture this, we write the payoff to action 1 as

$$U(\theta, l) = u(\theta, v(\theta, l)).$$

Whenever $\sum_{s<t} g\left((\theta_s - \theta_t)/\tau\right) > 0$, the estimated payoff for action 0 is simply 0, and that for action 1 is given by

$$\frac{\sum_{s<t} u(\theta_t, v(\theta_s, l_s)) \frac{1}{\tau} g\left(\frac{\theta_s - \theta_t}{\tau}\right)}{\sum_{s<t} \frac{1}{\tau} g\left(\frac{\theta_s - \theta_t}{\tau}\right)}.$$

In addition to the general assumptions from Section 3, we assume the following.

A5 (State Monotonicity) The payoffs $U(\theta, 0)$ and $U(\theta, I-1)$ are strictly increasing in θ.

A6 (Extremal Payoffs at Extremal Profiles) For all $l = 0, \ldots, I-1$ and all $\theta \in \Theta$, $U(\theta, 0) \leq U(\theta, l) \leq U(\theta, I-1)$.

[9] Frankel et al. (2003) prove existence, uniqueness, and monotonicity of equilibria in asymmetric global games with many actions. Although they assume common priors, their argument does not rely on commonality *per se*, and can in principle be extended to global games with heterogeneous priors of the form considered here. We restrict our attention to symmetric binary-action games to facilitate explicit characterization of the equilibrium.

A7 (Dominance Regions) There exist some $\underline{\theta}', \overline{\theta}' \in (\underline{\theta}, \overline{\theta})$ such that action 0 is dominant at every state below $\underline{\theta}'$ and action 1 is dominant at every state above $\overline{\theta}'$.

A8 (Continuity) The payoffs $U(\theta, 0)$ and $U(\theta, I-1)$ are continuous in θ.

Assumptions A5–A8 are variants of standard global game assumptions, though some details differ. Assumption A6 substantially weakens the strategic complementarity assumption typically used in global games since we do not require that $U(\theta, l') \geq U(\theta, l)$ if $0 < l < l' < I-1$. In fact, we do not impose any restrictions on the relative values of $U(\theta, l')$ for $l' \neq 0, I-1$, and the outcome of the model is independent of these values. In contrast, by choosing values of $U(\theta, l')$ for $l' \neq 0, I-1$ that violate strategic complementarity, one may readily construct games satisfying Assumption A6 for which the global games approach does not select a unique equilibrium.

Let G be the cumulative distribution function corresponding to the density g. Define the threshold θ^* to be the (unique) solution to

$$G(0)U(\theta, 0) + (1 - G(0))U(\theta, I-1) = 0. \tag{5}$$

The existence of this solution is guaranteed by the existence of dominance regions (Assumption A7), and its uniqueness by state monotonicity (Assumption A5).

Proposition 1. *For any $\delta > 0$, there exists $\overline{\tau} > 0$ such that for any $\tau \in (0, \overline{\tau})$, in the learning process with parameter τ, all players almost surely eventually choose action 0 whenever $\theta_t < \theta^* - \delta$ and action 1 whenever $\theta_t > \theta^* + \delta$.*

This proposition provides a stark contrast to learning in a fixed game. If instead of varying in each period the state θ is fixed over all periods, then the learning process reduces to standard fictitious play (as long as $g^i(\theta, \theta) > 0$). For any θ outside of the dominance regions, there are multiple long-run learning outcomes that depend on the initial strategies used by players in the game. For instance, if all players are initially coordinated on one of the two actions, then they continue to choose this action in every period. In contrast, Proposition 1 indicates that extrapolation from different past states may lead to a unique long-run outcome in many of these states θ, independent of the initial strategies players use in the learning process.

The following proof draws on techniques from the proofs of Propositions 2.1 and 2.2 in Morris and Shin (2003) for the corresponding result in global games. The first part of the proof characterizes contagion away from the small-τ limit. As in global games away from the small-noise limit, the dominant actions spread from the dominance regions, but if the distribution of states is not uniform, the contagion from above and below may not meet at a unique threshold. The second part of the proof shows that the intermediate region with multiple learning outcomes collapses to a single point in the small-τ limit exactly as in global games with vanishing noise. Intuitively, as τ becomes small, the distribution of states becomes locally uniform.

Proof of Proposition 1. Define $m_\tau(\theta, k)$ to be the expected payoff to action 1 for type θ in the modified game when the opponents play a threshold strategy with threshold k.

That is

$$m_\tau(\theta,k) \equiv \frac{\int_{\underline{\theta}}^{k}\phi(\theta')\frac{1}{\tau}g\left(\frac{\theta'-\theta}{\tau}\right)u(\theta,v(\theta',0))\,d\theta'+\int_{k}^{\overline{\theta}}\phi(\theta')\frac{1}{\tau}g\left(\frac{\theta'-\theta}{\tau}\right)u(\theta,v(\theta',I-1))\,d\theta'}{\int_\Theta \phi(\tilde{\theta})\frac{1}{\tau}g\left(\frac{\tilde{\theta}-\theta}{\tau}\right)d\tilde{\theta}}. \quad (6)$$

First, we prove that action 0 is serially dominated for $\theta > \overline{\theta}^*$ and action 1 is serially dominated for $\theta < \underline{\theta}^*$, where $\overline{\theta}^*$ and $\underline{\theta}^*$ are, respectively, the maximal and minimal roots of $m_\tau(\theta,\theta)=0$.[10] Note that the function $m_\tau(\theta,k)$ is continuous and decreasing in k. Moreover, for sufficiently small τ, the existence of dominance regions (Assumption A7) implies that $m_\tau(\theta,k)$ is negative for small enough values of θ and positive for large enough values.

