Queen's Economics Department Working Paper No. 981

Comparative Dynamics

James Bergin
Queen's University

Dan Bernhardt
University of Illinois

Department of Economics
Queen's University
94 University Avenue
Kingston, Ontario, Canada
K7L 3N6

2-1999

Comparative Dynamics

by

James Bergin

Department of Economics
Queen's University
Kingston
Canada K7L 3N6

Dan Bernhardt

Department of Economics
University of Illinois
Champaign, IL 61820
USA

This draft: February 1999

Abstract

This paper develops a dynamic evolutionary model in which agents make choices on the basis of relative performance criteria. We distinguish two classes of learned behavior: imitative dynamics and a new class of dynamics, "introspective dynamics." Under imitative dynamics, agents compare payoffs of different agents in the population and tend to adopt the strategies of those agents who earned greater payoffs — agents imitate more successful behavior in the population. Under introspective dynamics, agents compare their own current with past payoffs when they engaged in different actions, and tend to select actions that provide higher payoffs. With introspective dynamics, under weak regularity conditions (satisfied by a rich class of economic environments), the stochastically stable set of states is contained in the set of Nash equilibria, providing a novel rationale or justification for the prediction of Nash equilibrium behavior. With imitative dynamics, under mild regularity conditions (again satisfied in many economic environments) there is a unique stochastically stable state, but it is *not* a Nash equilibrium. Rather, each agent's action maximizes the *difference* between his payoffs and those of other agents. Paradoxically, comparing stochastically stable states across dynamics, agent payoffs are lower for imitative than introspective dynamics — mimicking best practice turns out to be counterproductive. We consider both forms of dynamics in the contexts of games satisfying strategic substitutes and strategic complements.

1 Introduction.

Underlying recent dynamic evolutionary models (e.g. Kandori, Mailath and Rob (1993) and Young (1993)) are fundamental questions: What is the appropriate way to model agents' responses to their experiences, and what behavioral predictions are implied by the resulting dynamic? In this paper we motivate agents' behavior in terms of the information they possess and the way they utilize this information.

We study models in which agents make choices on the basis of relative performance criteria. We call the associated class of dynamic adjustment rules *comparative dynamics*, the defining feature of which is that agents make choices based on a comparison of the performance of actions across the population or historically (or both). We distinguish two classes of learning behavior: learning from the experience of others and learning from one's own experience. The dynamic adjustment that arises from these forms of behavior we call *imitative* and *introspective* dynamics respectively. Underlying each dynamic are different specifications about what each agent knows and can do.

In an imitative dynamic, agents compare the payoffs from their own actions with those of other agents. Agents behave as if they believe other agents' experience is relevant for them, tending to imitate the actions of more successful agents. For an imitative dynamic to be a reasonable model of economic behavior, the economic environment must allow agents to see the actions and payoffs of those whom they imitate, and their payoff functions and feasible action spaces must correspond. This contrasts with an introspective dynamic, where an agent compares his payoffs from current and past actions, and tends to select actions that have had greater past payoffs. Agents behave as if they believe their own past experience is relevant for them, even though the actions of other agents may have changed over time, altering their payoffs. Hence, introspective dynamics may reasonably describe the economic behavior of heterogeneous agents, even when agents know neither population distributions nor how successful other agents have been. Both dynamics specify behavior that reflects comparative evaluations by agents: with imitative dynamics the comparison of actions and payoffs is across agents at a moment in time; with introspective dynamics the comparison is within an agent at different points in time.[1]

This paper characterizes how the choice of dynamic affects long–run equilibrium outcomes for a rich class of economic environments that encompasses: (1) Oligopoly output games, including both for homogeneous and differentiated goods, (2) R&D games with spillovers, (3) Team production, (4) Private provision of a public good or bad (e.g. tragedy of the commons), and (5) Production games with (e.g. negative pollution) externalities. Each dynamic makes sharp predictions about outcomes, predictions that accord with outcomes of models featuring greater agent rationality assumptions. However, the predicted outcomes are *very* different.

