Itai Arieli and H. Peyton Young

Stochastic learning dynamics and speed of convergence in population games

Article (Published version) (Refereed)

Original citation:
Arieli, Itai and Young, H. Peyton (2016) Stochastic learning dynamics and speed of convergence in population games. Econometrica, 84 (2). pp. 627-676. ISSN 0012-9682

DOI: 10.3982/ECTA10740

Econometrica, Vol. 84, No. 2 (March, 2016), 627–676

STOCHASTIC LEARNING DYNAMICS AND SPEED OF CONVERGENCE IN POPULATION GAMES

BY ITAI ARIELI AND H. PEYTON YOUNG[1]

We study how long it takes for large populations of interacting agents to come close to Nash equilibrium when they adapt their behavior using a stochastic better reply dynamic. Prior work considers this question mainly for 2×2 games and potential games; here we characterize convergence times for general weakly acyclic games, including coordination games, dominance solvable games, games with strategic complementarities, potential games, and many others with applications in economics, biology, and distributed control. If players' better replies are governed by idiosyncratic shocks, the convergence time can grow exponentially in the population size; moreover, this is true even in games with very simple payoff structures. However, if their responses are sufficiently correlated due to aggregate shocks, the convergence time is greatly accelerated; in fact, it is bounded for all sufficiently large populations. We provide explicit bounds on the speed of convergence as a function of key structural parameters including the number of strategies, the length of the better reply paths, the extent to which players can influence the payoffs of others, and the desired degree of approximation to Nash equilibrium.

KEYWORDS: Population games, better reply dynamics, convergence time.

1. OVERVIEW

NASH EQUILIBRIUM IS THE CENTRAL SOLUTION CONCEPT for noncooperative games, but many natural learning dynamics do not converge to Nash equilibrium without imposing strong conditions on the structure of the game and/or the players' level of rationality. Even in those situations where the learning dynamics do eventually lead to Nash equilibrium, the process may take so long that equilibrium is not a meaningful description of the players' behavior. In this paper, we study the convergence issue for population games, that is, games that are played by a large number of interacting players. These games have numerous applications in economics, biology, and distributed control (Hofbauer and Sigmund (1998), Sandholm (2010b), Marden and Shamma (2014)). Two key questions present themselves: are there natural learning rules that lead to Nash equilibrium for a reasonably general class of games? If so, how long does it take to approximate Nash equilibrium behavior starting from arbitrary initial conditions?

To date, the literature has focused largely on negative results. It is well known, for example, that there are no natural deterministic dynamics that converge to Nash equilibrium in general normal-form games. The basic dif-

[1]The authors thank Yakov Babichenko, Gabriel Kreindler, John Levy, Jason Marden, Bary Pradelski, William Sandholm, and Cheng Wan. We also thank the co-editor and several anonymous referees for constructive comments on an earlier draft. This research was supported by the Air Force Office of Scientific Research Grant # FA9550-09-1-0538 and by the Office of Naval Research Grant # N00014-09-1-0751.

 DOI: 10.3982/ECTA10740

ficulty is that, given virtually any deterministic dynamic, one can construct the payoffs in such a way that the process gets trapped in a cycle (Hofbauer and Swinkels (1996), Hart and Mas-Colell (2003), Hofbauer and Sandholm (2007)). Although stochastic learning algorithms can be designed that select Nash equilibria in the long run, their convergence time will, except in special cases, be very slow due to the fact that the entire space of strategies is repeatedly searched (Hart and Mas-Colell (2006), Foster and Young (2003, 2006), Germano and Lugosi (2007), Marden, Young, Arslan, and Shamma (2009), Young (2009), Marden and Shamma (2012), Babichenko (2012), Pradelski and Young (2012)).

There is, however, an important class of games where positive results hold, namely, games that can be represented by a global potential function. In this case, there are various decentralized algorithms that lead to Nash equilibrium quite rapidly (Shah and Shin (2010), Chien and Sinclair (2011), Kreindler and Young (2013), Borowski, Marden, and Frew (2013), Borowski and Marden (2014)). These results exploit the fact that better replies by individual players lead to monotonic increases in the potential function. A similar approach can be employed if the dynamical process has a Lyapunov function, which is sometimes the case even when the underlying game does not have a potential function (Ellison, Fudenberg, and Imhof (2014)).

Another class of games where positive results have been obtained are games that can be solved by the iterated elimination of strategies that are not contained in the minimal p-dominant set for some $p > 1/2$ (Tercieux (2006)). For such games, Oyama, Sandholm, and Tercieux (2015) showed that if players choose best responses to random samples of other players' actions, and if the distribution of sample sizes places sufficiently high probability on small samples, then the corresponding deterministic dynamics converge in bounded time to the unique equilibrium that results from the iterative elimination procedure.

Finally, rapid convergence can occur when agents are located at the vertices of a network and they respond only to the choices of their neighbors (Ellison (1993), Young (1998, 2011), Montanari and Saberi (2010), Kreindler and Young (2013)). The main focus of this literature is on the extent to which the network topology affects convergence time. However, the analysis is typically restricted to very simple games such as 2×2 coordination games, and it is not known whether the results extend to more general games.[2]

This paper examines the speed of convergence issue for the general case of weakly acyclic games with global interaction. There are numerous examples of such games that are not necessarily potential games, including n-person coordination games, games with strategic complementaries, dominance-solvable games, and many others with important applications in economics, computer science, and distributed control (Fabrikant, Jaggard, and Schapira (2013)). The key feature that these games share with potential games is that, from every

[2]Golub and Jackson (2012) studied the effect of the network topology on the speed with which agents reach a consensus when they update their beliefs based on their neighbors' beliefs.

initial state, there exists a better reply path to some Nash equilibrium. If the players can find such a path through some form of adaptive learning, there is a hope that they can reach an equilibrium (or at least the neighborhood of such an equilibrium) reasonably quickly.

Consider an n-person game G that is played by individuals who are drawn at random from n disjoint populations, each of size N.[3] In applications, G is often a two-person game, in which case pairs of individuals are matched at random to play the game. These are among the most common examples of population games in the literature. Here we shall call them *Nash population games*, since the idea was originally introduced by Nash as a way of motivating Nash equilibrium without invoking full rationality on the part of the players (Nash (1950)). In what follows, we develop a general framework for estimating the speed of convergence as a function of the structure of the underlying game G and the number N of individuals in each population. One of our key findings is that weak acyclicity does not in itself guarantee fast convergence when the population is large and the players' responses are subject to idiosyncratic independent shocks. By contrast, when the shocks are sufficiently correlated, convergence to equilibrium may occur quite rapidly.

For the sake of specificity, we shall focus on the important class of better reply processes variously known as "pairwise comparison revision protocols" or "pairwise comparison dynamics" (Björnerstedt and Weibull (1996), Sandholm (2010a)). These dynamics are very common in the literature, although their micro foundations are often not made explicit. Here we show that they have a natural motivation in terms of idiosyncratic switching costs. Specifically, we consider the following type of process: players revise their strategies asynchronously according to i.i.d. Poisson arrival processes. The arrival rate determines the frequency with which individuals update their strategies and serves as the benchmark against which other rates are measured.[4] When presented with a revision opportunity, a player compares his current payoff to the payoff from an alternative, randomly drawn strategy.[5] The player switches provided the payoff difference is higher than the switching cost, which is modeled as the realization of an *idiosyncratic* random variable. Thus, from an observer's standpoint, the player switches with a probability that is monotonically increasing in the payoff difference between his current strategy and a randomly selected alternative. We shall call such a process a *stochastic pairwise comparison dynamic.* One of the earliest examples of a stochastic pairwise comparison dynamic is a

[3]For a discussion of the origins and significance of this class of games, see Weibull (1995), Leonard (1994), Björnerstedt and Weibull (1996), and Sandholm (2009, 2010a).

[4]This is a standard assumption in the literature; see, among others, Shah and Shin (2010), Marden and Shamma (2012, 2014), Chien and Sinclar (2011), Kreindler and Young (2013). The specific context determines the frequency with which players update in real time.

[5]A variant of this procedure is to choose another player at random, and then to imitate his action with a probability that is increasing in its observed payoff, provided the latter is higher than the player's current payoff (Björnerstedt and Weibull (1996)).

model of traffic flow due to Smith (1984). In this model, drivers switch from one route to another with a probability that is proportional to the payoff difference between them.

Our results may be summarized as follows. Let G be a normal-form n-person game ($n \geq 2$) with a finite number of strategies for each player. Let N be the number of individuals in each of the n player positions, and let G^N be the population game in which the payoff to each individual is the expected payoff from playing against a group drawn uniformly at random from the other populations. The normal-form game G is *weakly acyclic* if, from any given strategy profile, there exists a better reply path to a Nash equilibrium (Young (1993)). However, the fact that G is weakly acyclic does not necessarily imply that G^N is weakly acyclic. This conclusion does hold for a generic set of payoffs defining G, but the usual form of genericity (no payoff ties) is insufficient. We introduce a new concept called *δ-genericity* that proves to be crucial not only for characterizing when weak acyclicity is inherited by the population game, but also for estimating the speed of convergence. This condition is considerably stronger and more delicate than the condition of no payoff ties, but it is still a generic condition, that is, the set of payoffs that satisfy δ-genericity for some $\delta > 0$ has full Lebesgue measure. We call the greatest such δ the *interdependence index* of G, because it measure the extent to which changes of strategy by members of one population can alter the payoffs to members of other populations.

We are interested in the following question: when G and G^N are weakly acyclic, and players update via a stochastic pairwise comparison dynamic, how long does it take for the dynamical system to approximate Nash equilibrium behavior? There are different ways that one can formulate the 'how long does it take' issue. One possibility is to consider the expected first time that the process closely resembles Nash behavior, but this is not satisfactory. The difficulty is that the process might *briefly* resemble a Nash equilibrium, but then move away from it. A more relevant concept is the time that it takes until expected behavior is close to Nash equilibrium over long periods of time. This concept allows for occasional departures from equilibrium (e.g., as the process transits from one equilibrium to another), but such departures must be rare.[6]

Our first main result shows that, under purely idiosyncratic random shocks, the convergence time can grow exponentially with N. Specifically, we construct a three-person game $\tilde{G}$ with a total of eight strategies, such that, for all sufficiently small $\varepsilon > 0$, the convergence time grows exponentially in N. In fact, this is true for a wide class of stochastic better reply dynamics including the stochastic replicator dynamics. This construction shows that results on the speed of

[6]A more demanding concept is the time it takes until the process comes close to Nash equilibrium and remains close in all subsequent periods. As we shall see, rapid convergence in this sense may not be achievable even when shocks are highly correlated.

convergence for potential games do not carry over to weakly acyclic games in general.

This finding complements recent results of Hart and Mansour (2010), who used the theory of communication complexity to construct N-person, weakly acyclic games such that the expected time for any uncoupled better reply dynamic to reach Nash equilibrium grows exponentially in N. Using similar techniques, Babichenko (2014) showed that it can take an exponential number of periods to reach an approximate Nash equilibrium.[7] Our results are quite different because the games constructed by these authors become increasingly complex as the number N of players grows. In our examples, the underlying game G is fixed, the Nash equilibria are trivial to compute, and the only variable is the population size N. Nevertheless, for a large class of better reply dynamics, it takes exponentially long to reach an approximate Nash equilibrium.

The second main contribution of the paper is to show that the speed of convergence can be greatly accelerated when the learning process is subjected to aggregate as well as idiosyncratic shocks. The nature of the aggregate shocks depends on the context. They could represent intermittent breakdowns in communications that temporarily prevent subgroups of players from learning about the payoffs available from alternative strategies. Or they could represent switching costs that make it unprofitable for all players currently using a given strategy i to switch to some alternative strategy j; for example, a strategy might represent the use of a given product, so that the switching cost affects all users of the product simultaneously.[8] Somewhat paradoxically, these shocks, which slow down the players' responses, can greatly reduce the convergence time; in fact, for a general class of shock distributions, the convergence time is bounded above for all sufficiently large N. More generally, we show how the speed of convergence depends on key structural parameters of the underlying game G, including the length of the better reply paths and the total number of strategies.

The plan of the paper is as follows. In Section 2, we introduce the concept of Nash population games, following Weibull's seminal treatment (Weibull (1995)). We also define what we mean by "coming close" to equilibrium. Given a Nash population game, we say that a distribution of behaviors is ε-*close* to Nash equilibrium if it constitutes an ε-equilibrium (in mixed strategies) with respect to the underlying game G. This definition does not require that everyone in the population play an ε-equilibrium, but it does require that, if

[7]Our framework also differs from Sandholm and Staudigl (2015), who studied the rate at which certain types of stochastic learning dynamics converge to the stochastically stable equilibrium. In games with multiple equilibria, the convergence time can also grow exponentially with N.

[8]In a similar spirit, Pradelski (2015) showed that by introducing aggregate shocks to a two-sided matching market, the convergence time becomes polynomial in the number of players, rather than exponential.

some players' behaviors are far from equilibrium, they must constitute a small fraction of the whole population. Section 3 introduces the class of better reply dynamics known as pairwise comparison dynamics (Sandholm (2010b)), which form the basis of most of our results. In Section 4, we exhibit a family of weakly acyclic, three-person Nash population games such that, given any pairwise comparison dynamic and any $\varepsilon > 0$, there exists a game G such that it takes *exponentially long* in the population size N for the learning process to come ε-close to a Nash equilibrium of G for the first time (see Theorem 1). A fortiori, it takes exponentially long to come close to equilibrium for an extended period of time. These games also illustrate a key difference between our approach and the use of mean-field dynamics to approximate the behavior of the stochastic processes for large N. Namely, for all finite N, convergence to a pure Nash equilibrium occurs almost surely in finite time in any of these games, whereas under the limiting mean-field dynamics, convergence may never occur. The reason is that the mean-field better reply dynamics can have multiple attractors, whereas in the population game with finite N there must exist a better reply path from every state to a pure Nash equilibrium.

In Section 5, we introduce aggregate shocks to the learning process, and show that, given any finite, weakly acyclic game G with generic payoffs, and any $\varepsilon > 0$, the convergence time in the population game G^N is bounded above for all sufficiently large population sizes N. In Section 6, we estimate the convergence time as a function of the degree of approximation ε and the structural parameters defining the underlying game G, including the total number of strategies and the length of the better reply paths. In addition, we introduce a novel concept called the *interdependence index*, δ_G, which measures the extent to which changes in strategy by members of one population alter the payoffs to members of other populations. This index is crucial for estimating the rate at which individuals change strategies and therefore the rate at which the stochastic dynamics evolve in G^N. Using a combination of results in stochastic approximation theory, together with novel techniques for measuring transition times in population games, we show that the expected convergence time to come close to ε-equilibrium is polynomial in ε^{-1}, and exponential in the number of strategies and in δ_G^{-1}.

2. PRELIMINARIES

Let $G = (\mathcal{P}, (S^p)_{p\in\mathcal{P}}, (u^p)_{p\in\mathcal{P}})$ be a normal-form n-player game, where $\mathcal{P}$ is the finite set of players, $|\mathcal{P}| = n \geq 2$.[9] Let S^p and u^p denote the strategy set and the payoff function of player $p \in \mathcal{P}$, respectively. For every player $p \in \mathcal{P}$, let $m^p = |S^p|$ and $M = \sum_{p\in\mathcal{P}} m^p$. Let $S = \prod_{p\in\mathcal{P}} S^p$ be the set of all pure strategy

[9]The case where there is a single population and G is symmetric can be analyzed using similar methods, but requires different notation. For expositional clarity, we shall restrict ourselves to the case $n \geq 2$.

profiles. We let $X^p = \Delta(S^p)$ denote the set of mixed strategies of player p. Let $\chi = \prod_{p\in\mathcal{P}} X^p$ be the Cartesian product of mixed strategies.

A sequence of pure strategy profiles $(s_1, \dots, s_k) \in \underbrace{S \times \cdots \times S}_{k \text{ times}}$ is called a *strict better reply path* if each successive pair (s_j, s_{j+1}) involves a unilateral change of strategies by exactly one player, and the change in strategy strictly increases that player's payoff.

DEFINITION 1: A game G is *weakly acyclic* if, for every strategy profile $s \in S$, there exists a strict better reply path to a Nash equilibrium. G is *strictly weakly acyclic* if, for every $s \in S$, there exists a strict better reply path to a strict Nash equilibrium. If G is weakly acyclic and has no payoff ties, then clearly G is strictly weakly acyclic.

Examples of weakly acyclic games include potential games, coordination games, games with strategic complementaries, dominance-solvable games, and many others (Fabrikant, Jaggard, and Schapira (2013)).

2.1. *Nash Population Games*

For every natural number N, the n-person game G gives rise to a population game G^N in the spirit of Nash as follows. Every "player" p represents a finite *population* of size N. The strategy set available to every member of population p is S^p. A *population state* is a vector $x \in \chi$, where, for each $p \in \mathcal{P}$ and each $i \in S^p$, the fraction x_i^p is the proportion of population p that chooses strategy i. Let χ^N denote the subset of states that results when each population has N members, that is,

$$\chi^N = \{x \in \chi : Nx \in \mathbb{N}^M\}.$$

Let $u_i^p(x)$ be the *payoff* to a member of population p who is playing strategy i in population state x, that is,

$$u_i^p(x) = u^p(e_i^p, x^{-p}).$$

Note that the payoff of any individual depends only on his own strategy and on the distribution of strategy choices in the other populations; it does not depend on the distribution in his *own* population. Let $U^p(x)$ be the vector of payoffs $(u_i^p(x))_{i\in S^p}$. Every pure Nash equilibrium of the game G^N can be interpreted as a mixed Nash equilibrium of the game G.

A population state x is a *population ε-equilibrium* (ε-equilibrium for short) if x is an ε-equilibrium of the original game G, that is, for every population p

$$d^p(x) = \max_{y^p \in X^p} u^p(y^p, x^{-p}) - u^p(x) \leq \varepsilon.$$

Intuitively, x is an ε-equilibrium if the proportion of any population that can significantly improve its payoff by changing strategies is small.