Let $\overline{\theta}_0=\overline{\theta}$, and for $k=1,2,\ldots$, recursively define $\overline{\theta}_k$ to be the maximal solution to the equation

$$m_\tau(\theta,\overline{\theta}_{k-1})=0.$$

Let S_k denote the set of strategies remaining for each player after k rounds of deletion of dominated strategies. We prove by induction that action 0 is dominated for all types of each player above $\overline{\theta}_k$ against profiles of strategies from the set S_{k-1}. Suppose that the claim holds for $k-1$. By Assumption A6, if opponents play strategies in S_{k-1}, then the payoff to action 1 for any type θ is at least as large as if every opponent played a cut-off strategy with threshold $\overline{\theta}_{k-1}$ (i.e. a strategy choosing action 0 at any type below $\overline{\theta}_{k-1}$ and action 1 at any type above $\overline{\theta}_{k-1}$). Hence the expected payoff for action 1 at θ is at least $m_\tau(\theta,\overline{\theta}_{k-1})$ regardless of which strategies from S_{k-1} are chosen by the opponents. This expected payoff must be positive above the maximal root $\overline{\theta}_k$ because $m_\tau(\theta,\cdot)$ is continuous everywhere and positive for sufficiently large θ. Therefore, action 0 is dominated above $\overline{\theta}_k$, as claimed.

Next, we show by induction that $(\overline{\theta}_k)_{k=1}^{\infty}$ is a nonincreasing sequence. Note first that $\overline{\theta}_1\le\overline{\theta}_0$ trivially because $\overline{\theta}_0$ lies at the upper boundary of Θ. Suppose that $\overline{\theta}_{k-1}\le\overline{\theta}_{k-2}$. Then $m_\tau(\theta,\overline{\theta}_{k-1})\ge m_\tau(\theta,\overline{\theta}_{k-2})$ because $m_\tau(\theta,k)$ decreases in k, and hence the maximal root of $m_\tau(\theta,\overline{\theta}_{k-1})=0$ must be weakly smaller than that of $m_\tau(\theta,\overline{\theta}_{k-2})=0$, which establishes the induction step.

The nonincreasing sequence $(\overline{\theta}_k)_{k=1}^{\infty}$ converges to some $\overline{\theta}^*$ which, from the continuity of m_τ, must be a solution to $m_\tau(\theta,\theta)=0$. Therefore, action 0 is indeed serially dominated at every type above $\overline{\theta}^*$. The symmetric argument from below establishes that action 1 is serially dominated below the minimal solution $\underline{\theta}^*$ of $m_\tau(\theta,\theta)=0$.

Note that since

$$\int_{\theta-\varepsilon}^{\theta+\varepsilon}\frac{1}{\tau}g\left(\frac{\theta'-\theta}{\tau}\right)d\theta'=\int_{-\varepsilon/\tau}^{\varepsilon/\tau}g(z)\,dz,$$

[10] One can alternatively prove this statement by applying Theorem 5 of Milgrom and Roberts (1990) to the *ex ante* game (in which heterogeneous priors are no longer an issue). However, the direct argument given here better illustrates the process of contagion, and avoids technical issues that arise in moving to the *ex ante* game.

given any $\delta > 0$ and $\varepsilon > 0$, there exists some $\overline{\tau} > 0$ such that for all $\tau \in (0, \overline{\tau})$,

$$\int_{\theta-\varepsilon}^{\theta+\varepsilon} \frac{1}{\tau} g\left(\frac{\theta' - \theta}{\tau}\right) d\theta' > 1 - \delta.$$

In particular, for any function ψ that is continuous at θ, we have

$$\lim_{\tau \to 0} \int_{\Theta} \psi(\theta') \frac{1}{\tau} g\left(\frac{\theta' - \theta}{\tau}\right) d\theta' = \psi(\theta), \tag{7}$$

and similarly

$$\lim_{\tau \to 0} \int_{-\infty}^{\theta} \psi(\theta') \frac{1}{\tau} g\left(\frac{\theta' - \theta}{\tau}\right) d\theta' = \psi(\theta) G(0) \tag{8}$$

and

$$\lim_{\tau \to 0} \int_{\theta}^{+\infty} \psi(\theta') \frac{1}{\tau} g\left(\frac{\theta' - \theta}{\tau}\right) d\theta' = \psi(\theta)(1 - G(0)). \tag{9}$$

Moreover the convergence of the limits in (7)–(9) is uniform over θ in some set X as long as the function $\psi(\theta)$ is uniformly continuous on X.

Applying (7)– (9) to the definition of $m_\tau(\theta, \theta)$ from (6) gives

$$\begin{aligned} \lim_{\tau \to 0} m_\tau(\theta, \theta) &= G(0) u(\theta, v(\theta, 0)) + (1 - G(0)) u(\theta, v(\theta, I-1)) \\ &= G(0) U(\theta, 0) + (1 - G(0)) U(\theta, I-1) \end{aligned}$$

on the open interval $(\underline{\theta}, \overline{\theta})$. The convergence is uniform on any compact subinterval of Θ since $\phi(\theta)$, $U(\theta, 0)$, and $U(\theta, I-1)$ are uniformly continuous on compact sets. We can choose such a compact subinterval $\overline{\Theta}$ of $(\underline{\theta}, \overline{\theta})$ to intersect with both dominance regions, so that all roots of $m_\tau(\theta, \theta) = 0$ must lie in $\overline{\Theta}$. Define $m(\theta) \equiv \lim_{\tau \to 0} m_\tau(\theta, \theta)$ for $\theta \in \overline{\Theta}$. Given any neighborhood N of the unique root θ^* of $m(\theta)$, there exists some $\varepsilon > 0$ such that $m(\theta)$ is uniformly bounded away from 0 by ε outside of N. Choosing $\overline{\tau} > 0$ small enough so that whenever $\tau < \overline{\tau}$, $m_\tau(\theta, \theta)$ is within ε of $m(\theta)$ everywhere on $\overline{\Theta}$ guarantees that $m_\tau(\theta, \theta)$ has no root in $\overline{\Theta} \setminus N$. □

The uniqueness of the learning outcome results from a process of contagion driven by the use of similarity in learning. The structure of the similarity function implicitly precludes discontinuities that could block the contagion process. For example, if there exists some state $\tilde{\theta}$ such that all players assign zero similarity to states lying on opposite sides of $\tilde{\theta}$, then no action can spread contagiously across $\tilde{\theta}$. Such discontinuities are exogenously assumed by Jehiel (2005), who models similarity as a partition of the state space. As noted above, the continuity of our similarity measure is intended to capture players' imperfect knowledge of the structure of the game. Note, however, that contagion may occur even with Jehiel's discontinuous similarity if players do not use the same similarity classes. If similarity classes differ, the spread of an action within an element of one player's similarity partition can cause the same action to spread across a boundary of another's partition.