With an imitative dynamic, the unique stochastically stable state is *not* a Nash equilibrium. Instead, each agent's action maximizes the *difference* between his payoffs and those of other agents. An example of this outcome is in Vega–Redondo (1997) who shows that imitative dynamics in a homogeneous good oligopoly economy lead to the Walrasian competitive equilibrium outcome, where price equals marginal

[1] This comparative flavor distinguishes the dynamics from best response or fictitious play dynamics in which each agent determines his future action by taking the distribution of actions (or average of history of actions) in the population as given and myopically optimizing against it.

cost. Since, imitative dynamics capture learned behavior rather than some higher degree of rationality, it is perhaps not surprising that Nash equilibria do not obtain. This contrasts with best response dynamics in which agents are rational, optimizing against the current population distribution, so that rest points are *necessarily* Nash equilibria.

We show that with introspective dynamics, the stochastically stable set of states is contained in the set of Nash equilibria. Thus, introspective dynamics provide a novel rationale for the prediction of Nash equilibrium behavior in the long run, a justification based solely on learning from past experience without higher order rationality assumptions. A key insight of our paper is that the outcome predicted by dynamics based on models of "learning from experience" depends crucially on the assumed structure of information. We then contrast agent payoffs across dynamics in the stochastically stable set of states and show that agent payoffs are lower for imitative than introspective dynamics for certain classes of games. This suggests a paradoxical result: in the long–run, mimicking best practice leads to lower agent payoffs.

The paper is organized as follows. In what follows we briefly review some of the related work. In section 2 we present the basic framework. Sections 3 and 4 develop formal models for introspective dynamics and imitative dynamics respectively. Section 5 provides the intuition underlying the different outcomes for the two dynamics. Section 6 discusses sufficient conditions on two important economic environments for our analysis to hold. Section 7 considers aggregate welfare associated with stable states under both dynamics and section 8 provides a collection of economic examples for which the analysis applies. Section 9 concludes.

Related research with an imitative flavor includes Schlag (1997) and Vega-Redondo (1997). Schlag considers a bandit problem in which an agent observes observes his own action choice and payoff in the previous period and that of some sampled individual. In this framework the best imitative rule has an agent imitate the observed individual with probability proportional to the amount by which the observed person does better. Under this rule the aggregate dynamic is approximately the replicator dynamic. Vega-Redondo considers a homogeneous good oligopoly game where agents imitate the most successful firms and shows that the stochastically stable outcome is that where agents select the competitive or Walrasian outcome.

A second related approach incorporates the spirit of stimulus–response models: actions that do well have higher probability of being chosen in subsequent periods. For example, Borgers and Sarin (1997) consider two agents playing a fixed game over time. A player's current state is a distribution over his strategies, and the player's experience (current action and payoff) determines next period's state for the player: next period's state is a weighted average of the current state and the state that puts all mass on the current choice, with the weight on current choice proportional to the payoff received from that choice. As the time between repetitions of the game becomes small, they show that the evolution of strategies follows the replicator dynamic.

Another approach develops aspirations–based models of behavior (Karandikar *et al.* (1998)), in which agents have a status quo action and an endogenous aspiration or prospect level against which actions are evaluated. In the prisoners' dilemma game, Karandikar *et al.* show with sufficiently slow updating of aspirations, both players cooperate most of the time.

2 Preliminaries.

The environment consists of n agents. The strategy space for agent i is a finite grid, $C^i = \{a_i, a_i + \Delta, a_i + 2\Delta, \ldots, a_i + k_i\Delta\}$, ($k_i$ integer valued, $\Delta > 0$), and agent i has payoff function, $\pi^i : C \to R$ ($C = \times_{i=1}^n C^i$). One may view the grid as the base strategy space or as a discretization of a continuous action space.

2.1 Static Equilibrium Concepts.

In a Nash equilibrium, agents can find no improving deviation in the set of possible alternatives:

Definition 1 *A vector $q \in C$ is a Nash equilibrium if for all agents i,*

$$\pi^i(q^i, q^{-i}) \geq \pi^i(\tilde{q}^i, q^{-i}), \quad \forall \tilde{q}^i \in C^i.$$

Denote the set of Nash equilibria by NE.