DEFINITION 2: For every $x \in \chi$, let $d(x)$ denote the *minimal* ε for which x is an ε-equilibrium. We shall say that $d(x)$ is the *deviation* of x from equilibrium, that is,

$$d(x) = \max_{p \in \mathcal{P}} d^p(x).$$

3. THE ADAPTIVE DYNAMIC

In this section, we introduce a natural class of updating procedures that define a stochastic dynamical system on the space χ^N. Suppose that every individual receives updating opportunities according to a Poisson arrival process with rate one per time period, and suppose that these processes are independent among the individuals. Thus, in expectation, there are N updates per period in each population. The speed of convergence of the process is measured relative to the underlying rate at which individuals update.[10]

Let $x \in \chi^N$ be the current state and $U^p(x)$ the vector of payoffs in that state. If a random member of a population p updates, the probability is x_i^p that he is currently playing strategy i. Assume that he switches to strategy j with probability ρ_{ij}^p, where $\sum_{j \neq i} \rho_{ij}^p \leq 1$. We shall assume that

(i) ρ_{ij}^p is Lipschitz continuous and depends only on $u_i^p(x)$ and $u_j^p(x)$;

(ii) $\rho_{ij}^p(u_i^p(x), u_j^p(x)) > 0 \Leftrightarrow u_j^p(x) > u_i^p(x)$.

We shall denote this stochastic process by $X^N(\cdot)$ and refer to it as a *stochastic pairwise comparison dynamic*; the matrix of transition functions $\rho = [\rho_{ij}^p]$ constitutes a "revision protocol" (Sandholm (2010b)).[11] It will also be convenient to write $\rho_{ij}^p(U^p(x))$ as a function of the entire vector $U^p(x)$ even though in fact it depends only on the two components $u_i^p(x)$ and $u_j^p(x)$. As noted by Sandholm (2010b), this class of dynamics has a number of desirable properties. In particular, the informational demands are very low: each individual need only compare his payoff with the potential payoff from an alternative strategy; he does not need to know the distribution of payoffs, or even the average payoff to members of his population, which would pose a heavier informational burden. It is assumed, however, that everyone knows the potential payoffs from alternative strategies.[12]

[10] This is the standard approach in the literature; see, among others, Hart and Mansour (2010), Shah and Shin (2010), Kreindler and Young (2013), Babichenko (2014), Marden and Shamma (2014).

[11] A more general definition of revision protocols was given by Björnerstedt and Weibull (1996), here we shall only consider the pairwise comparison format.

[12] One could assume instead that payoffs are observable with some error, or that individuals try out alternative strategies in order to estimate their payoffs. These and other variations can be analyzed using similar methods, but we shall not pursue them here.

One way of motivating these dynamics is in terms of switching costs. Consider a member of population p who is currently playing strategy i in state $x \in \chi$. He receives updating opportunities according to a Poisson arrival process with unit expectation. Given an updating opportunity, he draws an alternative strategy $j \neq i$ uniformly at random and compares its payoff $u_j^p(x)$ with his current payoff $u_i^p(x)$. Let c be the realization of an idiosyncratic switching cost distributed on an interval $[0, b^p]$ with c.d.f. $F^p(c)$. He then switches if and only if $c \leq u_j^p(x) - u_i^p(x)$. Thus, in each unit time interval, an updating individual in population p who is currently using strategy i switches to $j \neq i$ with probability

$$(1) \qquad \rho_{ij}^p\big(U^p(x)\big) = \frac{1}{\big(m^p - 1\big)} F^p\big(u_j^p(x) - u_i^p(x)\big).$$

Suppose that $F^p(c)$ has a density $f^p(c)$ that is bounded above and also bounded away from zero on $[0, b^p]$. Then $\rho_{ij}^p(U^p(x))$ is Lipschitz continuous, and it satisfies conditions (i) and (ii).

From Lemma 1 in Benaïm and Weibull (2003), we know that, as the size of the population grows, the behavior of the process $X^N(\cdot)$ can be approximated by the following *mean-field* differential equation on the space χ of population proportions:

$$(2) \qquad \forall p, \forall i, j \in S^p, \quad \dot{z}_i^p = \sum_{j \in S^p} z_j^p \rho_{ji}^p\big(U^p(z)\big) - z_i^p \rho_{ij}^p\big(U^p(z)\big).$$

A particularly simple example arises when $F^p(c)$ is the uniform distribution, that is, $F^p(c) = c/b^p$ for all $c \in [0, b^p]$. In this case,

$$(3) \qquad \rho_{ij}^p\big(U^p(x)\big) = \frac{\big[u_j^p(x) - u_i^p(x)\big]_+ \wedge b^p}{\big(m^p - 1\big)b^p}.$$

(In general, $a \wedge b$ denotes the minimum of a and b.) In other words, the rate of change between any two strategies is proportional to the payoff difference between them (subject to an upper bound). To avoid notational clutter, we shall consider the case where $(m^p - 1)b^p = 1$, and all payoffs $u_i^p(x)$ lie in the interval $[0, b^p]$. In this case, we can write

$$(4) \qquad \rho_{ij}^p\big(U^p(x)\big) = \big[u_j^p(x) - u_i^p(x)\big]_+.$$

This yields the following mean-field differential equation on the state space χ:

$$(5) \qquad \forall p, \forall i, j \in S^p, \quad \dot{z}_i^p = \sum_{j \in S^p} z_j^p \big[u_i^p(z) - u_j^p(z)\big]_+ - z_i^p \big[u_j^p(z) - u_i^p(z)\big]_+.$$

This is known as the *Smith dynamic* (Smith (1984)) and was originally proposed as a model of traffic flow. We shall take this as our benchmark example in what follows, but our results hold for every responsive revision protocol.

4. EQUILIBRIUM CONVERGENCE

Let G be a finite normal-form game and let G^N be the population game induced by G. Given a revision protocol ρ and a starting point in χ^N, recall that $X^N(\cdot)$ denotes the stochastic process defined by ρ. Recall that for every point $x \in \chi$, the *deviation* $d(x)$ is the minimal ε such that x constitutes a mixed ε-equilibrium of the game G. Thus $d(X^N(t))$ is a random variable that represents the deviation of the population from equilibrium at time t.

DEFINITION 3: *Equilibrium convergence* holds for G if, for every N, every revision protocol ρ, and every initial state $x \in \chi^N$,

$$\mathbb{P}\big(\exists t, d\big(X^N(t)\big) = 0\big) = 1.$$

Once the process reaches a Nash equilibrium, it is absorbed. Hence equilibrium convergence holds if and only if there exists a random time such that, from that time on, the process is at an equilibrium.

PROPOSITION 1: *Equilibrium convergence holds for a generic subset of weakly acyclic population games G.*

The proof of this proposition is given in Appendix C. We remark that the usual definition of genericity (no payoff ties) is not sufficient; a more delicate condition is needed for equilibrium convergence. We introduce this condition in Section 6.1, where we show that it also plays a key role in determining the speed of convergence.

4.1. *Convergence Time in Large Populations*

Our goal in this section is to study the convergence time as a function of the population size N. One might suppose that convergence occurs quite rapidly, as in potential games. It turns out, however, that in some weakly acyclic games the convergence time can be extremely slow. Consider the following example. For every $\delta, \gamma > 0$, let $\Gamma_{\gamma,\delta}$ be the following three-player game:

$\mathcal{L}$

	L	M	R
T	$-1, -1, 1$	$\gamma, 0, 1$	$0, \gamma, 1$
M	$0, \gamma, 0$	$-1, -1, 1$	$\gamma, 0, 1$
B	$\gamma, 0, 1$	$0, \gamma, 1$	$-1, -1, 1$

$\mathcal{R}$

	L	M	R
T	$0, 0, \delta$	$0, 0, \delta$	$0, 0, \delta$
M	$2, 2, \delta$	$0, 0, \delta$	$0, 0, \delta$
B	$0, 0, \delta$	$0, 0, \delta$	$0, 0, \delta$

This game is weakly acyclic: starting from any state where the third player plays $\mathcal{L}$, the first two players have a sequence of strict best replies that take them to $(M, L, \mathcal{L})$, at which point $\mathcal{R}$ is a strict best reply for the third player. Alternatively, if the initial state is not $(M, L, \mathcal{R})$ but the third player is playing $\mathcal{R}$, then it is a strict best reply for the third player to switch to $\mathcal{L}$. After that, the first two players have a sequence of strict best replies that takes the state to $(M, L, \mathcal{L})$, at which point $\mathcal{R}$ is a strict best reply for the third player, which takes the process to $(M, L, \mathcal{R})$.

Fix a revision protocol ρ. Let $X^N(\cdot)$ be the stochastic process associated with the preceding population game $\Gamma^N_{\gamma,\delta}$, where the members of each population update their strategy choices in accordance with ρ. Given $\varepsilon > 0$ and an initial state y, let $T^N(\varepsilon, y)$ be the first time t such that $d(X^N(t)) \leq \varepsilon$. Further, let

$$\bar{T}^N(\varepsilon) = \sup_{y \in \chi} E\big(T^N(\varepsilon, y)\big).$$

THEOREM 1: *Given any revision protocol* ρ, *there exist values* $\gamma, \delta, \varepsilon > 0$ *such that* $\bar{T}^N(\varepsilon)$ *grows exponentially with* N.

Before giving the detailed proof, we shall outline the overall argument. If populations 1 and 2 are playing the strategy combination (M, L), then population 3 would prefer $\mathcal{R}$ over $\mathcal{L}$ because the payoff gain is δ. However, if a proportion of at least δ of populations 1 and 2 are not playing (M, L), then population 3 prefers $\mathcal{L}$ to $\mathcal{R}$. The idea is to begin the process in a state such that: (i) population 3 is playing $\mathcal{L}$, and (ii) populations 1 and 2 are distributed among the cells of the cycle

$$\begin{aligned}(M, L) &\to (B, L) \to (B, M) \to (T, M) \to (T, R)\\ &\to (M, R) \to (M, L).\end{aligned}$$

The expected (deterministic) motion leads to sluggish movement around the cycle with a very low proportion of populations 1 and 2 in the diagonal cells of the left matrix, while population 3 continues to play $\mathcal{L}$. The stochastic process also follows this pattern with high probability for a long time, although eventually enough mass accumulates in the cell (M, L) of the left matrix to cause players to switch to strategies in the right matrix. Using a result in stochastic approximation theory due to Benaïm and Weibull (2003), we show that the expected waiting time until this happens is exponential in N.

Although we prove this result formally for pairwise comparison dynamics, a similar argument holds for a wide variety of better reply dynamics including the replicator dynamic. Indeed, consider any continuous better reply dynamic such that the expected rate of flow from a lower to a higher payoff strategy is strictly increasing in the payoff difference. Then the expected flow out of each cell in the above cycle is bounded away from zero. When the population size

N is large, it becomes extremely improbable that a large enough proportion of the population will accumulate in the particular cell (M, L), which is needed to trigger a shift from the left to the right matrix (and thus escape from the cycle).

PROOF OF THEOREM 1: Even though the game is not generic, it can be verified that the proof of the theorem remains valid under a slight perturbation of the payoff functions. For expositional clarity, we shall work with the nongeneric version.

Let ρ be the given revision protocol. Consider the subgame for players 1 and 2 when player 3 is held fixed at $\mathcal{L}$:

(6)

	L	M	R
T	$-1, -1$	$\gamma, 0$	$0, \gamma$
M	$0, \gamma$	$-1, -1$	$\gamma, 0$
B	$\gamma, 0$	$0, \gamma$	$-1, -1$

Let $Y = X^1 \times X^2$ be the product space of the mixed strategies of players 1 and 2. Consider the following differential equation with initial condition $z(0) = y \in Y$:

$$\text{for } p = 1, 2 \text{ and } i \in S^p, \quad \dot{z}_i^p = \sum_{j \in S^p} z_j^p \rho_{ji}^p\big(U^p(z)\big) - z_i^p \rho_{ij}^p\big(U^p(z)\big). \tag{7}$$

Let $\Phi^\gamma : \mathbb{R}_+ \times Y \to Y$ be the *semi-flow* of the differential equation (7) that corresponds to the game with parameter γ, that is, for every $t \geq 0$ and $y \in Y$:

$$\Phi^\gamma(t, y) = z(t)$$
$$\text{where } z(\cdot) \text{ is the solution of (7) with initial condition } z(0) = y.$$

Let $\mathcal{A} \subset Y$ be the set of states such that the diagonal strategy combinations have mass zero:

$$\mathcal{A} = \big\{y \in Y : y_T^1 y_L^2 + y_M^1 y_M^2 + y_B^1 y_R^2 = 0\big\}.$$

Consider the case $\gamma = 0$. We claim that $\mathcal{A}$ is an *attractor* of the semi-flow Φ^0; that is, $\mathcal{A}$ is a minimal set with the following properties:

1. $\mathcal{A}$ is *invariant*: for all $t \geq 0$, $\Phi^0(t, \mathcal{A}) = \mathcal{A}$.
2. There exists a neighborhood $\mathcal{U}$ of $\mathcal{A}$ such that

$$\lim_{t \to \infty} \sup_{y \in \mathcal{U}} \operatorname{dist}\big(\Phi^0(t, y), \mathcal{A}\big) = 0. \tag{8}$$

The first property follows at once from (7) and the fact that $\mathcal{A}$ is a subset of Nash equilibria when $\gamma = 0$. To establish the second, note that, when $\gamma = 0$, the corresponding game is a potential game with potential function

$$P(y^1, y^2) = -y^1 \cdot y^2 = -(y^1_T y^2_L + y^1_M y^2_M + y^1_B y^2_R).$$

The potential is weakly increasing along any solution of the dynamical system (7), and is strictly increasing if the starting point is not a Nash equilibrium. In addition, by Theorem 7.1.2 in Sandholm (2010b), any solution of (7) converges to a Nash equilibrium. The unique fully mixed Nash equilibrium $e = ((\frac{1}{3}, \frac{1}{3}, \frac{1}{3}), (\frac{1}{3}, \frac{1}{3}, \frac{1}{3}))$ has potential $-\frac{1}{3}$. All other Nash equilibria are partially mixed and have potential zero. Thus, whenever the starting point has potential greater than $-\frac{1}{3}$, the potential must increase to zero, and the limit state must be a Nash equilibrium. For every value $a \in \mathbb{R}$, let $\mathcal{U}_a$ be the set of states with potential strictly greater than a. The fact that $\mathcal{A}$ satisfies the second property in (8) follows by letting the open set $\mathcal{U} = \mathcal{U}_{-1/4}$.

By Theorem 9.B.5 in Sandholm (2010b), for all sufficiently small $\gamma \geq 0$, there exists an attractor $\mathcal{A}^\gamma$ of Φ^γ such that $\mathcal{A}^0 = \mathcal{A}$ and the map $\gamma \to \mathcal{A}^\gamma$ is upper-hemicontinuous. It follows that for all sufficiently small $\gamma > 0$, all elements $y \in \mathcal{A}^\gamma$ satisfy

$$y^1 \cdot y^2 \leq 1/10 \tag{9}$$

and

$$\lim_{t\to\infty} \sup_{w \in \mathcal{U}_{-1/4}} \operatorname{dist}(\Phi^\gamma(t, w), \mathcal{A}^\gamma) = 0. \tag{10}$$

Fix any $\gamma > 0$ such that (9) and (10) hold, and let $\mathcal{C} = \mathcal{A}^\gamma$. For every positive constant $r > 0$, let $\mathcal{C}^r$ be the set of points that lie within a distance r of $\mathcal{C}$. The proof of the theorem is based on two lemmas. The first lemma asserts that, among all states in $\mathcal{C}$, the proportion of the population playing the strategy pair (M, L) is bounded away from 1.

LEMMA 1: *There exists $\theta^* > 0$ such that, for every $y \in \mathcal{C}$,*

$$y^1_M y^2_L \leq 1 - \theta^*.$$

PROOF: We note first that the switching probability of population 2 between pairs of strategies is bounded above by some positive number τ. Hence the flow into strategy L (and a fortiori into (M, L)) is at most $\tau(1 - y^2_L)$ for every state y. Let y be such that y^2_L is larger than $\frac{1+(3/2)\gamma}{1+2\gamma}$. It can be checked that in any such state, strategy B yields $\frac{\gamma}{2}$ more than strategy M to population 1. Hence there is a positive number $\tau' > 0$ such that the outflow from M to B is larger than $\tau' y^1_M$. This holds in particular for every y that is sufficiently concentrated on (M, L). Hence there exists a number θ' such that, whenever $y^1_M y^2_L \geq 1 - \theta'$, the outflow to strategy B in population 1 is greater than the inflow to strategy L in population 2. Let $\theta^* = \frac{\theta'}{2}$. It therefore must hold that $y^1_M y^2_L \leq 1 - \theta^*$ for every $y \in \mathcal{C}$. This establishes our first claim. *Q.E.D.*

Using Lemma 1 and the properties of the attractor $\mathcal{C}$, we shall next show the following.

LEMMA 2: *There exist constants* $r, \delta > 0$, $T > 0$, *and* $\varepsilon > 0$ *with the following properties*:
1. $\Phi^\gamma(T, \mathcal{C}^{2r}) \subset \mathcal{C}^r$.
For every point $w \in \mathcal{C}^r$, *every time* $t \geq 0$, *and every point* y *such that* $\|y - \Phi^\gamma(t, w)\| \leq r$, *the following two conditions hold*:
2. y *assigns a proportion smaller than* $1 - 2\delta$ *to the profile* (M, L), *that is*, $y^1_M y^2_L < 1 - 2\delta$;
3. $d(y) > \varepsilon$.

We use these key properties to establish that the waiting time to reach an ε-equilibrium grows exponentially with the population size N.