Since our results focus on the long-run outcomes of learning, a natural question is whether convergence occurs sufficiently quickly for these outcomes to be relevant. Convergence of behavior is not uniform across states, and in the coordination environment studied here, is likely to be particularly slow close to the long-run threshold. However, states very close to the threshold occur only rarely, so convergence to the predicted behavior in most periods may occur relatively quickly. As mentioned above, Carlsson has simulated a bounded-memory variant of strategy-based learning in an environment similar to the one in Section 2. He finds approximate convergence on the order of hundreds of rounds even in the worst-case scenario in which initial play is biased entirely in favor of one action.[11] With unbiased initial strategies, convergence is generally faster.

Theorem 1 identifies a formal parallel between contagion through learning and contagion through incomplete information in the modified game. This connection explains in part why many features of the two kinds of contagion appear similar. However, the information structure of the modified game is inherently different from that of global games with a common prior. Consequently, important differences arise between the outcomes of contagion through learning and those of contagion in global games.

The equilibrium threshold in the standard binary action global game model is independent of the noise distribution, while the threshold in the similarity learning model depends on the similarity function g (which determines the noise distribution in the modified game). The noise-independence result in global games is driven by the common prior, which generates, in equilibrium, uniform beliefs over l at the threshold type regardless of the noise distribution. With learning by similarity, beliefs over l at the threshold in the modified game depend on g because of the heterogeneity of the priors. Proposition 1 allows us to identify how the learning outcome changes if we vary the similarity function. We say that a similarity function $\tilde{g}$ is *more optimistic* than g if $\tilde{G}(0) < G(0)$. Equation (5) immediately implies the following result.

Corollary 1. *Suppose that the similarity function $\tilde{g}$ is more optimistic than g. Then the threshold θ^* when players learn according to $\tilde{g}$ is less than or equal to that when players learn according to g. That is, in the narrow-similarity limit, players using $\tilde{g}$ coordinate on the Pareto dominant equilibrium for a (weakly) larger set of states than do players using g.*

Izmalkov and Yildiz (2008) obtain a similar result in a global game without common priors in which two players observe private signals $x^i = \theta + \varepsilon^i$ and payoffs are as in Table 1. They define the notion of investor sentiment to be the value $q = \Pr_i(x^{-i} > x^i \mid x^i)$. The unique symmetric equilibrium of their game is characterized by a threshold signal x^* satisfying the indifference condition $x^* + q = 0$ because the threshold type assigns probability q to the event that her opponent invests. Investor sentiment q is $\frac{1}{2}$ under the common prior specification, but can attain any value in $(0,1)$ under non-common priors. The definition of investor sentiment naturally extends to the class of modified

[11] Carlsson's simulated model differs from ours in that there is an initial phase in which players play fixed strategies, which slows learning by making initial beliefs persistent. The results reported here are based on an initial phase of one hundred periods.

games studied in the present section. In this setting, sentiment q is equal to $1-G(0)$, the belief that each player assigns to the true state exceeding her signal. Hence any sentiment $q \in (0,1)$ and any equilibrium threshold θ^* outside the dominance regions can be supported by some similarity function.

While the precise outcome of learning depends on the similarity function, qualitative comparative statics do not. Suppose that the payoffs $U(\theta, l; z)$ depend differentiably on an exogenous parameter z, and that the derivative $(\partial/\partial z)U(\theta, l; z)$ has the same sign for all θ and $l = 0, I-1$. Implicitly differentiating (5) gives the following result.

COROLLARY 2. *For any similarity function g, we have* $\operatorname{sgn}(\partial\theta^*/\partial z) = -\operatorname{sgn}(\partial U/\partial z)$.

One restriction on the similarity function that may seem natural is that θ and θ' should be perceived as similar to the same degree that θ' and θ are.

A9 (Symmetry) For every θ and θ', $g(\theta' - \theta) = g(\theta - \theta')$.

The fact that $G(0) = \frac{1}{2}$ for symmetric similarity functions implies the following result.

COROLLARY 3. *If the similarity function is symmetric, then the threshold θ^* solves*

$$\tfrac{1}{2}U(\theta^*, 0) + \tfrac{1}{2}U(\theta^*, I-1) = 0. \tag{10}$$

Even when similarity is symmetric, the outcome of contagion through learning generally differs from that in global games, where the threshold solves

$$\sum_{l=0}^{I-1} \frac{U(\theta^*, l)}{I} = 0. \tag{11}$$

The solutions of (10) and (11) generally differ if payoffs are not linear in l.

The difference between the threshold indifference conditions (10) and (11) indicates why the learning model requires weaker strategic complementarity than the global game model. The threshold of the learning model is independent of the payoffs $U(\cdot, l)$ for values of l other than 0 and $I-1$. In the modified game, the threshold player places zero probability on intermediate values of l, and thus full monotonicity of $U(\cdot, l)$ with respect to l is not necessary. In global games with common priors, the threshold player has uniform beliefs over l, and equilibrium uniqueness may fail without full strategic complementarity.

6. Related literature

Processes of learning from similar games have been examined in several papers, which typically define similarity by an equivalence relation on a given set of games (Germano 2007, Katz 1996 (Chapter 1), LiCalzi 1995, Mengel 2007). Stahl and Van Huyck (2002) demonstrate learning from similar games experimentally.

Jehiel (2005), Jehiel and Koessler (2008), and Eyster and Rabin (2005) study equilibrium concepts in which players best respond to coarsenings of their opponents' strategies, where the coarsenings arise from aggregation across similar states. These models

focus on interesting deviations from standard equilibrium behavior arising from persistent errors in beliefs. We, on the other hand, focus on the case in which these errors are small, leading to a selection among equilibria.