When studying imitative behavior it is natural to consider alternative notions of equilibrium where agents think in terms of relative payoffs: a choice for a player is good if the choices of other players are no more successful in payoff terms. Consideration of relative payoffs gives rise naturally to a "getting ahead of the Jones" behavior that may not obtain in models featuring greater agent rationality. In a dynamic environment, stable action profiles with imitative behavior must be such that deviations lead the deviator to consider the choice of an opponent superior.

Say that $q \in C$ is *envy-free* if $\forall i, j$, $\pi^i(q) = \pi^j(q)$: all agents receive the same payoff. In an envy-free equilibrium, if every agent can determine that no one is getting a higher payoff than they, then each agent is contented with the status quo. As a solution concept, this criterion is not demanding — in a symmetric game, any profile where all agents make the same choice is an envy-free equilibrium. This notion of choice based on relative payoffs may be extended to incorporate the impact of deviations on relative payoffs:

Definition 2 *A vector q is a relative equilibrium if for $j \neq i$*

$$\pi^i(q) = \pi^j(q)$$

and $\forall \tilde{q}^i \in C^i$, there exists a $j \neq i$ such that

$$\pi^i(\tilde{q}^i, q^{-i}) \leq \pi^j(q^j, \tilde{q}^i, q^{-\{i,j\}}).$$

Relative equilibrium identifies the variational restrictions on payoffs consistent with long run stability of imitative behavior.[2] Say that a relative equilibrium is strict if the inequality is strict for all $\tilde{q}^i \neq q$. Let

[2] Aumann (1961) provides another equilibrium notion called "twisted equilibrium", for games where players interests are partially or fully opposed, so that the payoffs of other players may enter into the strategic calculations of a player. See also Kats and Thisse (1992).

RE denote the set of relative equilibria and RE_S the set of strict relative equilibria. For the cross–agent comparison of payoffs by an agent i to make economic sense, agents in i's comparison group must have the same strategy spaces and the same payoff functions. A relative equilibrium is not only envy-free (take $\tilde{q}_i = q_i$, so for each pair i, j, $\pi^i(q) \leq \pi^j(q)$ implying $\pi^i(q) = \pi^j(q)$), but in addition, any deviation must make the deviator envious of some other agent's payoff. For example, in the context of symmetric actions, at a relative equilibrium a deviator prefers to return to the common action.

2.2 Dynamics and Mutations.

Nash equilibrium and relative equilibrium are static concepts; how equilibrium is reached involves a fuller description of agents — how they assess and react to circumstances. Dynamic models of individual behavior must specify each agents' information and how an agent reacts to or processes that information. Write s to denote a state[3] of the system and let $I^i(s)$ denotes the (partial) information available to i at this state. For example, with imitative dynamics, $I^i(s)$ includes at least the knowledge of the action and payoff profile in i's comparison group (or at least the highest payoff and associated action); while with introspective dynamics, $I^i(s)$ includes the knowledge of the agent's recent past history of actions and payoffs; and with best response dynamics it includes at least the distribution of actions in the population. A behavioral rule for agent i, φ_i, assigns a choice given i's information: $\varphi_i(s)$. Under best response dynamics, φ_i is the best response to the population distribution, whereas with imitation dynamics, φ_i is the choice of the most successful player (in the previous period).[4] At a state, s, the behavioral rules for each agent, $\varphi = (\varphi_1, \ldots, \varphi_n)$, are aggregated according to some rule, moving the system to a new state: $s' = \mathcal{A}(\varphi, s)$. Write $p^{\varphi}_{ss'}$ to denote the transition probability from s to s', and use P^{φ} to denote the corresponding transition matrix.

In this framework, mutations are modeled as perturbations in the choices of agents. At state s, the choice of i is $\varphi_i(s) \in C^i$. Under mutation, with high probability this is unchanged and with small probability some other choice in C^i is drawn. Formally, with probability $\alpha_{\varphi_i(s)} \approx 1$, $\varphi_i(s)$ is chosen, and with small probability $\alpha_q > 0$, choice q is made. The impact of mutations is to perturb the transition matrix P^{φ} to some matrix $\tilde{P}^{\varphi}$. An invariant distribution of the system $\tilde{P}^{\varphi}$ is approximately an invariant distribution of P^{φ} and is identified as a (stochastically) stable population distribution. In what follows, we examine stochastically stable distributions for the comparative dynamic systems described above.