PROOF OF LEMMA 2: By (10), for every small enough $r > 0$, there exists a large enough $T > 0$ such that

$$\forall w \in \mathcal{C}^{2r}, \quad \Phi^\gamma(T, w) \in \mathcal{C}^r, \tag{11}$$

which establishes claim 1. Let θ^* be the constant guaranteed by Lemma 1 and let $\delta = \frac{\theta^*}{4}$. It follows from equation (9) and the properties of the attractor $\mathcal{C}$ that there exists $r \in (0, \delta)$ such that:

$$\begin{aligned} &\text{If} \quad \|y - \Phi^\gamma(t, w)\| \leq r \quad \text{for some } w \in \mathcal{C}^{2r}, \\ &\quad \text{then} \quad y^1_M y^2_L \leq 1 - 2\delta \quad \text{and} \quad y^1 \cdot y^2 \leq 1/5. \end{aligned} \tag{12}$$

This establishes claim 2 in Lemma 2. To establish claim 3, it suffices to show that there exists $\varepsilon > 0$ such that any point y with inner product $y^1 \cdot y^2 \leq 1/5$ has a deviation greater than ε. To see this, note that the unique Nash equilibrium of the game in (6) is $e = ((1/3, 1/3, 1/3), (1/3, 1/3, 1/3))$ and $e^1 \cdot e^2 = \frac{1}{3}$. Therefore, any such y must be bounded away from e, hence the deviation of y must be bounded away from 0. This completes the proof of Lemma 2. *Q.E.D.*

We shall now prove Theorem 1. Let $\Gamma_{\gamma,\delta}$ be the game that corresponds to the constants γ and δ. For every $x \in \chi$, let x' be the projection of x onto the set Y. Let Ψ^γ be the semi-flow of the dynamic over the space χ. By Lemma 1 in Benaïm and Weibull (2003), there exists a constant $c > 0$ such that, for every $x \in \chi$ and all sufficiently large N,

$$\mathbb{P}\Big(\sup_{0 \leq t \leq T} \|\Psi^\gamma(t, x) - X^N(t)\| > r\Big) \leq \exp(-cN). \tag{13}$$

Note that, by definition of the game $\Gamma_{\gamma,\delta}$, if $x \in \chi$ assigns a proportion that is smaller than $1 - \delta$ to the profile (M, L), then $\mathcal{L}$ is the unique best reply for player 3. Hence as long as the population state x satisfies $x^1_M x^2_L < 1 - \delta$, no member of population 3 switches to strategy $\mathcal{R}$. Hence if $x = (x', (1, 0))$ for some $x' \in \mathcal{C}^{2r}$, then by claim 2 in Lemma 2 and the definition of Φ^γ, we have

$$\forall t \geq 0, \quad \Psi^\gamma(t,x) = \big(\Phi^\gamma(t,x),(1,0)\big).$$

Let $x = (x',(1,0)) \in \chi^N$ be a starting point of $X^N(\cdot)$ such that $x' \in C^{2r}$. We claim that if $\sup_{0\leq t\leq T} \|\Psi^\gamma(t,x) - X^N(t)\| \leq r$, then the following conditions hold with certainty:

$$\big(X^N(T)\big)' \in \mathcal{C}^{2r}, \tag{14}$$

$$\forall t \in [0,T], \quad y = \big(X^N(t)\big)' \quad \text{satisfies} \quad y_M^1 y_L^2 \leq 1 - 2\delta, \tag{15}$$

$$d\big(X^N(t)\big) > \varepsilon. \tag{16}$$

To verify (14), note that by claim 1 of Lemma 2,

$$\Psi^\gamma(T,x) = \big(\Phi^\gamma\big(t,x'\big),(1,0)\big) \in C^r \times \big\{(1,0)\big\}.$$

Since $\|X^N(T) - \Psi^\gamma(T,x)\| \leq r$, it follows that $(X^N(T))' \in \mathcal{C}^{2r}$. Condition (15) follows from claim 2 of Lemma 2. Condition (16) follows at once from claim 3 of Lemma 2. Hence by equation (13), the above three conditions hold with probability at least $1 - \exp(-cN)$.

Now divide time into blocks of size T. If there exists a time $t \geq 0$ such that the deviation of $X^N(t)$ is smaller than ε, then by the above three conditions, there must exist a $k \geq 0$ such that, for some $kT \leq t' \leq (k+1)T$,

$$\big\|X^N\big(t'\big) - \Psi^\gamma\big(t' - kT, X^N(kT)\big)\big\| > r.$$

Therefore, the expectation of the first time t_0 such that $d(X^N(t_0)) < \varepsilon$ is greater than $\exp(cN)T$. *Q.E.D.*

5. CONVERGENCE TIME UNDER AGGREGATE SHOCKS

5.1. *Aggregate Shocks*

The stochastic dynamic treated in the preceding section can be thought of as a better reply process with idiosyncratic shocks: whenever an individual revises, he chooses a new strategy with a *probability* that depends on its payoff gain relative to his current strategy. As we have seen, it can take exponentially long in the population size for average behavior to come close to Nash equilibrium even in very simple weakly acyclic games. In this section, we show that the convergence time can be greatly accelerated if, in addition to idiosyncratic shocks, there are aggregate shocks that affect all members of certain subpopulations at the same time. Such shocks can arise from interference to communications that temporarily prevent some groups of individuals from learning about the payoffs available from alternative strategies. Or they could arise from common payoff shocks that affect all members of a given subgroup at the same time. Suppose, for example, that x_i^p represents the proportion of population p

that is currently using a given product i with a payoff u_i^p. For each $j \neq i$, let c_{ij}^p be the realization of a stochastic switching cost that affects all i-users who are contemplating a switch to j. If $c_{ij}^p > u_j^p - u_i^p$, then the cost is prohibitive and no one wants to switch, whereas if $c_{ij}^p = 0$, they switch at the same rate as before.

In what follows, we shall make the simplifying assumption that these aggregate shocks are binary and i.i.d.[13] Specifically, we shall assume that for each population $p \in \mathcal{P}$, every pair of distinct strategies $i, j \in S^p$, there is a binary random variable $\alpha_{ij}^p(\cdot)$ that changes according to a Poisson process with unit arrival rate. At every arrival time t, $\alpha_{ij}^p(t)$ takes the value 0 or 1 with equal probability. Let $A^p(\cdot) = [\alpha_{ij}^p(\cdot)]_{i,j \in S^p}$ and let $\vec{A}(\cdot) = (A^1(\cdot), \dots, A^n(\cdot))$, where the variables α_{ij}^p are independent. When $\alpha_{ij}^p(t) = 1$, the switch rate from i to j is as before, and when $\alpha_{ij}^p(t) = 0$, it equals zero. Thus the random variables α_{ij}^p retard the switch rates in expectation. They represent aggregate or common shocks in the sense that $\alpha_{ij}^p(t)$ affects *all* of the individuals in the subpopulation p who are currently playing strategy i at time t.

We assume that individual changes of strategy are governed by i.i.d. Poisson arrival processes that are independent of the process $\vec{A}$. At every arrival time t, a member of some population is "activated." The probability that the activated individual is in population p and is currently playing the particular strategy i equals $\frac{x_i^p}{n}$. The switch rate of the activated individuals depends on the current shock realization $\vec{A}(t) = (\alpha_{ij}^p(t))_{i,j \in S^p}$. In particular, an individual switches from strategy i to strategy j with conditional probability $\alpha_{ij}^p(t)\rho_{ij}^p(U^p(x))$. We shall denote this process by $(\vec{A}(\cdot), X^N(\cdot))$, and sometimes write simply $X^N(\cdot)$ when the associated process $\vec{A}$ is understood.

Theorem 1 demonstrates that, for all sufficiently small $\varepsilon > 0$ there exists a starting point such that the expected first time that the process is in a state with deviation at most ε increases exponentially with the population size. We now show that when the process is subjected to aggregate shocks, due, for example, to interruptions in communication, convergence time can be greatly accelerated. In fact, we shall show that convergence is more rapid for an even more demanding notion of convergence time than the one used in Theorem 1.[14]

Given a population size N, a length of time L, and an initial state $x \in \chi^N$, let

$$D^{N,L}(x) = \frac{1}{L} \int_0^L d\big(X^N(s)\big)\, ds.$$

$D^{N,L}(x)$ is the *time average deviation* from equilibrium over a window of length L starting from state x.

[13]The analysis can be extended to many other distributions, including those that have both a common and an idiosyncratic component, but this would substantially complicate the notation without yielding major new insights.

[14]A fortiori, Theorem 1 continues to hold for this more stringent notion of convergence time.

DEFINITION 4: Given a game G, a revision protocol ρ, and an $\varepsilon > 0$, the associated sequence of processes $(\vec{A}, X^N)_{N\in\mathbb{N}}$ exhibits *fast convergence with precision* ε if there exists a number L and a positive integer N_ε such that

$$\forall L' \geq L, \forall N \geq N_\varepsilon, \forall x \in \chi^N, \quad E\big[D^{N,L'}(x)\big] \leq \varepsilon. \tag{17}$$

The *convergence time with precision* ε, L_ε, is the infimum over all such numbers L. The sequence $(\vec{A}(\cdot), X^N(\cdot))_{N\in\mathbb{N}}$ exhibits *fast convergence* if (17) holds for every $\varepsilon > 0$.

Another notion of fast convergence is the expected first passage time to reach the basin of attraction of some Nash equilibrium. This concept is more demanding, and in our setting it may become arbitrarily large as N grows. Under our definition, it suffices that behavior comes close to Nash equilibrium behavior over long periods of time; we do not insist that the process remains close to a given equilibrium forever.

THEOREM 2: *There exists a generic set $\mathcal{G}$ of weakly acyclic population games such that, for every revision protocol and every game $G \in \mathcal{G}$, the sequence of processes $(\vec{A}(\cdot), X^N(\cdot))_{N\in\mathbb{N}}$ exhibits fast convergence.*

5.2. *Proof Outline of Theorem 2*

Here we shall sketch the general gist of the argument; the proof is given in the next section. Given a pure strict Nash equilibrium y, define its ε-*basin* to be a neighborhood of y such that the deviation of every state is smaller than ε and the dynamic cannot exit from this neighborhood.[15]

Given $\varepsilon > 0$, let us say that a population state is "good" if its deviation from equilibrium is smaller than ε. It is "very good" if it lies in the ε-basin of some pure Nash equilibrium. Otherwise, the state is "bad." The first step in the proof is to show that, starting from any *bad* state x, there exists a continuous better reply path to a very good state. Surprisingly, this property is *not* guaranteed by weak acyclicity of the underlying normal-form game. However, it does hold for almost all weakly acyclic games, that is, for a set of payoffs having full Lebesgue measure. (The proof of this fact is quite delicate, and does not follow merely if there are no payoff ties; the details are given in Appendix C.)

The second step is to show that, with positive probability, there is a sequence of shocks such that the expected motion of the process is very close to the path defined in the first step. The third step is to show that, for every initial bad state x, there is a time T_x such that, whenever the process $X^N(\cdot)$ starts close enough to x and N is large enough, there is a positive probability that, by time $T_x + t$, the process will be in a very good state. Using compactness arguments,

[15]The existence of such a neighborhood is demonstrated in Section 5.3.

one can show that there is a population size N_ε and a time T_ε such that the preceding statement holds *uniformly* for all bad states x whenever $N \geq N_\varepsilon$. (The existence of a uniform time is crucial to the result, and relies heavily on the assumption of aggregate shocks.) Once the process is in a very good state, it is impossible to leave it. From these statements it follows that, after some time $T'_\varepsilon > T_\varepsilon$, the process is *not* in a bad state with high probability. Therefore, at T'_ε and all subsequent times, the process is in a good or very good state with high probability, hence its expected deviation is small.[16] From this, it follows that there is a bounded time L_ε such that the expected deviation is at most ε over any window of length at least L_ε. Moreover, this statement holds uniformly for all $N \geq N_\varepsilon$.

5.3. *Proof of Theorem 2*

Before commencing the proof, we shall need several auxiliary results. Let $v = \sum_{p\in\mathcal{P}} \binom{|S_p|}{2}$ and let $\beta : [0, T] \to \{0, 1\}^v$ be a piecewise constant function that results from a finite series of *shocks* $(\vec{A}(t))_{0\leq t\leq T}$ on the interval $[0, T]$. (All other realizations have total probability 0.) We shall call $(\beta(t))_{0\leq t\leq T}$ a *shock realization on* $[0, T]$. Given any $x \in \chi$, let $z : [0, T] \to \chi$ be the solution of the following differential equation:

$$(18)\qquad \forall p, \forall i, j \in S^p, \quad \dot{z}_i^p = \sum_{j\in S^p} \bigl[z_j^p \rho_{ji}^p\bigl(U^p(z)\bigr)\beta_{ji}(t) - z_i^p \rho_{ij}^p\bigl(U^p(z)\bigr)\beta_{ij}(t)\bigr],$$

$$z(0) = x.$$

Such a solution $z(\cdot)$ is called a *continuous better reply path*.

Fix a time $T > 0$. For any two shock realizations $\beta : [0, T] \to \{0, 1\}^v$ and $\gamma : [0, T] \to \{0, 1\}^v$, define the *distance* between β and γ on $[0, T]$ as follows:

$$d_T(\beta, \gamma) \equiv \mu\bigl(\{0 \leq t \leq T : \beta(t) \neq \gamma(t)\}\bigr),$$

where μ is a Lebesgue measure.

The following lemma provides an approximation of the distance between two continuous better reply paths as a function of the initial conditions and the distance between the corresponding two shock realizations $\beta : [0, T] \to \{0, 1\}^v$, and $\gamma : [0, T] \to \{0, 1\}^v$.

LEMMA 3: *Let $\beta : [0, T] \to \{0, 1\}^v$ and $\gamma : [0, T] \to \{0, 1\}^v$ be two shock realizations, and let $z(\cdot)$ and $y(\cdot)$ be the two continuous better reply paths that correspond to $\beta(\cdot)$ and $\gamma(\cdot)$, respectively, with initial states x_0 and y_0. There exists*

[16]Note, however, that one cannot conclude that the process reaches a *very* good state in bounded time for all $N \geq N_\varepsilon$. The difficulty is that, as N becomes large, the state space becomes larger and the process can get stuck for long stretches of time in states that are extremely close to equilibrium without being absorbed into the equilibrium.

a constant ν *such that, for all* $r, \eta > 0$,

$$d_T(\beta, \gamma) < r \text{ and } \|z_0 - y_0\| < \eta \Rightarrow \sup_{t\in[0,T]} \|z(t) - y(t)\| < (\eta + \nu r)\exp(\nu T).$$

The result shows that when the shock realizations are very close and the initial states are very close, the resulting dynamical paths are also very close over the finite interval $[0, T]$. The proof is a standard application of Grönwall's inequality (see Theorem 1.1 in Hartman (2002)) and is therefore omitted.

Next we shall define a stochastic process $z(\cdot) = (z(t))_{t\geq 0}$ on the state space χ that approximates the behavior of the process $(\vec{A}(\cdot), X^N(\cdot)) = ((\vec{A}(t), X^N(t)))_{t\geq 0}$ over any finite interval $[0, T]$ for all sufficiently large N. Given a shock realization $\alpha(\cdot) = (\alpha(t))_{0\leq t\leq T}$ on $[0, T]$, let $z(\cdot)$ obey the following differential equation on $[0, T]$:

$$\forall p, \forall i, j \in S^p, \quad \dot{z}_i^p = \sum_{j\in S^p} \bigl[z_j^p \rho_{ji}^p\bigl(U^p(z)\bigr)\alpha_{ji}^p - z_i^p \rho_{ij}^p\bigl(U^p(z)\bigr)\alpha_{ij}^p\bigr]. \tag{19}$$

The following is a variant of Lemma 1 in Benaïm and Weibull (2003) (the proof is given in Appendix A).

LEMMA 4: *For every* $\varepsilon > 0$ *and* $T > 0$, *and every solution* $z(\cdot)$ *of* (19), *there exists* $N_{T,\varepsilon}$ *such that*

$$\forall N \geq N_{T,\varepsilon}, \quad \mathbb{P}\Bigl(\sup_{t\in[0,T]} \|X^N(t) - z(t)\| > \varepsilon\Bigr) < \varepsilon. \tag{20}$$

Let $s = (i^p)_{p\in\mathcal{P}}$ be a strict Nash equilibrium. Given any $\varepsilon > 0$ there exists a number $\phi < 1$ such that, for every x satisfying $x_{i^p}^p > \phi$ for all $p \in \mathcal{P}$, the following two conditions hold: (i) i^p is the unique best response in state x by all members of every population $p \in \mathcal{P}$; (ii) $d(x) < \varepsilon$. Let $\phi_\varepsilon(s)$ be the infimum of all such ϕ.

DEFINITION 5: The *ε-basin* of a strict Nash equilibrium $s = (i^p)_{p\in\mathcal{P}}$ is the open set $\mathcal{B}_\varepsilon(s)$ of all $x \in \chi$ such that $x_{i^p}^p > \phi_\varepsilon(s)$. Let $\mathcal{B}_\varepsilon$ be the union of all such sets $\mathcal{B}_\varepsilon(s)$.

We claim that once the process $X^N(\cdot)$ enters $\mathcal{B}_\varepsilon$, it stays there forever. Suppose that $X^N(t) = x \in \mathcal{B}_\varepsilon(s)$, where $s = (i^p)_{p\in\mathcal{P}}$ is a strict Nash equilibrium. Thus, i^p is the unique best reply by all members of p, so $x_{i^p}^p$ cannot decrease under *any* better reply dynamic. Therefore, for every $t' \geq t$, $X^N(t') = y$ implies $y_{i^p}^p > \phi_\varepsilon(s)$. Hence $X^N(t') \in \mathcal{B}_\varepsilon(s)$ for all $t' \geq t$.

LEMMA 5: *There exists a generic subset of weakly acyclic games* $\mathcal{G}$ *such that, for every* $\varepsilon > 0$ *and for every state* $y \in \chi$ *that is not a Nash equilibrium*, *there exists* *a time* T_y *and a shock realization* $\beta : [0, T_y] \to \{0, 1\}^v$ *such that the solution of* (18) *with initial condition* $z(0) = y$ *enters* $\mathcal{B}_\varepsilon$ *by time* T_y.

With these lemmas in hand, we can now proceed to the proof of Theorem 2.