Carlsson (2004) proposes an evolutionary justification of global games equilibrium using strategy-based learning by similarity. Carlsson offers an informal argument to suggest that the learning process can be approximated by the best-response dynamics of a modified game. Theorem 1 above formalizes this result in terms of long-run outcomes. The outcome of the learning process coincides with the global game prediction in Carlsson's (2004) two-player model. With more than two players, Proposition 1 above indicates that the learning outcome generally shares only the qualitative features of the global game solution, while quantitatively they differ.

Argenziano and Gilboa (2005) study multiplicity of similarity-based learning outcomes in coordination problems. With finitely many states, the long-run outcome depends on historical accidents when games with dominant actions are sufficiently rare.

Milgrom and Roberts (1990) study supermodular games, of which the coordination environment studied here is a special case, and show that only serially undominated strategies are played in the long-run under a large class of adaptive dynamics. These dynamics, however, require that players adjust to the full strategies of their opponents. In games with large state spaces, where play of the game (at most) reveals the actions $s(\theta_t)$ assigned by strategies s to the particular states θ_t that are drawn, such dynamics are difficult to justify. The use of similarity in learning can be seen as generating "close to" adaptive dynamics, as reflected in the modified serially undominated result of Theorem 1.[12]

An alternative approach to learning in certain binary action supermodular games is offered by Beggs (2005). Play almost surely converges to an equilibrium of the game if players follow learning rules that adapt threshold strategies based on payoffs from similar types.

7. Discussion

7.1 *Sources of contagion*

Morris (1997) identifies a formal relationship between contagion across types in incomplete information games and contagion across players in local interaction games. Starting with some incomplete information game, one may reinterpret the types in the game as players situated in various locations. Each of these players interacts with some subset of the population, her *neighbors*, and must choose the same action against all opponents. Payoffs in the local interaction game are obtained by a weighted sum of payoffs from interactions with all neighbors, where the weights correspond to the posterior beliefs over opponents' types in the incomplete information game (see Morris 1997 for details).

[12] In addition, Samuelson and Zhang (1992), Nachbar (1990), and Heifetz et al. (2007) identify classes of learning processes under which players learn not to play serially dominated strategies. However, all three papers assume finite or real-valued strategy spaces. The strategy space in our environment, consisting of functions $s : \Theta \longrightarrow \{0,1\}$, is larger.

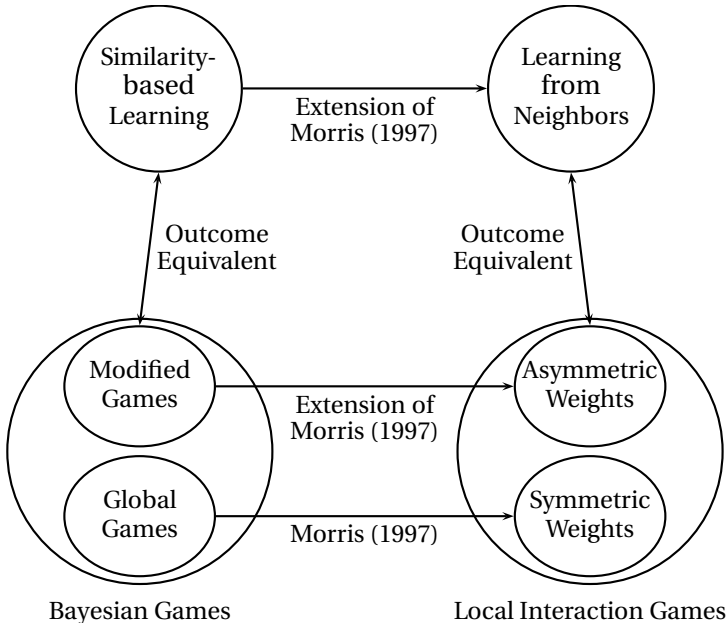


FIGURE 2. Atlas of Contagion.

The model of learning by similarity discussed here may be reinterpreted in the same way. Each player may instead be viewed as a population of players with locations $\theta \in \Theta$. In every period, players at a randomly drawn location are matched to play a game $\Gamma(\theta)$. Players estimate payoffs based on other players' experience at nearby locations. Thus learning by similarity corresponds to learning from neighbors. Since this is merely a formal reinterpretation, the modified game result also holds in this setting. As a game of incomplete information, the modified game may be reinterpreted according to Morris (1997) as a local interaction game. The only difference from the usual case is that the heterogeneous priors in the modified game correspond to *asymmetric* weighting of payoffs in the corresponding local interaction game; thus, for instance, player i's payoff may depend on player j's action even if player j's payoff does not depend on player i's action.

The formal connections among the three sources of contagion described here are summarized in Figure 2. Contagion through learning is related to contagion in Bayesian games through the equivalence of outcomes with the modified game (Theorem 1). The modified games that arise in this way differ from global games because of heterogeneous priors. Each of these may be reinterpreted according to Morris (1997) as local interaction games, with heterogeneous priors corresponding to asymmetric weights and common priors to symmetric weights. Learning by similarity may also be reinterpreted directly as learning from neighbors in local interaction, where the modified game result describes an equivalence of outcomes with certain local interaction games with asymmetric payoff weights.

7.2 *Learning with incomplete information*

An earlier version of this paper (Steiner and Stewart 2007) considers learning by similarity in games with incomplete information. The environment is close to that of Section 5, except that each player receives only a noisy signal $x_t^i = \theta_t + \sigma\varepsilon_t^i$ of the state θ_t in each period t. Players estimate payoffs based on payoffs from similar past types. The modified game result of Theorem 1 extends naturally to this setting, and may be used to demonstrate that there is a unique outcome of learning when both σ and τ are small (i.e. the noise in signals is small and players use narrow similarity functions). This outcome depends on the ratio σ/τ. If σ is small relative to τ, then we recover the complete information learning outcome of Proposition 1. If σ is large relative to τ, then we recover the usual global game solution.