We assume that agents have finite memories. Hence, all the information required to model the evolution of the system is contained in the vector $(\mathbf{q}, \Pi) = ([q_{t-l}, \pi_{t-l}], \ldots, [q_t, \pi_t])$. There is some redundancy here since knowledge of $\underline{q}_t$ allows one to compute π_t: Π may be computed from $\mathbf{q}$. Using this observation, one may define the state space as $S = C^{l+1}$, $C = [\times_{i=1}^{n} C^i]$.[5] Formally:

Definition 3 *A state of the system at time $t+1$ is a vector $s = \mathbf{q} \in S$ containing the actions in the past $l+1$ periods.*

---

[3] For example, a state could be the action chosen by each agent in the previous period distribution.

[4] This is not the only possible formulation of imitative behavior. An alternative is that agents imitate those more successful than themselves, but not necessarily the most successful. Our analysis extends to this case.

[5] Alternatively, one may carry payoffs in the state space. This is appropriate if the payoff function is unknown to players — they just observe payoffs. These issues do not affect the dynamics of the models developed here.

For $q \in C = \times_{i=1}^{n} C^i$, let s_q denote the state $(q, q, \ldots, q)$, the history where the choices in the previous l periods are all the same and define:

$$S_A = \{s \in S \mid s = s_q, \text{ some } q \in C\}.$$

Thus, S_A is the set of aggregate states where the choices are constant over time.

3 Introspective Dynamics.

We first develop a model of behavior in which players base future choices on the historically most successful decisions, comparing their own current and past performance in determining their best choice. This introspective model of behavior allows for the possibility that agents have quite limited information: each player need only know their own action history and the payoffs they received. At date $t+1$, agent i remembers his own[6] actions and payoffs from the previous $l+1 > 1$ periods: $\mathbf{q}^i(t) = (q^i_{t-l}, q^i_{t-l+1}, \ldots, q^i_t)$ and $\Pi^i(t) = (\pi^i_{t-l}, \pi^i_{t-l+1}, \ldots, \pi^i_t)$. We sometimes drop the t index and denote i's information by $(\mathbf{q}^i, \Pi^i)$. Agent i observes $(\mathbf{q}^i, \Pi^i)$ before choosing next period's action.

If at period t, the history or state is $s_t = \mathbf{q}(t)$, then define $B^i_t(\mathbf{q}(t))$ to be the set of choices for i, during the period $t-l$ to t that yielded the highest payoff for i:

$$B^i_t(\mathbf{q}(t)) = \{q^i_\tau \mid \pi^i_\tau = \max_{t-l \leq r \leq t} \pi^i_r,\ t-l \leq \tau \leq t\}$$

Agent i is said to optimize introspectively if i selects an action in period $t+1$ from the set $B^i_t(\mathbf{q}(t))$ when changing strategy. We model the decision selection process of agent i at time $t+1$ as a distribution $\gamma^i(t) = (\gamma^i_{t-l}, \ldots, \gamma^i_t)$, on the history of actions of the agent, with support $B^i_t(\mathbf{q}(t))$:

Definition 4 *Agent i optimizes introspectively at state $s = \mathbf{q}(t)$ if the choice of action for period $t+1$ is drawn from a distribution $\gamma^i(s)$ with support $B^i_t(\mathbf{q}(t))$.*

This simple procedure fully describes the agent's decision–making process: past experiences are compared and successful choices are repeated. We require no assumptions of inertia in behavior. The following remark notes that the results are very robust to alternative specifications of introspective behavior.