PROOF OF THEOREM 2: Let $\mathcal{G}$ be the generic subset of weakly acyclic games guaranteed by Lemma 5. Choose $G \in \mathcal{G}$. Given $\varepsilon > 0$, let $E^{\varepsilon/2}$ be the set of states in χ with deviation strictly smaller than $\frac{\varepsilon}{2}$. By construction, $\mathcal{B}_{\varepsilon/2} \subset E^{\varepsilon/2}$ and genericity ensures that $\mathcal{B}_{\varepsilon/2} \neq \emptyset$. Let $(E^{\varepsilon/2})^c$ denote the set of *bad* states. Given a bad state y, Lemma 5 implies that there is a time T_y and a shock realization $\beta_y : [0, T_y] \to \{0, 1\}^v$ such that, starting from y, the path $z_y(t)$ defined by (18) enters $\mathcal{B}_{\varepsilon/2}$ by time T_y. By Lemma 3, there exists $\theta_y > 0$ such that, whenever $\|y - y'\| < \theta_y$ and $\beta' : [0, T_y] \to \{0, 1\}^v$ is a shock realization such that $d_{T_y}(\beta, \beta') < \theta_y$, the solution to (18) with starting point y' and realization β' is *also* in the open set $\mathcal{B}_{\varepsilon/2}$ by time T_y.[17]

By Lemmas 4 and 5, there exists an open neighborhood C_y of y, a positive integer N_y, and a positive number r_y, such that, for all $N \geq N_y$,

$$\mathbb{P}\big(X^N(T_y) \in \mathcal{B}_{\varepsilon/2} : X^N(0) \in C_y\big) > r_y. \tag{21}$$

The family $\{C_y : y \in E_\varepsilon\}$ covers the set of bad states $(E^{\varepsilon/2})^c$. Since the latter is compact, there exists a finite covering

$$(E^\varepsilon)^c \subseteq \bigcup_{m=1}^{l} C_{y_m}.$$

Let

$$T_\varepsilon = \max\{T_{y_1}, \dots, T_{y_l}\}, \quad r_\varepsilon = \min\{r_{y_1}, \dots, r_{y_l}\}, \quad \text{and}$$
$$N_\varepsilon = \max\{N_{y_1}, \dots, N_{y_l}\}.$$

It follows from expression (21) that for all $N \geq N_\varepsilon$,

$$\mathbb{P}\big(\exists s \in [0, T_\varepsilon] \text{ s.t. } X_a^N(s) \in \mathcal{B}_{\varepsilon/2} : X^N(0) \in \big(E^{\varepsilon/2}\big)^c\big) \geq r_\varepsilon. \tag{22}$$

Fix $N \geq N_\varepsilon$. We have proved that whenever $X^N(t)$ is in a bad state, then by time $t + T_\varepsilon$ it has entered the absorbing set $\mathcal{B}_{\varepsilon/2}$ with probability at least r_ε.

Starting from an arbitrary state $X^N(0)$, let T_1 be the first time (if any) such that $X^N(T_1)$ is bad. Let T_2 be the first time (if any) *after* $T_1 + T_\varepsilon$ such that $X^N(T_2)$ is bad. In general, let T_{k+1} be the first time (if any) after time $T_k + T_\varepsilon$

[17] This argument holds for much more general shock distributions: it suffices that the shocks steer the process sufficiently close to the target path $z_y(t)$ with a probability that is bounded away from zero.

such that $X^N(T_{k+1})$ is bad. If the process has entered $\mathcal{B}_{\varepsilon/2}$ by time $T_k + T_\varepsilon$, then T_{k+1} will never occur. Hence, by equation (22), the probability that T_{k+1} occurs (given that T_k occurs) is *at most* $1 - r_\varepsilon$. It follows that the *expected* number of times T_k over the entire interval $[0, \infty)$ is bounded above by

$$\sum_{k=1}^{\infty} k(1-r_\varepsilon)^{k-1} = \frac{1}{r_\varepsilon^2}. \tag{23}$$

By construction, all of the bad times (if any) fall in the union of time intervals

$$S = \bigcup_{k=1}^{\infty} [T_k, T_k + T_\varepsilon]. \tag{24}$$

Fix a length of time $L > 0$. From (23) and (24), it follows that the expected proportion of bad times in the interval $[0, L]$ is bounded above by

$$\frac{T_\varepsilon}{r_\varepsilon^2 L}. \tag{25}$$

Let K be the maximal deviation among all bad states. The deviation of all other states is, by definition, at most $\frac{\varepsilon}{2}$. Now choose $L = \frac{2KT_\varepsilon}{\varepsilon r_\varepsilon^2}$. From (25), we deduce that the expected proportion of bad times on $[0, L]$ is at most $\frac{\varepsilon}{2K}$. Hence the expected deviation of the process on $[0, L]$ is at most

$$\frac{\varepsilon}{2K} \cdot K + \frac{\varepsilon}{2}\left(1 - \frac{\varepsilon}{2K}\right) < \varepsilon.$$

This also holds for all $L' \geq L$ and all $N \geq N_\varepsilon$. Hence the sequence of processes $(\vec{A}(\cdot), X^N(\cdot))_{N\in\mathbb{N}}$ exhibits fast convergence. *Q.E.D.*

6. THE SPEED OF CONVERGENCE

In this section, we shall derive an explicit bound on the speed of convergence as a function of certain structural properties of G, including the total number of strategies, the length of the better reply paths, the degree of approximation ε, and (crucially) the extent to which players can influence the payoffs of other players—a concept that we turn to next.

6.1. *The Notion of Genericity*

Fix a game structure, $(\mathcal{P}, (S^p)_{p\in\mathcal{P}})$, and let $S = \prod_{p\in\mathcal{P}} S^p$, which is assumed to be finite. A game G with this structure is determined by a vector of $n|S|$ payoffs

$$\big(u^p(s)\big)_{p\in\mathcal{P}, s\in S}.$$

A set $\mathcal{G}$ of such games is *generic* if the payoffs have full Lebesgue measure in $\mathbb{R}^{n|S|}$. Genericity is often framed in terms of "no payoff ties," but in the present situation we shall need a different (and more demanding) condition on the payoffs.

For every $x \in \chi$, two distinct populations $p, q \in \mathcal{P}$, and strategies $k \in S^q$ and $i \in S^p$, let $u_i^p(e_k^q, x^{-q})$ be the payoff to a member of population p who is playing strategy i when all members of population q are playing strategy k and all other populations are distributed in accordance with x.

Fix a population q, and let $x \in \chi$. Let $k, l \in S^q$ be two distinct strategies for population q. Let $p \neq q$ and let $i \in S^p$. Define

$$\Delta_i^{(k,l)}(x) = u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right).$$

$\Delta_i^{(k,l)}(x)$ represents the impact that members of population q have on those members of population p who are currently playing strategy i, when the former switch from l to k in state x.

Let $\mathbf{s}^p(x) \subset S^p$ be the support of x^p. Define the (k,l)-*impact* q has on p as follows:

$$\Delta_{(q\to p)}^{(k,l)}(x) = \max_{i,j\in \mathbf{s}^p(x)} \left|\Delta_i^{(k,l)}(x) - \Delta_j^{(k,l)}(x)\right|.$$

Note that when x is a pure state, $\Delta_{(q\to p)}^{(k,l)}(x) = 0$. Finally, define the *impact of population* q in state x as follows:

$$\max_{p\neq q} \min_{k,l\in S^q \ k\neq l} \Delta_{(q\to p)}^{(k,l)}(x). \tag{26}$$

To better understand the notion of impact, consider a state x where all members of every population $p \neq q$ are indifferent among all of the strategies that are used by some member of population p. Assume that a proportion of members of population q revise their strategy from l to k, and let y be the resulting state. Consider the case where the impact of population q is zero. In that case, it follows that all members of every population $p \neq q$ would still be indifferent among their strategies, because the switch has exactly the same impact on the payoffs of all strategies in S^p. On the other hand, if the impact of population q is positive, then, for some population p and some $i \neq j$, the difference in payoffs to those playing i and those playing j is positive in state y. Expression (26) provides a bound on these payoff differences.

For every state x and population p, let

$$\tilde{d}^p(x) = \max_{i\in \mathbf{s}^p(x)} u_i^p(x) - u^p(x). \tag{27}$$

Let $\tilde{d}(x) = \max_{p\in\mathcal{P}} \tilde{d}^p(x)$. Note that $\tilde{d}(x)$ measures the maximum positive gap between the payoff to some strategy that is played by a positive fraction of the population, and the average payoff to members of that population. In particular, if x is a Nash equilibrium, then $\tilde{d}(x) = d(x) = 0$. However, $\tilde{d}(x) = 0$ does *not* imply that x is a Nash equilibrium; in fact, $\tilde{d}(x) = 0$ for every pure state x, whether or not it is an equilibrium. In general, $\tilde{d}(x) \leq d(x)$ with equality whenever x is in the interior of the state space χ. We shall need the function $\tilde{d}(x)$ in order to formulate our condition of δ-genericity, which we turn to now.

DEFINITION 6: A game G is called δ*-generic* if the following two conditions hold for every population $q \in \mathcal{P}$:

1. For every two distinct pure strategy profiles, the associated payoffs differ by at least δ.
2. The impact of q in state x is at least δ whenever $\tilde{d}(x) \leq \delta$ and $|\mathbf{s}^p(x)| \geq 2$ for some $p \neq q$.

We remark that if we wanted the impact of q to be at least δ in *all* states x, then the condition would not hold generically (as may be shown by example).

DEFINITION 7: The *interdependence index* of G, δ_G, is the supremum of all $\delta \geq 0$ such that G is δ-generic.

PROPOSITION 2: *Given a game structure* $(\mathcal{P}, (S_p)_{p\in\mathcal{P}})$, *there exists a generic set of payoffs such that the associated game G has positive interdependence index* δ_G.

The proof of Proposition 2 is given in Appendix C.

6.2. *Bounding the Convergence Time*

Given a generic, weakly acyclic game G, we shall now establish a bound on the convergence time as a function of the following parameters: the precision level ε, the interdependence index δ_G, the number of players n, the total number of strategies M, and the "responsiveness" of the revision protocol ρ—a concept that we define as follows.

DEFINITION 8: A revision protocol ρ is *responsive* if there exists a positive number $\lambda > 0$ (the *response rate*) such that, for every state $x \in \chi$, population $p \in \mathcal{P}$, and every two distinct strategies $i, j \in S^p$,

$$\rho^p_{ij}\big(U^p(x)\big) \geq \lambda \cdot \big[u^p_j(x) - u^p_i(x)\big]_+. \tag{28}$$

This assumption guarantees that the switching rate between two different strategies, relative to the payoff difference between them, is bounded away from zero. The Smith protocol, for example, is responsive with $\lambda = 1$ (see expression (4)). More generally, consider any revision protocol that is generated by idiosyncratic switching costs as described in (1)–(3). If the switching cost distribution $F^p(c)$ has a density $f^p(c)$ that is bounded away from zero on its domain $[0, b^p]$, then the resulting dynamic is responsive.

To state our main result, we shall need the following notation. Let G be a weakly acyclic n-person population game where $n \geq 2$. For each pure strategy profile s, let B_s be the length of the shortest pure better reply path from s to a Nash equilibrium, and let $B = \max_s B_s$. Let M denote the total number of strategies in G. Recall that, given any small $\varepsilon > 0$, the *convergence time with precision* $\varepsilon > 0$, L_ε, is the infimum over all L such that the expected deviation of the process over the interval $[0, L']$ is at most ε for every $L' \geq L$ and for all sufficiently large N.

In what follows, it will be convenient to assume that the payoffs are normalized so that

$$\forall p, \forall i, \forall x, \quad 0 \leq u_i^p(x) \leq 1.$$

In addition, we shall assume that

$$\forall p, \forall i, \forall x, \quad \sum_{j \neq i} \left[u_j^p(x) - u_i^p(x)\right]_+ \leq 1.$$

THEOREM 3: *Let G have interdependence index $\delta > 0$, and let ρ be a responsive revision protocol with response rate $\lambda > 0$. There exists a constant K such that, for every $\varepsilon > 0$, the convergence time L_ε is at most*

$$K\left[\varepsilon^{-1} \exp\left(\frac{nM^2}{\lambda\delta} + B\right)\right]^{KnM^3/(\lambda\delta)^2}. \tag{29}$$

The proof of Theorem 3 is given in Appendix B. Here we shall outline the main ideas and how they relate to the variables in expression (29).

We begin by noting that the convergence time is polynomial in ε^{-1}, while it is exponential in M and B. Notice that B may be small even when M is large. For example, in an n-person coordination game, there is a pure better reply path of length $n - 1$ from any pure strategy profile to a Nash equilibrium, hence $B = n - 1$, whereas M can be arbitrarily large. However, there are other weakly acyclic games in which B is exponentially larger than n and M. For example, Hart and Mansour (2010) constructed weakly acyclic n-person coordination games in which each player has just two strategies (so $M = 2n$) but the length of the better reply paths is of order 2^n. It is precisely this type of example that differentiates our framework from theirs: in the Hart–Mansour examples the underlying game grows, whereas in our set-up the underlying game is fixed and the population size grows.

The idea of the proof of Theorem 3 is as follows. Suppose that the process starts in some state $x^0 \in \chi$ that is not an ε-equilibrium. If x^0 is close to a pure strategy profile s of G, the argument is straightforward: with some probability, the shock realization variables will be realized in such a way that the shocks steer the process close to a better reply path that runs from s through a series of pure strategy profiles to a strict Nash equilibrium s^* of G. Such an equilibrium exists because δ-genericity implies that every pure Nash equilibrium is strict and by weak acyclicity there exists at least one pure Nash equilibrium. By assumption, this path has at most B "legs" or segments, along each of which exactly one population is switching from a lower payoff strategy to a higher payoff strategy. If, on the other hand, x^0 is *not* close to a pure strategy profile of G, we show how to construct a better reply path of bounded length to the vicinity of a pure strategy profile of G, and then apply the preceding argument.

The remainder of the proof involves estimating two quantities: (i) how long it takes to move along each leg of the paths constructed above, and (ii) how likely it is that the shock realization variables are realized in such a way that the required paths are followed to a close approximation. The first quantity (the rate of travel) is bounded below by $\lambda\delta$ times the minimum size of the populations that are currently switching from lower to higher payoff strategies. Although this estimation would appear to be straightforward, it is in fact quite delicate. The difficulty is that the process can get bogged down on paths that are nearly flat (there are almost no potential payoff gains for any player) but the process is in the vicinity of an unstable Nash equilibrium and does not converge to it. The second quantity, namely the log probability of realizing a given target path, is bounded by the number of distinct legs along the path (each of which corresponds to a specific shock) times the number of independent exogenous shock variables. The latter is of order $\frac{M^2}{2}$.

Putting all of these estimates together, we obtain the bound in expression (29). A particular implication is that the convergence time is bounded independently of N by a polynomial in ε^{-1}, where ε is the desired degree of approximation to equilibrium. The bound depends exponentially on the size of the game M and on the length of the better reply paths B. The exponential dependence seems inevitable given previous results in the literature such as Hart and Mansour (2010) and Babichenko (2014). Faster convergence may hold if the process is governed by a Lyapunov function, but this is much more restrictive than the conditions assumed here.

7. CONCLUSION

In this paper, we have studied the speed of convergence in population games when individuals use simple adaptive learning rules and the population size is large. The framework applies to weakly acyclic games, which include coordination games, games with strategic complementarities, dominance-solvable games, potential games, and many others with application to economics, biology, and distributed control.

Our focus has been on stochastic better reply rules in which individuals shift between strategies with probabilities that depend on the potential gain in payoff from making the switch. When these switching probabilities result from idiosyncratic shocks, weak acyclicity is not sufficient to achieve fast convergence; indeed, Theorem 1 shows that there exist very simple weakly acyclic games such that the expected time to come close to Nash equilibrium grows exponentially with the population size. This result is similar in spirit to earlier work on the computational complexity of learning Nash equilibrium (see, in particular, Hart and Mansour (2010) and Babichenko (2014)); the difference is that here we show that the problem persists even for games with extremely simple payoff structures. The nature of the argument is also fundamentally different from these earlier papers, which rely on results in communication complexity; here we use stochastic dynamical systems theory to obtain the result.

When the learning process is subjected to aggregate shocks, however, the convergence time can be greatly reduced; in fact, under suitable conditions, the convergence time is bounded above for all sufficiently large populations. Such shocks might result from intermittent interruptions to communication, or they might represent stochastic switching costs that retard the rate at which groups switch between strategies. For expositional simplicity, we have modeled these shocks as independent binary random variables, but similar results hold for many other distributions. The crucial property is that the shocks steer the process close to a target better reply path of the deterministic dynamic with positive probability.

The framework proposed here can also be extended to population games that are not representable as Nash population games. In this case, the analog of weak acyclicity is that, from any initial state, there exists a continuous better reply path that leads to the interior of the basin of attraction of some Nash equilibrium. Under suitable conditions on the aggregate shock distribution, the stochastic adjustment process will travel near such a path with positive probability. As the proof of Theorem 3 shows, the expected time it takes to traverse such a path depends critically on the payoff gains along the path. In the case of Nash population games, the interdependence index provides a lower bound on the payoff gains and hence on the expected convergence time. Analogous conditions on payoff gains along the better reply paths govern the expected convergence time in the more general case.

APPENDIX A: PROOF OF THE AUXILIARY RESULTS OF SECTION 5.2

LEMMA 4: *For every $\varepsilon > 0$, and $T > 0$, and every solution $z(\cdot)$ of (19), there exists $N_{T,\varepsilon}$ such that*

$$\forall N > N_{T,\varepsilon}, \quad \mathbb{P}\Big(\sup_{t\in[0,T]} \big\| X^N(t) - z(t)\big\| > \varepsilon\Big) < \varepsilon. \tag{30}$$

PROOF: Lemma 1 in Benaïm and Weibull (2003) implies the following:

CLAIM 4: *Let ρ be a revision protocol for the game G, and let $X^N(\cdot)$ be the stochastic process corresponding to ρ starting at state x. Let $\xi(t, x)$ be the semi-flow of the differential equation defined by* (2), *and let*

$$D^N(T, x) = \max_{0\le t\le T} \left\| X^N(t) - \xi(t, x) \right\|_\infty.$$

There exists a scalar $c(T)$ and a constant $\nu > 0$ such that, for any $\varepsilon > 0$, $T > 0$, and $N > \frac{\exp(\nu T)\nu T}{\varepsilon}$:

$$\mathbb{P}_x\big(D^N(T, x) \ge \varepsilon\big) \le 2N \exp\big(-\varepsilon^2 c(T)N\big), \tag{31}$$

where

$$c(T) = \frac{\exp(-2BT)}{8TA}.$$

(*Here A and B are constants that depend on the Lipschitz constant of ρ.*)

Let F be the event that the state at time $t = 0$ is (α, x) and during the time interval $[0, T]$ the shock realization remains constant. Let $(Y^N(t))_{0\le t\le T}$ be the process corresponding to the revision protocol $\bar\rho$ given by

$$\forall p, \forall i, j \in S^p, \quad \bar\rho^p_{ij}\big(U^p(x)\big) = \alpha^p_{ij} \cdot \rho^p_{ij}\big(U^p(x)\big). \tag{32}$$

Conditional on the event F, the process $(X^N(t))_{0\le t\le T}$ has the same distribution as the process $(Y^N(t))_{0\le t\le T}$.