8. Conclusion

The theory of global games has shown that relaxing the common knowledge assumption in games can lead to a process of contagion that generates a unique selection among multiple equilibria. This paper identifies a similar effect that arises under learning if we relax the assumption that players learn from repeated play in exactly the same game. Moreover, the learning outcome is formally related to the equilibrium of a global game with subjective priors, which we call the modified game. While the connection to the modified game is very general, the set of learning outcomes may be difficult to identify in games outside the coordination environment studied here. The unique outcome of learning in this environment relies on the dominance solvability of the modified game. In more general settings, learning outcomes correspond to rationalizable profiles of the game when beliefs are perturbed in a particular way that depends on the similarity function. In a different setting, Weinstein and Yildiz (2007) show that for any finite type, given any rationalizable action, there exists a perturbation of beliefs in the universal type space for which this action is uniquely rationalizable. A natural question, then, is whether the corresponding result holds under learning by similarity in general classes of games; in other words, given any state in the original game, whether any equilibrium may be uniquely selected by an appropriate choice of similarity function.

Appendix

Lemma A.1. *For any $\varepsilon > 0$, there exists some $\delta > 0$ such that changing the opponents' strategies on a set of type profiles of Lebesgue measure at most δ changes the expected payoff of every type of player i from each action a^i by at most ε.*

Proof. Denote i's expected payoff from action a^i at type θ against the profile s^{-i} by

$$\tilde{U}^i(\theta, a^i, s^{-i}) = \frac{\int_\Theta u^i(\theta, a^i, v^i(\theta', a^i, s^{-i}(\theta')))\phi(\theta')g^i(\theta', \theta)\,d\theta'}{\int_\Theta \phi(\tilde{\theta})g^i(\tilde{\theta}, \theta)\,d\tilde{\theta}}.$$

The denominator is bounded above zero because it is continuous and positive by Assumption A3, and hence attains a positive minimum on the compact set Θ. Recall that

the functions u^i, g^i, and ϕ are bounded by assumption. Hence there exists a constant K such that if s^{-i} changes only on a set of measure δ, then the numerator changes by at most $K\delta$. □

LEMMA A.2. *Fix a profile* $\mathbf{a}$ *and an arbitrary action* a^i. *For any* $\delta > 0$, *there exists some* $\pi > 0$ *such that the set of types of player* i *for which action* a^i *is dominated but not* π*-dominated under the profile* $\mathbf{a}$ *has measure at most* δ.

PROOF. Consider any decreasing sequence $\pi_1, \pi_2, \ldots$ such that $\lim_{n\to\infty} \pi_n = 0$. Let $\Theta(n)$ denote the set of types for which action a^i is π_n-dominated under $\mathbf{a}$, and let $\overline{\Theta}$ denote the set of types for which a^i is dominated under $\mathbf{a}$. Then $\Theta(n)$ is a monotone sequence of sets, and it suffices to show that $\lim_{n\to\infty} \Theta(n) = \overline{\Theta}$.

Suppose for contradiction that $\overline{\Theta} \backslash \lim_{n\to\infty} \Theta(n)$ contains some type θ. Then there exists some action $a^{i\prime}$ that dominates a^i at θ under the profile $\mathbf{a}$, but does not π-dominate a^i at θ under $\mathbf{a}$ for any $\pi > 0$. Hence we have

$$\inf_{s^{-i} \in \mathbf{a}^{-i}} \tilde{U}^i(\theta, a^{i\prime}, s^{-i}) - \tilde{U}^i(\theta, a^i, s^{-i}) = 0,$$

where we abuse notation by writing $s^{-i} \in \mathbf{a}^{-i}$ to mean that s^{-i} is consistent with $\mathbf{a}^{-i}$. Define a strategy profile $\overline{s}^{-i}$ by choosing

$$\overline{s}^{-i}(\theta') \in \operatorname*{argmin}_{a^{-i} \in \mathbf{a}^{-i}(\theta')} u^i(\theta, a^{i\prime}, v^i(\theta', a^{i\prime}, a^{-i})) - u^i(\theta, a^i, v^i(\theta', a^i, a^{-i}))$$

for each θ'. The profile $\overline{s}^{-i}$ is consistent with $\mathbf{a}^{-i}$, and satisfies

$$\tilde{U}^i(\theta, a^{i\prime}, \overline{s}^{-i}) - \tilde{U}^i(\theta, a^i, \overline{s}^{-i}) = \inf_{s^{-i} \in \mathbf{a}^{-i}} \tilde{U}^i(\theta, a^{i\prime}, s^{-i}) - \tilde{U}^i(\theta, a^i, s^{-i}) = 0,$$

contradicting that $a^{i\prime}$ dominates a^i at θ under $\mathbf{a}$. □

LEMMA A.3. *For any* k *and* θ, *we have* $\mathbf{a}^i_{k,\pi}(\theta) \subseteq \mathbf{a}^i_{k,\pi'}(\theta)$ *whenever* $\pi \leq \pi'$.

PROOF. Note that the statement is trivial for $k = 0$. Suppose for induction that the statement holds for k (for all θ). We need to show that if a^i is π'-dominated at θ under $\mathbf{a}_{k,\pi'}$ then a^i is π-dominated at θ under $\mathbf{a}_{k,\pi}$. Accordingly, suppose that a^i is π'-dominated at θ under $\mathbf{a}_{k,\pi'}$; that is, there exists $a^{i\prime} \in \mathbf{a}^i_{k,\pi'}$ for which

$$\tilde{U}(\theta, a^{i\prime}, s^{-i}) - \tilde{U}(\theta, a^i, s^{-i}) > \pi' \text{ for all } s^{-i} \text{ consistent with } \mathbf{a}^{-i}_{k,\pi'}. \tag{12}$$