Let $\beta : [0, T] \to \{0, 1\}^v$ be any shock realization. Using Claim 4, we shall approximate the distance between $X^N(T)$ and $z(T)$ given that the shock realization is β. We shall express this approximate distance as a function of the $k + 1$ distinct shocks of β in $[0, T]$. Let $(\tau_1, \ldots, \tau_k)$ be the sequence of distinct times at which the shock realization changes. It follows from equation (32) that along each interval $[\tau_l, \tau_{l+1})$, the process $X^N(\cdot)$ is distributed according to the stochastic process generated by the revision protocol $\bar\rho$ defined by

$$\forall p, \forall i, j \in S^p, \quad \bar\rho^p_{ij}\big(U^p(x)\big) = \rho^p_{ij}\big(U^p(x)\big)\beta^p_{ij}(\tau_l),$$

with initial condition $X^N(\tau_l)$. Note also that for any realization of the shock realization β, the protocol $\bar\rho$ has Lipschitz constants that are no greater than the Lipschitz constant for ρ.

Let $s(\cdot)$ be the piecewise continuous process defined as follows:

$$\dot s^p_i(t) = \sum_{j\in S^p} s^p_j(t)\rho^p_{ji}\big(U^p\big(s(t)\big)\big)\beta^p_{ji}(t) - \sum_{j\in S^p} s^p_i(t)\rho^p_{ij}\big(U^p\big(s(t)\big)\big)\beta^p_{ij}(t),$$

and for all $1 \le l \le k$, let $s(\tau_l) = X^N(\tau_l)$. Let $\tau_0 = 0$ and $\tau_{k+1} = T$. We then have

$$\begin{aligned}
(33)\qquad & \sup_{0\le t\le T} \|X^N(t) - z(t)\| \\
&\le \sup_{0\le t\le T} \big[\|X^N(t) - s(t)\| + \|s(t) - z(t)\|\big] \\
&= \max_{1\le l\le k+1} \sup_{\tau_{l-1}\le t\le \tau_l} \big[\|X^N(t) - s(t)\| + \|s(t) - z(t)\|\big].
\end{aligned}$$

Hence

$$\begin{aligned}
& \mathbb{P}\Big(\sup_{0\le t\le T} \|X^N(t) - z(t)\| > \varepsilon : \big(\beta(t)\big)_{0\le t\le T}\Big) \\
&\le \sum_{l=1}^{k+1} \mathbb{P}\Big(\sup_{\tau_{l-1}\le t\le \tau_l} \big[\|X^N(t) - s(t)\| + \|s(t) - z(t)\|\big] > \varepsilon\Big) \\
(34)\qquad &\le \sum_{l=1}^{k+1} \mathbb{P}\Big(\sup_{\tau_{l-1}\le t\le \tau_l} \|X^N(t) - s(t)\| > \frac{\varepsilon}{2} : \big(\beta(t)\big)_{0\le t\le T}\Big) \\
(35)\qquad &\quad + \sum_{l=1}^{k+1} \mathbb{P}\Big(\sup_{\tau_{l-1}\le t\le \tau_l} \|s(t) - z(t)\| > \frac{\varepsilon}{2} : \big(\beta(t)\big)_{0\le t\le T}\Big).
\end{aligned}$$

Note that expression (34) goes to zero in N, and by Claim 4 it does so uniformly for all β that have at most $k+1$ realizations. Therefore, there exists N_k such that, for every $N > N_k$ and for every shock realization with $k+1$ shocks or less, (34) is less than $\frac{\varepsilon}{2}$.

We claim that expression (35) also goes to zero in N. By Lemma 3, there exists a real number ν such that

$$\begin{aligned}
& \sup_{\tau_{l-1}\le t\le \tau_l} \|s(t) - z(t)\| \\
&\le \exp\big(\nu(\tau_l - \tau_{l-1})\big)\|s(\tau_{l-1}) - z(\tau_{l-1})\|.
\end{aligned}$$

Inductively, we obtain

$$\begin{aligned}
& \sup_{0\le t\le T} \|s(t) - z(t)\| \le \sum_{l=2}^{k+1} \exp(\nu\tau_l)\|s(\tau_1) - z(\tau_1)\| \\
&\le k\exp(\nu T)\|X^N(\tau_1) - z(\tau_1)\|,
\end{aligned}$$

which, by Claim 4, goes to zero uniformly in N. Since the random number of distinct shock realizations on $[0, T]$ is finite with probability 1, it follows that there exists $N_{T,\varepsilon}$ such that

$$\forall N > N_{T,\varepsilon}, \quad \mathbb{P}\Big(\sup_{t\in[0,T]} \big\|X^N(t) - z(t)\big\| > \varepsilon\Big) < \varepsilon.$$

This completes the proof of Lemma 4. *Q.E.D.*

LEMMA 5: *For every state $y \in \chi$ that is not an equilibrium, there exists a time T_y and a shock realization $\beta : [0, T_y] \to \{0, 1\}^v$ such that $z_y(T_y) \in \mathcal{B}_\varepsilon$.*

This result follows from Corollary 1, Lemma 10, and Lemma 9 where we provide an explicit construction of such a path.

APPENDIX B: PROOF OF THEOREM 3

For the sake of clarity, we shall restrict attention to the Smith revision protocol. The same proof applies with minor modifications to any responsive revision protocol. We shall henceforth write $X^N(\cdot)$ instead of $(\vec{A}(\cdot), X^N(\cdot))$. Recall that, by assumption, all payoffs lie in the unit interval.

Fix a generic, weakly acyclic game G with interdependence index $\delta_G = \delta > 0$. The value of δ will be fixed throughout the proof. Let $b = \max_{p\in\mathcal{P}} |S^p|$ be the maximal number of pure strategies available to any player. For every pure profile $s = (i^p)_{p\in\mathcal{P}} \in S$ and every state x, let $x(s) = \prod_{p\in\mathcal{P}} x^p_{i^p}$ denote the proportion of players associated with the profile s. Let

$$\big\|x^p - y^p\big\|_1 = \sum_{i\in S^p} \big|x^p_i - y^p_i\big|,$$

and

$$\|x - y\|_1 = \sum_{s\in S} \big|x(s) - y(s)\big|.$$

CLAIM 5: *Let $a = \frac{1}{8n}$. For every $\nu > 0$ and all $x, y \in \chi$,*

$$\forall p \in \mathcal{P}, \quad \big\|x^p - y^p\big\|_1 \le a\nu \quad \textit{implies} \quad \|x - y\|_1 \le \frac{\nu}{8}. \tag{36}$$

PROOF: First we shall establish the result for two populations. Let $x^1, y^1 \in \Delta^{m_1}$ be a pair of mixed strategies for player 1, and let $x^2, y^2 \in \Delta^{m_2}$ be a pair of mixed strategies for player 2. We shall show that if $\|x^1 - y^1\|_1 = \sum_{i=1}^{m_1} |x^1_i - y^1_i| \le \nu$ and $\|x^2 - y^2\| = \sum_{j=1}^{m_2} |x^2_j - y^2_j| \le \nu$, then

$$\sum_{i,j} \big|x^1_i x^2_j - y^1_i y^2_j\big| \le 2\nu.$$

By the triangle inequality,

$$\sum_{i,j}|x_i^1x_j^2 - y_i^1y_j^2| \leq \sum_{i,j}|x_i^1x_j^2 - x_i^1y_j^2| + \sum_{i,j}|x_i^1y_j^2 - y_i^1y_j^2|.$$

The left-hand-side summation equals

$$\sum_{i,j}|x_i^1x_j^2 - x_i^1y_j^2| = \sum_i x_i^1\sum_j|x_j^2 - y_j^2| = \sum_i x_i^1\|x^2 - y^2\|_1$$
$$= \|x^2 - y^2\|_1 \leq \nu.$$

Similarly,

$$\sum_{i,j}|x_i^1y_j^2 - y_i^1y_j^2| = \|x^1 - y^1\|_1 \leq \nu.$$

Hence

$$\sum_{i,j}|x_i^1x_j^2 - y_i^1y_j^2| \leq 2\nu.$$

For general n, it follows by induction that if, for every $p \in \mathcal{P}$, $\|x^p - y^p\|_1 \leq \nu$, then $\|x - y\|_1 < n\nu$. This concludes the proof of the claim. *Q.E.D.*

DEFINITION 9: For every population state x, let

$$\tilde{\mathbf{s}}^p(x) = \left\{i \in S^p : x_i^p \geq \frac{a\delta}{b}\right\}. \tag{37}$$

We shall say that $\tilde{\mathbf{s}}^p(x)$ consists of the strategies in S^p that are played by a *sizeable proportion* of the population as defined by the lower bound $\frac{a\delta}{b}$.

For every population p and state x, let

$$\bar{d}^p(x) = \max_{i\in\tilde{\mathbf{s}}^p(x)} u_i^p(x) - u^p(x).$$

(Recall that for every population p and state x, $u^p(x)$ denotes the average payoff to the members of p.) Let

$$\bar{d}(x) = \max_{p\in\mathcal{P}} \bar{d}^p(x).$$

Given $x \in \chi$, $q \in \mathcal{P}$, and $h > 0$, let $l, k \in S^q$ be two distinct strategies such that $x_l^q \geq h > 0$. Let $\tilde{x} = \tilde{x}(h, l, k, q)$ be the population state obtained from x

when a proportion h of population q switches from strategy l to strategy k, that is,

$$(38)\qquad \forall p \in \mathcal{P}, \quad \tilde{x}^p = \begin{cases} x^p & \text{if } p \neq q, \\ x^q + h\big(e_k^q - e_l^q\big) & \text{if } p = q. \end{cases}$$

LEMMA 6: *Given $x \in \chi$ such that $\bar{d}(x) \leq \frac{\delta}{2}$, let $\tilde{x} \in \chi$ be defined as in (38). Assume there exists at least one population different from q in which two distinct strategies are played by sizeable proportions. Then for at least one of these populations p and two distinct strategies $i, j \in \tilde{\mathbf{s}}^p(x)$,*

$$\big|\big[u_i^p(x) - u_i^p(\tilde{x})\big] - \big[u_j^p(x) - u_j^p(\tilde{x})\big]\big| \geq \frac{h\delta}{2}.$$

PROOF: We shall start with an observation that follows directly from the definition of $\tilde{\mathbf{s}}^p(x)$.

OBSERVATION 6: *For every population p and every $x^p \in X^p$, there exists $z^p \in X^p$ such that $\|z^p - x^p\|_1 \leq a\delta$, $z_i^p \geq x_i^p$ for every $i \in \tilde{\mathbf{s}}^p(x^p)$, and $z_i^p = 0$ for every $i \notin \tilde{\mathbf{s}}^p(x^p)$.*

To prove Lemma 6, choose z such that, for every $p \neq q$, the distribution z^p satisfies the conditions of Observation 6 with respect to x^p, and let $z^q = x^q$. Equation (36) implies that $\|z - x\|_1 \leq \frac{\delta}{8}$. Since all payoffs lie in the unit interval, it holds for every population p and strategy i, that

$$\big|u_i^p(x) - u_i^p(z)\big| \leq \frac{\delta}{8},$$

and that

$$\big|u^p(x) - u^p(z)\big| \leq \frac{\delta}{8}.$$

Therefore, since $\bar{d}(x) \leq \frac{\delta}{2}$, it follows that $\tilde{d}(z) \leq \frac{3\delta}{4}$. Since G is δ-generic, there exists a population $p \neq q$ and $i, j \in \mathbf{s}^p(z) = \tilde{\mathbf{s}}^p(x)$ such that

$$(39)\qquad \big|\big[u_i^p\big(e_k^q, z^{-q}\big) - u_i^p\big(e_l^q, z^{-q}\big)\big] - \big[u_j^p\big(e_k^q, z^{-q}\big) - u_j^p\big(e_l^q, z^{-q}\big)\big]\big| \geq \delta.$$

By the above, it follows that

$$\big|u_i^p\big(e_k^q, z^{-q}\big) - u_i^p\big(e_k^q, x^{-q}\big)\big| \leq \frac{\delta}{8}.$$

Similarly,

$$\left|u_i^p\left(e_l^q, z^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right)\right| \le \frac{\delta}{8},$$
$$\left|u_j^p\left(e_k^q, z^{-q}\right) - u_j^p\left(e_k^q, x^{-q}\right)\right| \le \frac{\delta}{8},$$
$$\left|u_j^p\left(e_l^q, z^{-q}\right) - u_j^p\left(e_l^q, x^{-q}\right)\right| \le \frac{\delta}{8}.$$

Therefore, by (39), we have

$$\left|\left[u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right)\right] - \left[u_j^p\left(e_k^q, x^{-q}\right) - u_j^p\left(e_l^q, x^{-q}\right)\right]\right| \ge \frac{\delta}{2}.$$

By definition of $\tilde{x}$, it follows that

$$\begin{aligned} u_i^p(x) - u_i^p(\tilde{x}) &= \left(x_k^q - \tilde{x}_k^q\right)u_i^p\left(e_k^q, x^{-q}\right) + \left(x_l^q - \tilde{x}_l^q\right)u_i^p\left(e_l^q, x^{-q}\right) \\ &= h\left[u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right)\right]. \end{aligned}$$

Similarly, we have $u_j^p(x) - u_j^p(\tilde{x}) = h[u_j^p(e_k^q, x^{-q}) - u_j^p(e_l^q, x^{-q})]$. Therefore,

$$\begin{aligned} &\left|\left[u_i^p(x) - u_i^p(\tilde{x})\right] - \left[u_j^p(x) - u_j^p(\tilde{x})\right]\right| \\ &\quad = h\left|\left[u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right)\right] - \left[u_j^p\left(e_k^q, x^{-q}\right) - u_j^p\left(e_l^q, x^{-q}\right)\right]\right| \\ &\quad \ge \frac{h\delta}{2}. \end{aligned}$$

This establishes Lemma 6. *Q.E.D.*

DEFINITION 10: Call a state $x \in \chi$ *nearly pure* if, in every population $p \in \mathcal{P}$, a unique strategy is played by a sizeable proportion, that is, $|\tilde{\mathbf{s}}^p(x)| = 1$ for all $p \in \mathcal{P}$.

If $x \in \chi$ is a nearly pure state, then by (37) at least $1 - a\delta$ of every population p is playing a unique strategy $i^p \in S^p$.

Recall that under the Smith dynamic, a path $z : [0, T] \to \chi$ is called *a continuous better reply path* with initial conditions $z(0) = x$ if there exists a shock realization $\beta : [0, T] \to \{0, 1\}^v$ such that

(40) $$\forall p, \forall i \in S^p,$$
$$\dot{z}_i^p = \sum_{j \in S^p} \left[z_j^p\left[u_i^p(z) - u_j^p(z)\right]_+ \beta_{ji}(t) - z_i^p\left[u_j^p(z) - u_i^p(z)\right]_+ \beta_{ij}(t)\right].$$

DEFINITION 11: For every $r > 0$, let A_r be the set of states $x \in \chi$ for which there exists a population q and two distinct strategies $k, l \in \tilde{\mathbf{s}}^q(x)$ such that $u_k^q(x) \geq u_l^q(x) + r$. For every population state $x \in \chi$, let $\sigma(x)$ denote the total *number* of strategies that are played by sizeable proportions of the respective populations:

$$\sigma(x) = \sum_{p \in \mathcal{P}} \left|\tilde{\mathbf{s}}^p(x)\right|.$$

Fix $\varepsilon \leq 1$ and recall that $\delta := \delta_G$ is the interdependence index of the given game G. Our main goal is to uniformly bound the convergence time L_ε for all sufficiently large population sizes N.

Let x be such that $d(x) \geq \frac{\varepsilon}{2}$. Let $r^* = \frac{a\delta^2}{16b}$. In the next few lemmas, we shall bound the elapsed time to get from such a state x to $\mathcal{B}_{\varepsilon/2}(s)$, the $\frac{\varepsilon}{2}$-basin of a strict Nash equilibrium s, via a continuous better reply path. Lemma 7, Lemma 8, and Corollary 1 bound the elapsed time to get from x to a state in A_{r^*}. Lemma 9 bounds the elapsed time to get from a state in A_{r^*} to a nearly pure state. Lemma 10 bounds the elapsed time to get from a nearly pure state to a nearly pure state in $\mathcal{B}_{a\delta}(s)$ of some strict Nash equilibrium s.

The constant r^* plays a central role when ε is small. The reasoning in Lemma 9 can be used to bound the elapsed time to get from a state x satisfying $d(x) \geq \frac{\varepsilon}{2}$ directly to a nearly pure state without going first to A_{r^*}. The difficulty is that this bound is poor for $\varepsilon \ll r^*$ and the polynomial dependence of the waiting time on ε^{-1} cannot be derived in this way. For this reason, we bound the waiting time to get to a nearly pure state from x in two steps. First, relying on δ-genericity, we provide an efficient bound for the process to go from x to A_{r^*}. Then we bound the waiting time to go from A_{r^*} to a nearly pure state. This yields an escape route from x to A_{r^*} that establishes a bound that is polynomial in ε^{-1}.

LEMMA 7: *Let $x \in \chi$ be such that $d(x) \geq \frac{\varepsilon}{2}$. Assume that there exist at least two distinct populations p, q such that $|\tilde{\mathbf{s}}^p(x)| \geq 2$ and $|\tilde{\mathbf{s}}^q(x)| \geq 2$. There exists a time $T \leq 1$, and a continuous better reply path $z : [0, T] \to \chi$ starting at x with a single shock, such that $z(T) = y \in A_{\delta\varepsilon/(16b)}$.*

PROOF: If $x \in A_{\delta\varepsilon/(16b)}$, then we are done. If $x \notin A_{\delta\varepsilon/(16b)}$, then, by definition of $A_{\varepsilon\delta/(16b)}$, for every population p and two sizeable strategies $i, j \in \tilde{\mathbf{s}}^p(x)$,

$$u_i^p(x) < u_j^p(x) + \frac{\delta\varepsilon}{16b}. \tag{41}$$

Since $d(x) \geq \frac{\varepsilon}{2}$, there exists a population q and a strategy $k \in S^q$ such that

$$u_k^q(x) \geq u^q(x) + \frac{\varepsilon}{2}. \tag{42}$$

It follows that

$$\sum_{j\in S^q} x_j^q\big(u_k^q(x) - u_j^q(x)\big) \geq \frac{\varepsilon}{2}.$$

Hence there exists a strategy l such that

$$x_l^q\big(u_k^q(x) - u_l^q(x)\big) \geq \frac{\varepsilon}{2b}. \tag{43}$$

Since all payoffs lie in the unit interval, it follows from (43) that $x_l^q \geq \frac{\varepsilon}{2b}$. Define a continuous better reply path $z(\cdot)$ from x with the coefficients $\beta_{lk}^q = 1$ and $\beta_{ij}^p = 0$ otherwise. Let T be the first time t such that $z_l^p(t) = x_l^p - \frac{\varepsilon}{4b}$. By (43), $\dot{z}_l^q \leq -\frac{\varepsilon}{4b}$ so long as $z_l^q \geq \frac{x_l^q}{2}$, hence $T \leq 1$.