Since, by the inductive hypothesis, we have $\mathbf{a}^{-i}_{k,\pi} \subseteq \mathbf{a}^{-i}_{k,\pi'}$, (12) implies

$$\tilde{U}(\theta, a^{i\prime}, s^{-i}) - \tilde{U}(\theta, a^i, s^{-i}) > \pi \text{ for all } s^{-i} \text{ consistent with } \mathbf{a}^{-i}_{k,\pi}. \tag{13}$$

If $a^{i\prime} \in \mathbf{a}^i_{k,\pi}(\theta)$, then we are done. Otherwise, there exists some $a^{i\prime\prime} \in \mathbf{a}^i_{k,\pi}(\theta)$ such that $a^{i\prime\prime}$ dominates $a^{i\prime}$ at θ under the profile $\mathbf{a}_{k,\pi}$. Thus we have

$$\tilde{U}(\theta, a^{i\prime\prime}, s^{-i}) - \tilde{U}(\theta, a^{i\prime}, s^{-i}) > 0 \text{ for all } s^{-i} \text{ consistent with } \mathbf{a}^{-i}_{k,\pi}.$$

Combining this with (13) gives the result. □

PROOF OF LEMMA 1. Given any type θ for which $a^i \in \mathbf{a}^i_{k-1,0} \setminus \mathbf{a}^i_{k,0}$, there exists some $\pi(\theta) > 0$ such that, against any profile s^{-i} consistent with $\mathbf{a}^{-i}_{k-1,0}$, the expected payoff for some action $a^{i\prime} \in \mathbf{a}^i_{k,0}(\theta)$ is at least $\pi(\theta)$ greater than that for action a^i. By Lemma A.3, we have $a^{i\prime} \in \mathbf{a}^i_{k,\pi}(\theta)$ for all $\pi > 0$, and hence it suffices to show that $a^{i\prime}$ dominates a^i under the profile $\mathbf{a}_{k-1,\pi}$ for some $\pi > 0$. By Lemma A.1, it suffices to show that given any $\delta > 0$, there exists some $\pi > 0$ small enough so that, for any player i, $\mathbf{a}^i_{k-1,\pi}$ differs from $\mathbf{a}^i_{k-1,0}$ on a set of measure at most δ.

We proceed by induction. The result is trivial for $k = 1$. For $k > 1$, assume for induction that the result is true for $k-1$; that is, assume that for any $\delta > 0$, there exists some $\pi > 0$ for which $\mathbf{a}^i_{k-2,\pi}$ differs from $\mathbf{a}^i_{k-2,0}$ on a set of measure at most δ.

By Lemma A.2, given $\delta > 0$, we can choose $\pi' > 0$ small enough so that the set of types of player i for which the action a^i is dominated but not π'-dominated under $\mathbf{a}_{k-2,0}$ has measure at most δ. By Lemma A.1, starting from any strategy profile, there exists some $\delta' > 0$ such that changing the actions of at most a measure of δ' of the opponents' types changes the expected payoff of each type of player i by at most $\pi'/4$. By the inductive hypothesis, we can choose $\pi'' > 0$ such that $\mathbf{a}^{-i}_{k-2,\pi''}$ differs from $\mathbf{a}^{-i}_{k-2,0}$ on a set of types of measure at most δ'. Consider $\pi = \min\{\pi'/2, \pi''\}$. We need to show that $\mathbf{a}^i_{k-1,\pi}$ differs from $\mathbf{a}^i_{k-1,0}$ on a set of types of measure at most δ.

Consider any type θ and actions $a^i, a^{i\prime} \in \mathbf{a}^{-i}_{k-2,0}(\theta)$. By Lemma A.3, a^i and $a^{i\prime}$ also belong to $\mathbf{a}^{-i}_{k-2,\pi}(\theta)$ for all $\pi > 0$. Also by Lemma A.3, $\mathbf{a}^{-i}_{k-2,\pi}(\theta') \subseteq \mathbf{a}^{-i}_{k-2,\pi''}(\theta')$ for all θ'. Therefore, we have

$$\sup_{s^{-i} \in \mathbf{a}^{-i}_{k-2,\pi}} [\tilde{U}(\theta, a^i, s^{-i}) - \tilde{U}(\theta, a^{i\prime}, s^{-i})] \leq \sup_{s^{-i} \in \mathbf{a}^{-i}_{k-2,\pi''}} [\tilde{U}(\theta, a^i, s^{-i}) - \tilde{U}(\theta, a^{i\prime}, s^{-i})], \quad (14)$$

where, as above, we write $s^{-i} \in \mathbf{a}^{-i}_{k,\pi}$ to mean that the strategy profile s^{-i} is consistent with $\mathbf{a}^{-i}_{k,\pi}$. By the definition of π'', we have

$$\sup_{s^{-i} \in \mathbf{a}^{-i}_{k-2,\pi''}} [\tilde{U}(\theta, a^i, s^{-i}) - \tilde{U}(\theta, a^{i\prime}, s^{-i})] \leq \sup_{s^{-i} \in \mathbf{a}^{-i}_{k-2,0}} [\tilde{U}(\theta, a^i, s^{-i}) - \tilde{U}(\theta, a^{i\prime}, s^{-i}) + \tfrac{1}{2}\pi']. \quad (15)$$

If action a^i is π'-dominated by $a^{i\prime}$ for type θ of player i under $\mathbf{a}_{k-2,0}$, then

$$\sup_{s^{-i} \in \mathbf{a}^{-i}_{k-2,0}} [\tilde{U}(\theta, a^i, s^{-i}) - \tilde{U}(\theta, a^{i\prime}, s^{-i}) + \tfrac{1}{2}\pi'] < -\tfrac{1}{2}\pi'.$$

Combining (14) and (15) gives the following: if action a^i is π'-dominated for i at type θ under $\mathbf{a}_{k-2,0}$, then a^i must be $\pi'/2$-dominated, and hence also π-dominated under $\mathbf{a}_{k-2,\pi}$. Therefore, if, at some θ, a^i is dominated under $\mathbf{a}_{k-2,0}$ but not π-dominated under $\mathbf{a}_{k-2,\pi}$, then a^i is dominated under $\mathbf{a}_{k-2,0}$ but not π'-dominated under $\mathbf{a}_{k-2,0}$. The latter can happen only on a set of types of measure δ.