We claim that since $x \notin A_{\delta\varepsilon/(16b)}$, it must be the case that $\bar{d}(x) \leq \frac{\delta}{2}$. Suppose by way of contradiction that $\bar{d}(x) > \frac{\delta}{2}$. Then for some strategy $k \in \tilde{\mathbf{s}}^p(x)$,

$$u_k^p(x) > u^p(x) + \frac{\delta}{2}.$$

A similar consideration as in equation (43) above shows that there exists a strategy $l \in S^p$ such that

$$x_l^p\big(u_k^p(x) - u_l^p(x)\big) > \frac{\delta}{2b}. \tag{44}$$

Since all payoffs lie in the unit interval, equation (44) implies that $x_l^p > \frac{\delta}{2b} > \frac{a\delta}{b}$ and $u_k^p(x) > u_l^p(x) + \frac{\delta}{2b}$. Hence, in particular, $x \in A_{\delta/(2b)}$. Since $\varepsilon \leq 1$, it follows that $x \in A_{\delta\varepsilon/(16b)}$, a contradiction.

Note that $z(T)$ plays the role of $\tilde{x}$ in Lemma 6 for $h = \frac{\varepsilon}{4b}$. Let $z(T) = \tilde{x}$. Since $\bar{d}(x) \leq \frac{\delta}{2}$, Lemma 6 implies that there exists a population $p \neq q$ and $i, j \in \tilde{\mathbf{s}}^p(x)$ such that

$$\big[u_i^p(x) - u_i^p(\tilde{x})\big] - \big[u_j^p(x) - u_j^p(\tilde{x})\big] \geq \frac{\delta\varepsilon}{8b},$$

which is equivalent to

$$\big[u_i^p(x) - u_j^p(x)\big] - \big[u_i^p(\tilde{x}) - u_j^p(\tilde{x})\big] \geq \frac{\delta\varepsilon}{8b}. \tag{45}$$

Inequality (41) implies that $u_i^p(x) - u_j^p(x) < \frac{\delta\varepsilon}{16b}$. It follows from this and (45) that

$$u_j^p(\tilde{x}) > u_i^p(\tilde{x}) + \frac{\delta\varepsilon}{16b}.$$

Hence $\tilde{x} \in A_{\delta\varepsilon/(16b)}$, as was to be shown. *Q.E.D.*

Let $x \in A_r$, where $0 < r \leq r^* = \frac{a\delta^2}{16b}$. Assume there exist at least two distinct populations at x, in each of which two strategies are played by sizeable proportions. The role of the next lemma is to estimate the elapsed time to get from x by a continuous better reply path to a state in A_{2r}.

LEMMA 8: *Let $0 < r \leq r^* = \frac{a\delta^2}{16b}$ and let $x \in A_r$. Assume that there exist at least two distinct populations p, q such that $|\tilde{\mathbf{s}}^p(x)| \geq 2$ and $|\tilde{\mathbf{s}}^q(x)| \geq 2$. There exist a time $T \leq \frac{16b}{a\delta^2}$, and a continuous better reply path $z : [0, T] \to \chi$ with a single shock such that $z(T) = y \in A_{2r}$.*

PROOF: We again use δ-genericity and Lemma 6. If $x \in A_{2r}$, we have nothing to prove. Thus we can assume that $x \notin A_{2r}$. Since $x \in A_r$, there exist a population q and two strategies $k, l \in \tilde{\mathbf{s}}^q(x)$ such that

$$u_k^q(x) \geq u_l^q(x) + r. \tag{46}$$

Define a continuous better reply path starting at x, such that $\beta_{lk}^q = 1$ and 0 otherwise. Let T be the first time such that $z_k^q(T) = x_k^q + \frac{8r}{\delta}$ and let $\tilde{x} = z(T)$. Note that in order to get from x to $\tilde{x}$, we need to transfer a proportion of $\frac{8r}{\delta}$ individuals from strategy l to strategy k. To estimate how long it takes, note that $r \leq \frac{a\delta^2}{16b}$ and $x_l^q \geq \frac{a\delta}{b}$, hence $\tilde{x}_l^q \geq \frac{a\delta}{2b}$. By construction of the better reply path $z(\cdot)$,

$$\dot{z}_l^q = z_l^q\big(u_l^q(x) - u_k^q(x)\big).$$

From (46), it follows that $\dot{z}_l^q \leq -\frac{ra\delta}{2b}$ as long as $z_l^q \geq \frac{a\delta}{2b}$. Since $z_l^q(t) \geq \frac{a\delta}{2b}$ for every $t \leq T$, we have

$$T \leq \frac{8r}{\delta} \cdot \frac{2b}{ra\delta} = \frac{16b}{a\delta^2}.$$

Since $x \notin A_{2r}$ and $r \leq r^* = \frac{a\delta^2}{16b}$, it must be the case that $\bar{d}(x) \leq \frac{\delta}{2}$. Otherwise, a similar derivation to that of equation (44) shows that $x \in A_{\delta/(2b)}$. Since $a, \delta \leq 1$, it follows that $x \in A_{2r^*} \subseteq A_{2r}$, a contradiction.

We can therefore apply Lemma 6 with $h = \frac{8r}{\delta}$. There exist a population $p \neq q$ and two strategies $i, j \in \tilde{\mathbf{s}}^p(x)$ such that

$$\big[u_i^p(x) - u_i^p(\tilde{x})\big] - \big[u_j^p(x) - u_j^p(\tilde{x})\big] \geq 4r.$$

Therefore

$$\big[u_j^p(\tilde{x}) - u_i^p(\tilde{x})\big] \geq 4r - \big[u_i^p(x) - u_j^p(x)\big].$$

Since $x \notin A_{2r}$,

$$u_j^p(x) + 2r > u_i^p(x).$$

Hence

$$u_j^p(\tilde{x}) \geq u_i^p(\tilde{x}) + 2r,$$

which implies that $\tilde{x} \in A_{2r}$. This concludes the proof of Lemma 8. *Q.E.D.*

REMARK B.1: Let x be a nearly pure state. By definition, there exists a pure strategy profile $s = (i^p)_{p\in\mathcal{P}} \in S$ such that a proportion at least $(1 - a\delta)$ of every population p is playing the strategy i^p. Define the pure state $y \in \chi$ such that $y^p = e^p_{i^p}$ for every p. By (36), $\|x - y\|_1 \leq \frac{\delta}{8}$. Since all payoffs lie in the unit interval, we also have

$$\forall p \in \mathcal{P}, \forall i \in S^p, \quad \left|u_i^p(x) - u_i^p(y)\right| \leq \frac{\delta}{8}. \tag{47}$$

Since y is a pure state and G has interdependence index δ,

$$\forall p \in \mathcal{P}, \forall i, j \in S^p, i \neq j, \quad \left|u_i^p(y) - u_j^p(y)\right| \geq \delta.$$

Therefore by inequality (47),

$$\forall p \in \mathcal{P}, \forall i, j \in S^p, i \neq j, \tag{48}$$
$$\left|u_i^p(x) - u_j^p(x)\right| \geq \left|u_i^p(y) - u_j^p(y)\right| - \frac{\delta}{4} \geq \frac{3\delta}{4}.$$

The following corollary of Lemma 7 and Lemma 8 bounds the elapsed time to get from a state $x \notin A_{r^*}$ that is not nearly pure to a state in A_{r^*}.

COROLLARY 1: *Let x be such that $d(x) \geq \frac{\varepsilon}{2}$. Assume that x is not nearly pure. There exist a time $T \leq 1 + 2\ln(\varepsilon^{-1})\frac{16b}{a\delta^2}$ and a continuous better reply path $z : [0, T] \to \chi$ starting at x, with at most $2\ln(\varepsilon^{-1}) + 1$ shocks, such that $z(T) = y \in A_{r^*}$.*

PROOF: If, at x, there exists a *unique* population p for which two strategies $i, j \in \tilde{\mathbf{s}}^p(x)$ are played by sizeable proportions, then by equation (48),

$$u_i^p(x) \geq u_j^p(x) + \frac{3\delta}{4}.$$

In particular, it follows that $x \in A_{a\delta^2/(16b)} = A_{r^*}$.

Case 1: $\bar{d}(x) > \frac{\delta}{2}$.

A similar argument to that in Lemma 8 shows that $x \in A_{\delta/(2b)} \subset A_{r^*}$.

Case 2: $\bar{d}(x) \leq \frac{\delta}{2}$.

There exist at least two distinct populations p, q such that $|\tilde{\mathbf{s}}^p(x)| \geq 2$, $|\tilde{\mathbf{s}}^q(x)| \geq 2$, and $\bar{d}(x) \leq \frac{\delta}{2}$. Lemma 7 implies that there exist a time $T \leq 1$ and a continuous better reply path $z : [0, T] \to \chi$ starting at x with a single shock, such that $z(T) = y_0 \in A_{\delta\varepsilon/(16b)}$.

If, at y_0, there exist a unique population p and two strategies $i, j \in \tilde{\mathbf{s}}^p(y_0)$ that are played by sizeable proportions, or if $\bar{d}(y) > \frac{\delta}{2}$, then we conclude as above that $y \in A_{a\delta^2/(16b)}$. Otherwise, there exist at least two distinct populations p, q such that $|\tilde{\mathbf{s}}^p(y)| \geq 2$, $|\tilde{\mathbf{s}}^q(y)| \geq 2$, and $\bar{d}(y) \leq \frac{\delta}{2}$. Lemma 8 implies that there exist a time $T \leq \frac{16b}{a\delta^2}$ and a continuous better reply path $z : [0, T] \to \chi$ starting at y_0, with a single shock, such that $z(T) = y_1 \in A_{2(\delta\varepsilon/(16b))}$.

We can apply this argument again. Namely, if there exist a unique population p and two strategies $i, j \in \tilde{\mathbf{s}}^p(y_1)$ that are played by sizeable proportions, or if $\bar{d}(y_1) > \frac{\delta}{2}$, then $y_1 \in A_{a\delta^2/(16b)}$. Otherwise, by Lemma 8, there exist a time $T \leq \frac{16b}{a\delta^2}$ and a continuous better reply path $z : [0, T] \to \chi$ starting at y_1, with a single shock, such that $z(T) = y_2 \in A_{4(\delta\varepsilon/(16b))}$.

By repeatedly applying the preceding argument, we conclude that there exist a time $T \leq 1 + k\frac{16b}{a\delta^2}$ and a continuous better reply path $z : [0, T] \to \chi$ starting at x, with at most $k + 1$ shocks, such that either $z(T) \in A_{a\delta^2/(16b)} = A_{r^*}$ or $z(T) \in A_{2^k(\delta\varepsilon/(16b))}$. Note that $2^k \frac{\delta\varepsilon}{16b} \geq \frac{a\delta^2}{16b}$ if

$$(49) \qquad k \geq \frac{\ln(\varepsilon^{-1}) + \ln(a\delta)}{\ln(2)}.$$

By Claim 5, $a = \frac{1}{8n} < 1$, and by assumption on the payoffs, $\delta \leq 1$. Hence, $\ln(a\delta) < 0$. Furthermore, $\ln(2) > \frac{1}{2}$. We conclude that if $k \geq 2\ln(\varepsilon^{-1})$, then (49) is satisfied and hence $2^k \frac{\delta\varepsilon}{16b} \geq \frac{a\delta^2}{16b}$. Hence, by time $1 + 2\ln(\varepsilon^{-1})\frac{16b}{a\delta^2}$, the process has reached $A_{a\delta^2/(16b)}$ with at most $1 + 2\ln(\varepsilon^{-1})$ shocks. This concludes the proof of Corollary 1. *Q.E.D.*

The role of the next lemma is to bound the elapsed time to get from $x \in A_{r^*}$ to a nearly pure state.

LEMMA 9: *Let $x \in A_{r^*}$. There exists a continuous better reply path $z : [0, T] \to \chi$ such that $z(0) = x$, there are at most $2M$ shocks in $[0, T]$, $z(T) = y$ is a nearly pure state, and $T \leq 2M\frac{64b^2}{a^2\delta^3}$.*

PROOF: *Case 1*: There exists a *unique* population p with more than one strategy that is played by a sizeable proportion.

In this case, all other populations are playing a nearly pure strategy. By Remark B.1, there is a unique strategy $i \in S^p$ such that, for every strategy $j \in S^p$ different from i,

$$(50) \qquad u_i^p(x) \geq u_j^p(x) + \frac{3\delta}{4}.$$

Define a continuous better reply path $z(\cdot)$ starting at x by letting $\beta^p_{ji} = 1$ for every $j \neq i$ and 0 otherwise. Thus $z(\cdot)$ reaches a nearly pure state in a time that is bounded above by $\frac{2b}{a\delta^2}$.

If Case 1 does not hold, then there exist at least two distinct populations at x, in each of which two strategies are played by sizeable proportions. Since $x \in A_{a\delta^2/(16b)}$, there exist a population q and two strategies $k, l \in \tilde{\mathbf{s}}^q(x)$ such that

$$u^q_k(x) \geq u^q_l(x) + \frac{a\delta^2}{16b}. \tag{51}$$

Let $h = x^q_l - \frac{a\delta}{2b}$ and let $\tilde{x}$ be defined as in equation (38) (note that $h \geq \frac{a\delta}{2b}$ since $l \in \tilde{\mathbf{s}}^q(x)$).

Case 2a: $\tilde{x} \in A_{a\delta^2/(16b)}$.

Define a continuous better reply path $z : [0, \infty) \to \chi$ that starts at x, and let $\beta^q_{lk} = 1$ and 0 otherwise. Recall that $\tilde{x}$ is obtained from x by a transfer of a proportion of h from strategy l to strategy k. Hence there exists a unique time T' such that $z(T') = \tilde{x}$. Clearly, $\sigma(\tilde{x}) = \sigma(x) - 1$. Since $\dot{z}^q_l \leq -(\frac{a\delta^2}{16b}\frac{a\delta}{2b})$ so long as $z^q_l \geq \frac{a\delta}{2b}$, and since $h \leq 1$, we get that

$$T' \leq \frac{2b}{a\delta} \cdot \frac{16b}{a\delta^2} = \frac{32b^2}{a^2\delta^3}.$$

Case 2b: $\tilde{x} \notin A_{a\delta^2/(16b)}$.

Let $z(\cdot)$ be the continuous path defined in Case 2a. Let t_0 be any time $t < T'$ such that $z(t_0) \in A_{a\delta^2/(32b)}$, and $z(t_0) \notin A_{a\delta^2/(16b)}$. (Such t_0 exists since $x \in A_{a\delta^2/(16b)}$ and $\tilde{x} \notin A_{a\delta^2/(16b)}$.) Let $w = z(t_0)$. Note that $\sigma(w) \leq \sigma(x)$. If there exists a unique population p at w with more than one strategy that is played by a sizeable proportion, then we are back in Case 1. If there is more than one such population, it follows as in the proof of Lemma 7 that $\bar{d}(w) \leq \frac{\delta}{2}$.

Since $w \in A_{a\delta^2/(32b)}$, there exist a population q and two strategies $k, l \in \tilde{\mathbf{s}}^q(w)$ such that

$$u^q_k(w) \geq u^q_l(w) + \frac{a\delta^2}{32b}.$$

Let $h = w^q_l - \frac{a\delta}{2b}$. Note that $h \geq \frac{a\delta}{2b}$ since $l \in \tilde{\mathbf{s}}^q(w)$. Let $\tilde{w} \in \chi$ be as defined in equation (38) for this value of h. By construction, $\tilde{w}^q_l = \frac{a\delta}{2b} < \frac{a\delta}{b}$ and hence $\sigma(\tilde{w}) = \sigma(w) - 1 \leq \sigma(x) - 1$. Define a continuous better reply path $z(\cdot)$ from w to $\tilde{w}$ by letting $\beta^q_{lk} = 1$ and $\beta^p_{ij} = 0$ otherwise. Let T be the first time such that $z(T) = \tilde{\omega}$. An argument like the one given for Case 2a shows that $T \leq \frac{64b^2}{a^2\delta^3}$. Recall that there exists at least one population different from q in which two distinct strategies are played by sizeable proportions. Therefore, since $\bar{d}(w) \leq \frac{\delta}{2}$, an argument like the one given for Lemma 8 shows that $\tilde{w} \in A_{a\delta^2/(16b)}$.

The two cases considered above demonstrate that there exists a continuous better reply path $z : [0, 2T] \to \chi$ starting at x, with at most two shocks, such that $T \le \frac{64b^2}{a^2\delta^3}$, and the state $z(2T)$ is either nearly pure or the following two conditions hold: (i) $\sigma(z(2T)) \le \sigma(x) - 1$, (ii) $z(2T) \in A_{a\delta^2/(16b)}$. A repeated application of the argument shows that we can construct a better reply path from x to a nearly pure state such that: (i) there are at most $2M$ shocks along the path, and (ii) the length of the path between two successive shocks is at most $\frac{64b^2}{a^2\delta^3}$. *Q.E.D.*

REMARK B.2: Let $s = (i^p)_{p\mathcal{P}}$ be a strict Nash equilibrium, and let $\varepsilon \le a\delta$. Recall that $\mathcal{B}_\varepsilon(s)$ is characterized by the minimal ϕ_ε such that, if a state x satisfies $x^p_{i^p} > \phi_\varepsilon$ for every population p, then i^p is a unique best reply at x and the deviation of x is at most ε. Let x satisfy $x^p_{i^p} \ge 1 - \varepsilon \ge 1 - a\delta$ for every population p. By equation (48) of Remark B.1,

$$\forall p \in \mathcal{P}, \forall i \ne i^p, \quad u^p_{i^p}(x) \ge u^p_i(x) + \frac{3\delta}{4}. \tag{52}$$

Therefore, i^p is the unique best reply at x for every population p. We claim that $x \in \mathcal{B}_\varepsilon(s)$. To see this, note that since all payoffs lie in the unit interval, it follows that the deviation $d^p(x)$ is bounded above by the proportion that is not playing i^p. Hence the deviation $d^p(x)$ satisfies $d^p(x) \le \varepsilon$.