We have shown that the set of types θ for which $a^i \in \mathbf{a}^i_{k-2,\pi}(\theta)$ but $a^i \notin \mathbf{a}^i_{k-2,0}(\theta)$ has measure at most δ. The result now follows since the number of players and the number of actions are both finite. □

PROOF OF THEOREM 1. (i) Assume for induction that there almost surely exists some period after which the strategies s^i_t are consistent with $\mathbf{a}^i_{k-1,\pi}$.

In the first step, we consider payoff estimates at a fixed state θ^*. Suppose that action a^i is π-dominated by some action $a^{i\prime}$ for type θ^* of player i in the modified game under the profile $\mathbf{a}_{k-1,\pi}$. Let

$$\pi'(\theta^*) = \pi \int_\Theta g^i(\theta, \theta^*)\, d\Phi(\theta).$$

We show that, under the learning process, there almost surely exists some period after which the estimated payoff to action $a^{i\prime}$ at θ^* exceeds that to action a^i by at least π. This is the case if

$$\frac{1}{t}\sum_{s<t}\left[u^i(\theta^*, a^{i\prime}, v^i(\theta_s, a^{i\prime}, a^{-i}_s)) - u^i(\theta^*, a^i, v^i(\theta_s, a^i, a^{-i}_s))\right] g^i(\theta_s, \theta^*) > \pi'(\theta^*). \tag{16}$$

For each θ, θ' and a^{-i}, let

$$\Delta(\theta, \theta', a^{-i}) = u^i(\theta, a^{i\prime}, v^i(\theta', a^{i\prime}, a^{-i})) - u^i(\theta, a^i, v^i(\theta', a^i, a^{-i})).$$

Keeping θ^* fixed, choose the strategy profile $s^{-i}_{\min}(\theta)$ to minimize the payoff advantage of $a^{i\prime}$ over a^i at θ^*; that is,

$$s^{-i}_{\min}(\theta) \in \operatorname*{arg\,min}_{a^{-i} \in \mathbf{a}^{-i}_{k-1,\pi}(\theta)} \Delta(\theta^*, \theta, a^{-i}).$$

Define a random variable

$$X = \Delta(\theta^*, \theta, s^{-i}_{\min}(\theta)) g^i(\theta, \theta^*),$$

where the distribution of X is induced from the distribution Φ of θ.

By the inductive hypothesis, opponents play actions in $\mathbf{a}^{-i}_{k-1,\pi}(\theta_s)$ in every period $s \geq t_0$ for some t_0. For large enough $t > t_0$, periods up to t_0 receive an arbitrarily small weight in each player's payoff estimates. Thus assuming that $a^{-i}_s \in \mathbf{a}^{-i}_{k-1,\pi}(\theta_s)$ for all s introduces only an arbitrarily small error in the payoff estimates. Note that for any history $(\theta_s, a^i_s, a^{-i}_s)^{t-1}_{s=1}$ in which $a^{-i}_s \in \mathbf{a}^{-i}_{k-1,\pi}(\theta_s)$, we have

$$\frac{1}{t}\sum_{s<t}\Delta(\theta^*, \theta_s, a^{-i}_s) g^i(\theta_s, \theta^*) \geq \frac{1}{t}\sum_{s<t}\Delta(\theta^*, \theta_s, s^{-i}_{\min}(\theta_s)) g^i(\theta_s, \theta^*),$$

so it suffices to prove that (16) holds when $a^{-i}_s = s^{-i}_{\min}(\theta_s)$ for every s.

By the Strong Law of Large Numbers, the weighted payoff difference

$$\frac{1}{t}\sum_{s<t}\Delta(\theta^*, \theta_s, s^{-i}_{\min}(\theta_s)) g^i(\theta_s, \theta^*)$$

almost surely tends to the expectation of X, that is, to

$$\int_{\Theta} \Delta(\theta^*, \theta, s^{-i}_{\min}(\theta)) g^i(\theta, \theta^*)\, d\Phi(\theta).$$

By the assumption that $a^{i\prime}$ π-dominates a^i, the last expression is greater than $\pi'(\theta^*)$. Therefore, there almost surely exists some period T such that (16) holds for every $t > T$, as desired.

In the second step, we show that there exists some $\delta > 0$ such that if (16) holds at θ^*, then $s^i_t(\theta) \neq a^i$ for all $\theta \in (\theta^* - \delta, \theta^* + \delta)$. Let

$$\pi' = \inf_{\theta \in \Theta} \pi'(\theta).$$

By Assumption A3, $\pi'(\theta)$ is positive everywhere, and since it is continuous, the compactness of Θ guarantees that π' is bounded away from zero.

Define the function

$$\begin{aligned} k(h_t, \theta) &= \frac{1}{t} \sum_{s<t} \left[u^i(\theta, a^{i\prime}, v^i(\theta_s, a^{i\prime}, a^{-i}_s)) - u^i(\theta, a^i, v^i(\theta_s, a^i, a^{-i}_s)) \right] g^i(\theta_s, \theta) \\ &= \frac{1}{t} \sum_{s<t} \Delta(\theta, \theta_s, a^{-i}_s) g^i(\theta_s, \theta) \end{aligned}$$

for finite histories $h_t = (\theta_s, a^i_s, a^{-i}_s)_{s<t}$ and states θ. Player i does not choose action a^i at state θ following history h_t if $k(h_t, \theta) > 0$. Thus the second step will be complete if we show that, for some $\delta > 0$, after any history h_t, we have

$$\left| k(h_t, \theta) - k(h_t, \theta') \right| < \pi'$$

whenever $|\theta - \theta'| < \delta$.

By Assumption A4, $\Delta(\theta, \theta_s, a^{-i}_s) g^i(\theta_s, \theta)$ is uniformly continuous in θ over all θ_s and a^{-i}_s. Hence the average $k(h_t, \theta)$ is also continuous in θ uniformly over all values of θ and all histories h_t, as needed.