LEMMA 10: *For every nearly pure state x, there exists a continuous better reply path $z : [0, T] \to \chi$ starting at x with at most B shocks, such that $z(T) = w$ is a nearly pure state that lies in $\mathcal{B}_{a\delta}$, and $T \le \frac{B4b}{3a\delta^2}$.*

PROOF: Since x is a nearly pure state, there exists a pure strategy profile $s = (i^p)_{p\in\mathcal{P}}$ such that $x^p_{i^p} \ge 1 - a\delta$ for every population p. Assume first that s is a pure Nash equilibrium. By δ-genericity, the payoff to every two distinct pure strategy profiles is different for every player. It follows that every pure Nash equilibrium is strict. From Remark B.2, it follows that $x \in \mathcal{B}_{a\delta}(s)$, so in this case we are done.

Suppose, on the other hand, that s is not a Nash equilibrium. By weak acyclicity, there exists a pure strict better reply path in G from s to *some* Nash equilibrium s', which by the above must be a strict Nash equilibrium. Denote this path by $(s_1, \dots, s_k)$. Clearly, the length of this path is bounded by B. We shall construct a continuous better reply path that stays close to this pure better reply path and is of bounded length.

Let $s_2 = (j^p)_{p\in\mathcal{P}}$ be the second element in the pure better reply path. Let y_2 be such that, for every population p, $y^p_2 = e^p_{j^p}$. By definition of a better reply path and δ-genericity, there exists a unique population p such that $i^p \ne j^p$ and

$$u^p_{j^p}(y) \ge u^p_{i^p}(y) + \delta = u^p_{i^p}(y_2) + \delta.$$

Again by equation (48) we obtain

$$u^p_{j^p}(x) \geq u^p_{i^p}(x) + \frac{3\delta}{4}. \tag{53}$$

Define the first leg of the continuous better reply path $z : [0, t_2] \to \chi$ by letting $z(0) = x$ and $\beta^p_{i^p j^p}(t) = 1$ for every t. Set all other β^q_{kl} equal to 0. Note that by (53),

$$\forall t \in [0, t_2], \quad u^p_{j^p}\big(z(t)\big) \geq u^p_{i^p}\big(z(t)\big) + \frac{3\delta}{4}.$$

Let t_2 be the first time t such that $z^p_{i_p}(t) = \frac{a\delta}{b}$. We claim that

$$\forall t \in [0, t_2], \quad \dot{z}^p_{i_p}(t) \leq -\frac{a\delta}{b} \cdot \frac{3\delta}{4}.$$

To see this, note that, by (40), the derivative $\dot{z}^p_{i^p}(t)$ is bounded by $z^p_{i^p}(t)$ times $u^p_{i^p}(z(t)) - u^p_{j^p}(z(t))$. Therefore, $t_2 \leq \frac{4b}{3a\delta^2}$. Let $x_2 = z(t_2)$. By construction, x_2 is a nearly pure state such that $\|x_2 - y_2\|_1 \leq \frac{\delta}{8}$. We can repeat the same argument iteratively and extend the continuous better reply path until it reaches a nearly pure state that lies in $\mathcal{B}_{a\delta}(s')$.

The length of the pure better reply path is bounded by B, and the length of every leg in the better reply path is at most $\frac{4b}{3a\delta^2}$. Therefore, the overall length of the continuous better reply path is at most $\frac{B4b}{3a\delta^2}$. This concludes the proof of Lemma 10. *Q.E.D.*

Call a state x *strictly pure* if x is nearly pure and there exists a strict Nash equilibrium $s = (i^p)_{p\in\mathcal{P}}$ such that the unique strategy played by a sizeable proportion of every population p is i_p.

LEMMA 11: *Let x be a strictly pure state, and let $s = (i^p)_{p\in\mathcal{P}}$ be the corresponding strict Nash equilibrium. Assume that for every population p, it holds that $x^p_{i^p} \geq 1 - r$. There exist a time $T \leq \frac{4}{3\delta}$, and a continuous better reply path $z : [0, T] \to \chi$ such that $z(0) = x$, $z(T) = y$, and for every population p, $y^p_{i^p} \geq 1 - \frac{r}{2}$.*

PROOF: Equation (48) of Remark B.1 implies that, for every population p and strategy $j \neq i_p$,

$$u^p_{i^p}(x) \geq u^p_j(x) + \frac{3\delta}{4}. \tag{54}$$

Define a continuous better reply path $z(\cdot)$ by letting $\beta^p_{ji^p} = 1$ for every population p and $j \neq i^p$, and 0 otherwise. Note that by equation (54) for every population p and $j \neq i^p$,

$$\dot{z}_j^p(t) \leq -\frac{3\delta}{4} z_j^p(t).$$

Therefore, as long as $z_j^p(t) \geq \frac{x_j^p}{2}$,

$$\dot{z}_j^p(t) \leq -\frac{3\delta}{4} \cdot \frac{x_j^p}{2}. \tag{55}$$

Let $T = \frac{4}{3\delta}$, and let $y = z(T)$. It follows from equation (55) that $y_j^p \leq \frac{x_j^p}{2}$ for every population p and strategy $j \neq i^p$. Hence for every population p,

$$y_{i^p}^p = 1 - \sum_{j \neq i^p} y_j^p \geq 1 - \frac{\sum_{j \neq i^p} x_j^p}{2} \geq 1 - \frac{r}{2}.$$

This concludes the proof of Lemma 11. *Q.E.D.*

LEMMA 12: *Let $\frac{\varepsilon}{2} \leq \frac{a\delta^2}{16b}$, and let x be such that $d(x) \geq \frac{\varepsilon}{2}$. Assume that $X^N(0) = x$. There exists a continuous better reply path $z : [0, T] \to \chi$ starting at x, with at most $1 + 3\ln(\varepsilon^{-1}) + 2M + B$ shocks such that $z(T) \in \mathcal{B}_{\varepsilon/2}$ and*

$$T = \frac{Cb}{a\delta^2}\left[\ln\left(\varepsilon^{-1}\right) + B + \frac{Mb}{a\delta}\right], \tag{56}$$

for some constant $C > 0$.

PROOF: Corollary 1 implies that there exist a time $T_1 \leq 1 + 2\ln(\varepsilon^{-1})\frac{16b}{a\delta^2}$ and a continuous better reply path $z : [0, T_1] \to \chi$ starting at x with at most $2\ln(\varepsilon^{-1}) + 1$ shocks such that $z(T_1) = y \in A_{a\delta^2/(16b)}$.

Let $y \in A_{a\delta^2/(16b)}$. By Lemma 9, there exist a time $T_2 \leq 2M\frac{64b^2}{a^2\delta^3}$, and a continuous better reply path $z : [0, T_2] \to \chi$ starting at y with at most $2M$ shocks such that $z(T_2) = w$ is a nearly pure state.

Let w be a nearly pure state. By Lemma 10, there exist a time $T_3 \leq \frac{B4b}{3a\delta^2}$ and a continuous better reply path $z : [0, T_3] \to \chi$ starting at w with at most B shocks such that $z(T_3)$ is a nearly pure state and $z(T_3) \in \mathcal{B}_{a\delta}(s)$ for some strict Nash equilibrium $s = (i^p)_{p \in \mathcal{P}}$.

Let $w' \in \mathcal{B}_{a\delta}(s)$ be a nearly pure state. By definition, $(w')_{i^p}^p \geq 1 - a\delta$ for every population p. By Lemma 11, there exist a time $T' \leq \frac{4}{3\delta}$ and a continuous better reply path $z : [0, T'] \to \chi$ starting at w' with a single shock such that $z(T') = w''$ where $(w'')_{i^p}^p \geq 1 - \frac{a\delta}{2}$. By applying this argument repeatedly, we conclude that there exist a time $T_4 \leq \ln(\varepsilon^{-1})\frac{4}{3\delta}$ and a continuous better reply path starting at

w', with at most $\ln(\varepsilon^{-1})$ shocks such that $y = z(T_4)$ and $y_{ip}^p \geq 1 - \frac{\varepsilon}{2}$. Hence, by Remark B.2, $y \in \mathcal{B}_{\varepsilon/2}$.

Overall, we have shown that, from any state x such that $d(x) \geq \frac{\varepsilon}{2}$, there exist a time

$$T \leq 1 + 2\ln(\varepsilon^{-1})\frac{16b}{a\delta^2} + 2M\frac{16b^2}{a^2\delta^3} + \frac{B4b}{3a\delta^2} + \ln(\varepsilon^{-1})\frac{4}{3\delta},$$

and a continuous better reply path $z : [0, T] \to \chi$ with at most $1 + 2\ln(\varepsilon^{-1}) + 2M + B + \ln(\varepsilon^{-1})$ shocks such that $z(T) \in \mathcal{B}_{\varepsilon/2}$. This concludes the proof of Lemma 12. *Q.E.D.*

Lemma 12 implies the following.

COROLLARY 2: *There exist a constant K', and a time*

$$T = \frac{K'}{a\delta^2}\left[Bb + \ln(\varepsilon^{-1}) + \frac{Mb^2}{a\delta}\right], \tag{57}$$

such that for all sufficiently large N, if $X^N(0) = x$ with $d(x) \geq \frac{\varepsilon}{2}$, then $X^N(T) \in \mathcal{B}_{\varepsilon/2}$ with probability at least $\exp(-K'Tv)$ *by time T.*

PROOF: First we shall estimate the probability that the stochastic process $z(\cdot)$ defined in equation (19) reaches $\mathcal{B}_{\varepsilon/2}$ by time T. On each leg of the continuous better reply path, the shock variables must take on a specific realization and stay fixed until the process reaches the next leg. Since the number of shock variables is v, the length of the continuous better reply path is T, and the number of distinct legs is at most $1 + 3\ln(\varepsilon^{-1}) + 2M + B$, these events occur with probability at least $\exp(-K_1 v[\ln(\varepsilon^{-1}) + 2M + B])\exp(-K_1 vT)$ for some constant $K_1 > 0$.

By Lemma 4, the stochastic process $X^N(\cdot)$ lies arbitrarily close to $z(\cdot)$ with a probability that goes to 1 with N. Hence we can find a constant $K_2 > 0$ such that, for all sufficiently large N, the process $X^N(\cdot)$ reaches $\mathcal{B}_{\varepsilon/2}$ by time T with probability at least

$$\exp(-K_2 v[\ln(\varepsilon^{-1}) + 2M + B])\exp(-K_2 vT).$$

Finally, since $T = \frac{Cb}{a\delta^2}[\ln(\varepsilon^{-1}) + B + \frac{Mb}{a\delta}]$ and both $a, \delta \leq 1$, there is a constant K' such that

$$\exp(-K_2 v[\ln(\varepsilon^{-1}) + 2M + B])\exp(-K_2 vT) \leq \exp(-K' vT).$$

This concludes the proof of the corollary. *Q.E.D.*

THEOREM 3: Let G be a weakly acyclic game with interdependence index $\delta > 0$, and let ρ be a responsive revision protocol with response rate $\lambda > 0$. There exists a constant K independent of G, such that, for every $\varepsilon > 0$, the convergence time L_ε is at most

$$K\left[\varepsilon^{-1}\exp\left(\frac{nM^2}{\lambda\delta}+B\right)\right]^{KnM^3/(\lambda\delta)^2}. \tag{58}$$

PROOF: Given the hypothesis and $\varepsilon > 0$, consider the process $(X^N(t))_{t\geq 0}$ starting from an arbitrary state $X^N(0)$. We shall say that a time t is *bad* if $d(X^N(t)) \geq \frac{\varepsilon}{2}$; otherwise, t is *good*.

Corollary 2 of Lemma 12 shows that there are a time T and a probability $q = \exp(-K'Tv)$ such that, if t is bad, then the probability is at least q that $X^N(t+T) \in \mathcal{B}_{\varepsilon/2}$, and hence all times from $t+T$ on are good.

As in the proof of Theorem 2, it follows that for any length of time $L > 0$, the expected proportion of bad times in the interval $[0, L]$ is at most

$$\frac{T}{Lq^2}. \tag{59}$$

Since the deviation of the process is bounded by 1 for all bad times the expected deviation of the process on $[0, L]$ is less than ε if $\frac{T}{Lq^2} \leq \frac{\varepsilon}{2}$. Hence the convergence time L_ε satisfies the inequality

$$L_\varepsilon \leq \frac{2T}{\varepsilon q^2}. \tag{60}$$

Since $q = \exp(-K'Tv)$, we have

$$L_\varepsilon \leq 2\varepsilon^{-1}T\exp(2K'Tv). \tag{61}$$

Since $T \leq \exp(T)$, there is a constant K'' such that

$$L_\varepsilon \leq 2\varepsilon^{-1}\exp(K''Tv). \tag{62}$$

From Corollary 2, we also know that

$$T = \frac{K'}{a\delta^2}\left[Bb + \ln(\varepsilon^{-1}) + \frac{Mb^2}{a\delta}\right]. \tag{63}$$

From (62) and (63), we deduce that there is a constant $\tilde{K}$ such that

$$L_\varepsilon \leq \left[\varepsilon^{-1}\exp\left(Bb + \frac{Mb^2}{a\delta}\right)\right]^{\tilde{K}v/(a\delta^2)}. \tag{64}$$

The final step is to bound a, b, and v. We know from Claim 5 that $a = \frac{1}{8n}$. The maximum number of strategies available to any given population is certainly less than the total number of strategies, hence $b < M$. The number of shock variables, v, is less than the total number of pairs of strategies, hence $v < \frac{M^2}{2}$. Finally let us recall that we chose $\lambda = 1$ to economize on notation, hence we need to replace δ by $\lambda\delta$. Making these substitutions, we deduce that for a suitably defined constant K,

$$L_\varepsilon \leq K\left[\varepsilon^{-1}\exp\left(\frac{nM^2}{\lambda\delta} + B\right)\right]^{KnM^3/(\lambda\delta)^2}. \tag{65}$$

This concludes the proof of Theorem 3. *Q.E.D.*

APPENDIX C: PROOF OF PROPOSITION 1 AND PROPOSITION 2

C.1. *Proof of Proposition 2*

Let $G = (\mathcal{P}, (S^p)_{p\in\mathcal{P}})$ be a game structure. The players in G are the elements of $\mathcal{P}$. Call $\tilde{G} = (\mathcal{P}, (\tilde{S}^p)_{p\in\mathcal{P}})$ a *subgame* of G if $\tilde{G}$ is obtained from G by restricting the strategy set of every player p to the nonempty subset $\tilde{S}^p \subseteq S^p$. A subgame is *nontrivial* if, for at least two players $p_1, p_2 \in \mathcal{P}$, the size of $\tilde{S}^{p_1}$ and $\tilde{S}^{p_2}$ is at least 2.

LEMMA 13: *Let $\tilde{G} = (\mathcal{P}, (\tilde{S}^p)_{p\in\mathcal{P}})$ be a subgame of G. Fix a player q with distinct strategies $\{k, l\} \not\subset \tilde{S}^q$. There exists a generic set of payoffs for G, such that for every player $p \neq q$, every pair of distinct strategies $i, j \in \tilde{S}^p$, and every equilibrium x of $\tilde{G}$,*

$$\left|\left[u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right)\right] - \left[u_j^p\left(e_k^q, x^{-q}\right) - u_j^p\left(e_l^q, x^{-q}\right)\right]\right| > 0. \tag{66}$$

PROOF: Fix a player $p \neq q$ and two distinct strategies $i, j \in \tilde{S}^p$.

Case 1: $\{k, l\} \cap \tilde{S}^q = \emptyset$.

By a known result in game theory, the subgame $\tilde{G}$ has finitely many equilibria for a full Lebesgue measure set of payoffs for $\tilde{G}$ (see Harsanyi (1973)). Fix such a payoff vector, and let E be the corresponding finite set of equilibria in $\tilde{G}$. Let Γ_k^p be the vector space of all payoffs to player p in G when player q plays strategy k, and let Γ_l^p be similarly defined.

Let $\bar{\Gamma}_{\tilde{G}}$ denote the subspace of payoffs to strategy profiles *other* than those defining $\tilde{G}$. We claim that, for every $x \in E$, there is a generic set of payoffs in $\bar{\Gamma}_{\tilde{G}}$ such that inequality (66) holds strictly. To see this, note that for a fixed $x^{-q} \in X^{-q}$, the set of all payoffs that satisfy (66) as an equality defines a lower dimensional subspace of payoffs in $\Gamma_k^p \times \Gamma_l^p$, which is a subspace of $\bar{\Gamma}_{\tilde{G}}$. Therefore, for any given $x \in E$, the set of payoffs that satisfy (66) as an equality has

Lebesgue measure zero in $\bar{\Gamma}_{\tilde{G}}$. Since E is finite for a generic set of payoffs in $\tilde{G}$, and inequality (66) holds strictly in any equilibrium of $\tilde{G}$ for a generic set of payoffs in $\bar{\Gamma}_{\tilde{G}}$, it follows from Fubini's theorem that there is a generic set of payoffs for G such that inequality (66) holds strictly for every equilibrium of $\tilde{G}$. This concludes the proof of Lemma 13 for case 1.

Case 2: $\{k, l\} \cap \tilde{S}^q \neq \emptyset$.

Without loss of generality, assume that $k \in \tilde{S}^q$ and $l \notin \tilde{S}^q$. As before, there is a generic set of payoffs in $\tilde{G}$ for which $\tilde{G}$ has a finite number of equilibria. Fix any such payoffs u for $\tilde{G}$ and let E denote the finite set of equilibria. Fix $x \in E$. Let $u_i^p(e_k^q, x^{-q}) = \alpha$ and $u_j^p(e_k^q, x^{-q}) = \beta$. Inequality (66) is satisfied as an equality if and only if

$$u_j^p(e_l^q, x^{-q}) - u_i^p(e_l^q, x^{-q}) = u_j^p(e_k^q, x^{-q}) - u_i^p(e_k^q, x^{-q}) = \beta - \alpha.$$

This equality defines a lower dimensional hyperplane in Γ_l^p, and hence has Lebesgue measure zero in $\bar{\Gamma}_{\tilde{G}}$. An application of Fubini's theorem establishes Lemma 13 for case 2. *Q.E.D.*

LEMMA 14: *Let $\tilde{G} = (\mathcal{P}, (\tilde{S}^p)_{p\in\mathcal{P}})$ be a nontrivial subgame of G. Given a player q with distinct strategies $\{k, l\} \subset \tilde{S}^q$, there is a generic set of payoffs for G such that, for every fully mixed Nash equilibrium of $\tilde{G}$, there exist a player p and two distinct strategies $i, j \in \tilde{S}^p$ such that inequality* (66) *holds.*

PROOF: $\tilde{G}$ has finitely many equilibria for a full Lebesgue measure set of payoffs for $\tilde{G}$. Fix such a payoff vector. If there are no fully mixed equilibria in $\tilde{G}$, we have nothing to prove. Otherwise, let E' be the finite set of fully mixed equilibria of $\tilde{G}$.

Assume by way of contradiction that (66) is violated for some $x \in E'$, every player $p \neq q$, and all $i, j \in \tilde{S}^p$. Let $0 < h < x_l^q$. Define a new mixed strategy y^q for player q as follows:

$$y_m^q = \begin{cases} x_m^q & \text{if } m \neq k, l, \\ x_l^q - h & \text{if } m = l, \\ x_k^q + h & \text{if } m = k. \end{cases}$$

Let $y = (y^q, x^{-q})$. For every player $p \neq q$ and $i, j \in \tilde{S}^p$,

$$u_i^p(y) = u_i^p(x) + h\big(u_i^p(e_k^q, x^{-q}) - u_i^p(e_l^q, x^{-q})\big) \tag{67}$$

$$= u_j^p(x) + h\big(u_j^p(e_k^q, x^{-q}) - u_j^p(e_l^q, x^{-q})\big) = u_j^p(y). \tag{68}$$

Equality (67) follows from the definition of y. Since x is a fully mixed Nash equilibrium, $u_i^p(x) = u_j^p(x)$. By assumption, (66) does not hold for x, hence

$u_i^p(e_k^q, x^{-q}) - u_i^p(e_l^q, x^{-q}) = u_j^p(e_k^q, x^{-q}) - u_i^p(e_l^q, x^{-q})$, from which equality (68) follows. Therefore, y is also a fully mixed equilibrium of $\tilde{G}$. Thus we can generate infinitely many equilibria of $\tilde{G}$, which contradicts the assumption that $\tilde{G}$ has finitely many equilibria. *Q.E.D.*

PROOF OF PROPOSITION 2: For any $x^p \in X^p$, let $\mathbf{s}^p(x) \subseteq S^p$ be the set of strategies in the support of x^p. Let Γ be a full measure set of payoffs such that both Lemma 13 and Lemma 14 hold with respect to every player q and every pair of distinct strategies $\{k, l\} \subset S^q$. Assume further that, for every payoff vector in Γ, every two distinct pure strategy profiles yield different payoffs for every player p. Let $u \in \Gamma$. We shall show that there exists a constant $\delta > 0$ such that G is δ-generic. The first condition of δ-genericity clearly holds for some $\delta > 0$ (see Definition 6). As for the second condition, assume by contradiction that it does not hold. Then there exists a sequence $\{x_m\}_{m=1}^{\infty}$ of mixed strategy profiles such that the following two properties hold:

(i) for every m,

$$\tilde{d}(x_m) \leq \frac{1}{m}; \tag{69}$$

(ii) for every m, there exist a player q_m and $\{k_m, l_m\} \subset S^{q_m}$ such that, for every player $p \neq q_m$ with two distinct strategies $i, j \in \mathbf{s}^p(x_m)$,

$$\begin{aligned} &\left|\left[u_i^p\left(e_{k_m}^{q_m}, x_m^{-q_m}\right) - u_i^p\left(e_{l_m}^{q_m}, x_m^{-q_m}\right)\right] - \left[u_j^p\left(e_{k_m}^{q_m}, x_m^{-q_m}\right) - u_j^p\left(e_{l_m}^{q_m}, x_m^{-q_m}\right)\right]\right| \\ &\quad \leq \frac{1}{m}. \end{aligned} \tag{70}$$

By taking subsequences, we can assume that x_m converges to some profile x. We can further assume by taking subsequences that the q_m's are constant, say $q_m = q$, and the $\{k_m, l_m\}$ are constant, say $\{k_m, l_m\} = \{k, l\}$. We can further assume that $\mathbf{s}^p(x_m)$ is fixed for every player p and m.

Case 1: $\{k, l\} \not\subset \mathbf{s}^q(x)$.

Define a subgame $\tilde{G}$ of G by letting $\tilde{S}^p = \mathbf{s}^p(x_m)$ for every player $p \neq q$, and let $\tilde{S}^q = \mathbf{s}^p(x)$. By (69), x is an equilibrium of $\tilde{G}$. Therefore, Lemma 13 implies that, for every player $p \neq q$ and $i, j \in S^p$,

$$\left|\left[u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right)\right] - \left[u_j^p\left(e_k^q, x^{-q}\right) - u_j^p\left(e_l^q, x^{-q}\right)\right]\right| > 0.$$

This stands in contradiction to inequality (70).

Case 2: $\{k, l\} \subset \mathbf{s}^q(x)$.

Define a subgame $\bar{G}$ by letting $\bar{S}^p = \mathbf{s}^p(x)$ for every player p. We claim that there exists a player $p \neq q$ such that $|\mathbf{s}^p(x)| \geq 2$. To see this, note that x is an equilibrium of $\bar{G}$ such that $x_l^q, x_k^q > 0$. Suppose by way of contradiction that $|\mathbf{s}^p(x)| = 1$ for every $p \neq q$. By assumption, every two pure strategy profiles yield a different payoff for player q. We conclude that q has a unique best reply at x, which is impossible because x is an equilibrium and both x_k^q, x_l^q are positive. Hence $\bar{G}$ is nontrivial. Since x is a fully mixed equilibrium of $\bar{G}$, Lemma 14 implies that there exist a player p and two distinct strategies $i, j \in \bar{S}^p$ such that

$$\left|\left[u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right)\right] - \left[u_j^p\left(e_k^q, x^{-q}\right) - u_j^p\left(e_l^q, x^{-q}\right)\right]\right| > 0.$$

This again contradicts inequality (70), and completes the proof of Proposition 2. *Q.E.D.*

C.2. *Proof of Proposition 1*

PROPOSITION 1: *Equilibrium convergence holds for a generic subset of weakly acyclic population games G.*

PROOF: We start by showing that if, for every N, the game G^N is weakly acyclic, then equilibrium convergence holds. Let ρ be a revision protocol. If G^N is weakly acyclic, then for every population state $x \in \chi^N$, there exists a better reply path to some pure Nash equilibrium y_x of G^N. This path has positive probability under the corresponding stochastic process $X^N(\cdot)$. Hence, for every state $x \in \chi^N$ and every time t, there exists a probability $p_x > 0$ such that

$$\mathbb{P}\big(X^N(t+1) = y_x | X^N(t) = y\big) = p_x.$$

Let $p = \min_{x \in \chi^N} p_x$. It follows that for every integer T, the probability is at most $(1-p)^T$ that the process has not reached an equilibrium state by time T. Therefore, equilibrium convergence holds for G.

It remains to be shown that if G is weakly acyclic and δ-generic for some $\delta > 0$ (see Section 6.1, Definition 6), then, for every N, the game G^N is weakly acyclic. Thus, we need to show that, for every N and every $x \in \chi^N$, there exists a better reply path to an equilibrium of G^N.

Let $\sigma(x) = \sum_{p \in \mathcal{P}} |\mathbf{s}^p(x)|$ be the size of the support of x, that is, the number of pairs (i, p) such that $x_i^p > 0$. We shall prove the claim by induction on $\sigma(x)$. The smallest value of $\sigma(x)$ is n. In such a state, all players in each population p are playing the same pure strategy. Let $s \in S$ denote the corresponding pure strategy tuple. By definition of a weakly acyclic game, there exists a better reply path $(s_1, \ldots, s_k) \in S^k$ in G such that $s_1 = s$ and s_k is an equilibrium. We can now define a better reply path in G^N: at every stage, all members of the corresponding population revise their strategy choice to the one prescribed by the better reply path in G. This better reply path terminates at s_k, which is an equilibrium of G^N.

Now let $x \in \chi^N$ be a state such that $\sigma(x) = c > n$. If x is an equilibrium, then we have nothing to prove. If x is not an equilibrium, we shall show that there exists a better reply path from x to some state y such that $\sigma(y) \le c - 1$. We can then use the induction hypothesis to complete the proof.

Since x is not an equilibrium, there must be a population q and a pure best reply strategy $k \in S^q$ such that $u_k^q(x) > u^q(x)$. If there exists such a population q with $|\mathbf{s}^q(x)| \ge 2$, then there must be a strategy $l \in S^q$ with $x_l^q > 0$ such that $u_k^q(x) > u_l^q(x)$. In this case, we can define a better reply path from x by letting all members of population q revise their strategy choice to k, and the resulting state y must satisfy $\sigma(y) \le c - 1$.

If this case does not hold, then for every population q that is not in equilibrium (i.e., some members are playing a suboptimal strategy), $|\mathbf{s}^q(x)| = 1$. It follows that $\tilde{d}(x) = 0$ (see expression (27) and the definition of $\tilde{d}$ immediately following). Thus there is at least one out-of-equilibrium population q *all* of whose members are playing a suboptimal strategy l, where $u_k^q(x) > u_l^q(x)$ for some $k \in S^q$. Let w be the state obtained from x by having all members of population q revise their strategy to k. Since $\tilde{d}(x) = 0 < \delta$, δ-genericity implies that the impact of q on some population $p \ne q$ with $|\mathbf{s}^p(x)| \ge 2$ is at least δ. Hence there exist two distinct strategies $i, j \in \mathbf{s}^p(x)$ such that

$$\left|u_i^p\left(e_k^q, x^{-q}\right) - u_i^p\left(e_l^q, x^{-q}\right) - \left[u_j^p\left(e_k^q, x^{-q}\right) - u_j^p\left(e_l^q, x^{-q}\right)\right]\right| \ge \delta.$$

This is equivalent to

$$\left|\left[u_i^p(w) - u_i^p(x)\right] - \left[u_j^p(w) - u_j^p(x)\right]\right| \ge \delta. \tag{71}$$

Since $\tilde{d}(x) = 0$, $\tilde{d}^p(x) = 0$, and therefore $u_i^p(x) = u_j^p(x)$. Hence equation (71) implies

$$\left|u_i^p(w) - u_j^p(w)\right| > 0.$$

Thus $\sigma(x) = \sigma(w) = c$ and we are back in the earlier case which has already been established. This completes the proof of Proposition 1. *Q.E.D.*

REFERENCES

BABICHENKO, Y. (2012): "Completely Uncoupled Dynamics and Nash Equilibria," *Games and Economic Behavior*, 76, 1–14. [628]

——— (2014): "Query Complexity of Approximate Nash Equilibria," in *STOC'14—Proceedings of the 46th Annual ACM Symposium on Theory of Computing*. New York: ACM, 535–544. [631, 634,651,652]

BENAÏM, M., AND J. WEIBULL (2003): "Deterministic Approximation of Stochastic Evolution in Games," *Econometrica*, 71, 873–903. [635,637,640,645,652]

BJORNERSTEDT, J., AND J. WEIBULL (1996): "Nash Equilibrium and Evolution by Imitation," in *The Rational Foundations of Economic Behavior*, ed. by K. Arrow. London: Macmillan, 155–171. [629,634]

BOROWSKI, H., AND J. R. MARDEN (2014): "Fast Convergence for Time-Varying Semi-Anonymous Potential Games," Unpublished Manuscript. [628]

BOROWSKI, H., J. R. MARDEN, AND E. W. FREW (2013): "Fast Convergence in Semi-Anonymous Potential Games," in *IEEE Conference on Decision and Control*. [628]

CHIEN, S., AND A. SINCLAIR (2011): "Convergence to Approximate Nash Equilibria in Congestion Games," *Games and Economic Behavior*, 71, 315–327. [628,629]

ELLISON, G. (1993): "Learning, Local Interaction and Coordination," *Econometrica*, 61 (5), 1047–1071. [628]

ELLISON, G., D. FUDENBERG, AND L. A. IMHOF (2014): "Fast Convergence in Evolutionary Models: A Lyapunov Approach," Unpublished Manuscript. [628]

FABRIKANT, A., A. D. JAGGARD, AND M. SCHAPIRA (2013): "On the Structure of Weakly Acyclic Games," *Theory of Computing Systems*, 53, 107–122. [628,633]

FOSTER, D. P., AND H. P. YOUNG (2003): "Learning, Hypothesis Testing, and Nash Equilibrium," *Games and Economic Behavior*, 45, 73–96. [628]

——— (2006): "Regret Testing: Learning to Play Nash Equilibrium Without Knowing You Have an Opponent," *Theoretical Economics*, 1, 341–367. [628]

GERMANO, F., AND G. LUGOSI (2007): "Global Nash Convergence of Foster and Young's Regret Testing," *Games and Economic Behavior*, 60, 135–154. [628]

GOLUB, B., AND M. O. JACKSON (2012): "How Homophily Affects the Speed of Learning and Best Response Dynamics," *Quarterly Journal of Economics*, 127, 1287–1338. [628]

HARSANYI, J. (1973): "Oddness of the Number of Equilibrium Points: A New Proof," *International Journal of Game Theory*, 2, 235–250. [670]

HART, S., AND Y. MANSOUR (2010): "How Long to Equilibrium? The Communication Complexity of Uncoupled Equilibrium Procedures," *Games and Economic Behavior*, 69, 107–126. [631,634,650-652]

HART, S., AND A. MAS-COLELL (2003): "Uncoupled Dynamics Do not Lead to Nash Equilibrium," *American Economic Review*, 93, 1830–1836. [628]

——— (2006): "Stochastic Uncoupled Dynamics and Nash Equilibrium," *Games and Economic Behavior*, 57, 286–303. [628]

HARTMAN, P. (2002): *Ordinary Differential Equations* (Second Ed.). Philadelphia, PA: SIAM. [645]

HOFBAUER, J., AND W. SANDHOLM (2007): "Evolution in Games With Randomly Disturbed Payoffs," *Journal of Economic Theory*, 132, 47–69. [628]

HOFBAUER, J., AND K. SIGMUND (1998): *Evolutionary Games and Population Dynamics*. Cambridge: Cambridge University Press. [627]

HOFBAUER, J., AND J. SWINKELS (1996): "A Universal Shapley Example," Discussion Paper, University of Vienna and Northwestern University. [628]

KREINDLER, G. E., AND H. P. YOUNG (2013): "Fast Convergence in Evolutionary Equilibrium Selection," *Games and Economic Behavior*, 80, 39–67. [628,629,634]

LEONARD, R. J. (1994): "Reading Cournot, Reading Nash: The Creation and Stabilisation of the Nash Equilibrium," *The Economic Journal*, 104, 492–511. [629]

MARDEN, J. R., AND J. S. SHAMMA (2012): "Revisiting Log-Linear Learning: Asynchrony, Completeness, and Payoff-Based Implementation," *Games and Economic Behavior*, 75, 788–808. [628,629]

——— (2014): "Game Theory and Distributed Control," in *Handbook of Game Theory*, Vol. 4, ed. by H. P. Young and S. Zamir. Amsterdam: Elsevier. [627,629,634]

MARDEN, J. R., H. P. YOUNG, G. ARSLAN, AND J. S. SHAMMA (2009): "Payoff-Based Dynamics for Multiplayer Weakly Acyclic Games," *SIAM Journal on Control and Optimization*, 48 (1), 373–396. [628]

MONTANARI, A., AND A. SABERI (2010): "The Spread of Innovations in Social Networks," *Proceedings of the National Academy of Sciences of the United States of America*, 107, 20196–20201. [628]

NASH, J. (1950): "Non-Cooperative Games," Ph.D. Thesis, Mathematics Department, Princeton University. [629]

OYAMA, D., W. H. SANDHOLM, AND O. TERCIEUX (2015): "Sampling Best Response Dynamics and Deterministic Equilibrium Selection," *Theoretical Economics*, 10, 243–281. [628]

PRADELSKI, B. S. R. (2015): "Decentralized Dynamics and Fast Convergence in the Assignment Game," Discussion Paper 700, Department of Economics, University of Oxford. [631]
PRADELSKI, B. S. R., AND H. P. YOUNG (2012): "Learning Efficient Nash Equilibrium in Distributed Systems," *Games and Economic Behavior*, 75, 882–897. [628]
SANDHOLM, W. H. (2009): "Large Population Potential Games," *Journal of Economic Theory*, 144, 1710–1725. [629]
——— (2010a): "Pairwise Comparison Dynamics and Evolutionary Foundations of Nash Equilibrium," *Games*, 1, 3–17. [629]
——— (2010b): *Population Games and Evolutionary Dynamics*. Cambridge, MA: MIT Press. [627,632,634,639]
SANDHOLM, W. H., AND M. STAUDIGL (2015): "Large Deviations and Stochastic Stability in the Small Noise Double Limit," *Theoretical Economics* (forthcoming). [631]
SHAH, D., AND J. SHIN (2010): "Dynamics in Congestion Games," in *ACM SIGMETRICS*. [628, 629,634]
SMITH, M. J. (1984): "The Stability of a Dynamic Model of Traffic Assignment—An Application of a Method of Lyapunov," *Transportation Science*, 18, 245–252. [630,636]
TERCIEUX, O. (2006): "p-Best Response Sets," *Journal of Economic Theory*, 131, 45–70. [628]
WEIBULL, J. (1995): "The Mass-Action Interpretation of Nash Equilibrium," Report. [629,631]
YOUNG, H. P. (1993): "The Evolution of Conventions," *Econometrica*, 61, 57–84. [630]
——— (1998): *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*. Princeton, NJ: Princeton University Press. [628]
——— (2009): "Learning by Trial and Error," *Games and Economic Behavior*, 65, 626–643. [628]
——— (2011): "The Dynamics of Social Innovation," *Proceedings of the National Academy of Sciences of the United States of America*, 108 (Suppl. 4), 21285–21291. [628]

The Faculty of Industrial Engineering and Management, Technion—Israel Institute of Technology, Technion City, Haifa, 3200003, Israel; iarieli@tx.technion.ac.il
and
Nuffield College, Oxford, OX1 1NF, U.K.; peyton.young@economics.ox.ac.uk.

Co-editor Matthew O. Jackson handled this manuscript.

Manuscript received April, 2012; final revision received August, 2015.