Finally, partition the set Θ into a finite number of subsets $\Theta_1, \ldots, \Theta_m$, each of diameter less than δ (where δ is chosen given π' as in the second step above). Consider any of these subsets Θ_l. If Θ_l contains some θ^* at which a^i is π-dominated by $a^{i\prime}$ (under some profile), then by the first step, $k(h_t, \theta^*)$ is eventually larger than π'. By the second step, $k(h_t, \theta)$ is therefore positive for *all* types in Θ_l. Hence there almost surely exists some period T_l after which player i never plays action a^i at *any* state in Θ_l. Since there are only finitely many sets Θ_l, there almost surely exists some T after which player i never plays action a^i at any state θ for which it is serially π-dominated for the corresponding type in the modified game. The result now follows from the finiteness of the action and player sets.

(ii) First we claim that for any k, the probability that play is consistent with k rounds of IEDS in the modified game approaches one as time tends to infinity. In the proof of Lemma 1, we show that for any $\delta > 0$, there exists some $\pi > 0$ such that the set of

types θ at which k rounds of IEDS differ from k rounds of IEπDS has measure at most δ. Thus the probability that θ_t lies in this set can be made arbitrarily small by choosing π to be sufficiently small. Note that outside of this set, play under the learning process is almost surely eventually consistent with k rounds of IEDS by part (i) of the theorem. This proves the claim.

For each $k = 1,2,\ldots$, denote by Θ_k^i the set of types of player i for which all serially dominated actions are eliminated within the first k rounds of IEDS. Let $\Theta_k = \bigcap_{i=1,\ldots,I} \Theta_k^i$. The sequence of sets Θ_k is nondecreasing in k, and converges to the set Θ. Hence the measure of the set $\Theta \setminus \Theta_k$ converges to zero as k tends to infinity. From the previous paragraph, the probability that play is consistent with IEDS on Θ_k tends to one over time, and the probability that the state θ_t lies in $\Theta \setminus \Theta_k$ can be made arbitrarily small by choosing k to be sufficiently large. □

References

Argenziano, Rossella and Itzhak Gilboa (2005), "History as a coordination device." Unpublished paper, Department of Economics, University of Essex. [450]

Beggs, Alan (2005), "Learning in Bayesian games with binary actions." Economics Series Working Paper 232, University of Oxford, Department of Economics. [450]

Carlsson, Hans (2004), "Rational and adaptive play in two-person global games." Unpublished paper, Department of Economics, Lund University. [443, 450]

Carlsson, Hans and Eric van Damme (1993), "Global games and equilibrium selection." *Econometrica*, 61, 989–1018. [433, 434, 436]

Eyster, Erik and Matthew Rabin (2005), "Cursed equilibrium." *Econometrica*, 73, 1623–1672. [449]

Frankel, David M., Stephen Morris, and Ady Pauzner (2003), "Equilibrium selection in global games with strategic complementarities." *Journal of Economic Theory*, 108, 1–44. [444]

Germano, Fabrizio (2007), "Stochastic evolution of rules for playing finite normal form games." *Theory and Decision*, 62, 311–333. [449]

Gilboa, Itzhak and David Schmeidler (2001), *A Theory of Case-Based Decisions*. Cambridge University Press, Cambridge. [433]

Heifetz, Aviad, Chris Shannon, and Yossi Spiegel (2007), "The dynamic evolution of preferences." *Economic Theory*, 32, 251–286. [450]

Izmalkov, Sergei and Muhamet Yildiz (2008), "Investor sentiments." Unpublished paper, Department of Economics, Massachusetts Institute of Technology. [448]

Jehiel, Philippe (2005), "Analogy-based expectation equilibrium." *Journal of Economic Theory*, 123, 81–104. [447, 449]

Jehiel, Philippe and Frederic Koessler (2008), "Revisiting games of incomplete information with analogy-based expectations." *Games and Economic Behavior*, 62, 533–557. [449]

Katz, Kimberly F. (1996), *Three Applications of Game Theory*. Ph.D. thesis, University of Pennsylvania. [449]

LiCalzi, Marco (1995), "Fictitious play by cases." *Games and Economic Behavior*, 11, 64–89. [449]

Mengel, Friederike (2007), "Learning across games." Working Paper WP-AD 2007-05, Instituto Valenciano de Investigaciones Económicas. [449]

Milgrom, Paul and John Roberts (1990), "Rationalizability, learning, and equilibrium in games with strategic complementarities." *Econometrica*, 58, 1255–1277. [446, 450]

Morris, Stephen (1997), "Interaction games: A unified analysis of incomplete information, local interaction and random matching." Working Paper 97-08-072, Santa Fe Institute. [450, 451]

Morris, Stephen (2000), "Contagion." *Review of Economic Studies*, 67, 57–78. [433]

Morris, Stephen and Hyun Song Shin (2003), "Global games: Theory and applications." In *Advances in Economics and Econometrics, Volume 1* (Mathias Dewatripont, Lars Peter Hansen, and Stephen J. Turnovsky, eds.), 56–114, Cambridge University Press, Cambridge. [434, 445]

Nachbar, John H. (1990), "'Evolutionary' selection dynamics in games: Convergence and limit properties." *International Journal of Game Theory*, 19, 59–89. [450]

Samuelson, Larry and Jianbo Zhang (1992), "Evolutionary stability in asymmetric games." *Journal of Economic Theory*, 57, 363–391. [450]

Stahl, Dale O. and John Van Huyck (2002), "Learning conditional behavior in similar stag hunt games." Unpublished paper. [449]

Steiner, Jakub and Colin Stewart (2007), "Learning by similarity in coordination problems." Working Paper 324, Center for Economic Research and Graduate Education, Economics Institute, Charles University. [452]

Weinstein, Jonathan and Muhamet Yildiz (2007), "A structure theorem for rationalizability with application to robust predictions of refinements." *Econometrica*, 75, 365–400. [452]

Submitted 2007-12-14. Final version accepted 2008-8-11. Available online 2008-8-20.