Remark 1 This is not the only plausible model of introspective optimization. An issue is the way in which an agent evaluates different action choices when the same action choice generated different payoffs at different dates because of experimentation by other agents. For example, suppose action q generated past payoffs of π^q_t, π^q_{t-1}, while action $\tilde{q}$ generated payoff $\pi^{\tilde{q}}_{t-2}$, where $\pi^q_{t-1} > \pi^{\tilde{q}}_{t-2} > \pi^q_t$: How does the agent evaluate the choices of q and $\tilde{q}$? The formulation that we consider has the agent select action q over $\tilde{q}$. An alternative formulation would have an agent place a positive probability on playing an action, if the maximum payoff from playing that action exceeded the minimum payoff from playing any other action over the past l periods. This formulation admits more choice candidates. The *main* observation to

[6] Throughout, we use superscripts to denote specific agents and subscripts to denote time periods or to label distinct vectors in C: q^i_t is the choice of i at time t, q_a and q_b are elements of C. Similarly, $\underline{q} = (q^1, \ldots, q^n)$ and $\underline{q}_t = (q^1_t, \ldots, q^n_t)$. If the vector of actions taken by agents is $(q^1, \ldots, q^n)$, then agent i receives the payoff $\pi^i(q^1, \ldots, q^n)$. At time t, the action choice–payoff profile is $(q_t, \pi_t) = (q^1_t, \ldots, q^n_t; \pi^1_t, \ldots, \pi^n_t)$.

be made is that the results in the paper are *not* sensitive to the formulation of introspective optimization *provided* the introspective dynamic satisfies two conditions: (a) There is positive probability that the choice yielding the highest payoff for agent in the previous $l+1$ periods is chosen next period and (b) There is zero probability that a strictly dominated choice (a choice that always yields a lower payoff in the previous $l+1$ periods than some other action) is chosen. ■

Remark 2 It is worth noting that each of the possible formulations of introspective behavior demands little in the way of calculation and rationality from agents — agents must only be able to compare past experiences and identify which past actions were more successful. This behavior arises naturally in environments where agents believe past payoffs are good predictors of future payoffs. ■

Combining agents' decisions determines a law governing the evolution of states. Given state $s_t = \mathbf{q}(t) = (q_{t-l}, \ldots, q_t)$, agents make decisions that lead to a selection of q_{t+1}, and the new state $s_{t+1} = \mathbf{q}(t+1) = (q_{t-l+1}, \ldots, q_{t+1})$. Given states s and s', let $p_{ss'}$ be the probability of reaching state s' next period, given that the current state is s.

Definition 5 *Let $P = \{p_{ss'}\}_{s,s' \in S}$. Call the matrix P the introspective dynamics transition matrix.*

Mutations modify this in the following way. Given a state $s = (q_1, \ldots, q_{l+1})$, suppose that next period the state reached under introspective dynamics is $s' = (q_2, \ldots, q_{l+1}, q')$, so that in particular, there is a positive probability that agent i will choose q_i'. Under mutation, with probability $(1-\epsilon)$, q_i' is chosen; and with probability ϵ the next choice for agent i is drawn from a distribution θ^i with full support on C^i. If $\tilde{q}_i$ is agent i's action choice the next period, then

$$\tilde{q}_i = \begin{cases} q_i'' & \text{with probability } (1-\epsilon) + \epsilon\theta^i_{q_i'} \\ \hat{q}_i & \text{with probability } \epsilon\theta^i_{\hat{q}_i},\ \hat{q}_i \neq q_i'. \end{cases}$$

Remark 3 In the present context the interpretation of mutations as experimentation in conjunction with introspective optimization is quite natural: agents usually choose actions that were more successful in the past, but occasionally try new things. This melds naturally with the observation that the distribution, γ^i, may be state dependent, so that, for example, recent history may be assigned more weight. We may also allow the distribution over experiments θ^i to be state dependent, $\theta^i(s)$, without affecting the results. In that case, agents experiment with small probability, ϵ, but given that they do experiment, they choose in a state–dependent way (so that, for example, they may be more likely to take actions that are close to their past action). ■

The associated (perturbed) transition process is denoted P_ϵ. Note that P_ϵ is not strictly positive because at state $s = (\bar{q}_1, \ldots, \bar{q}_{l+1})$, if $s' = (\hat{q}_1, \ldots, \hat{q}_{l+1})$, and $(\bar{q}_2, \ldots, \bar{q}_{l+1}) \neq (\hat{q}_1, \ldots, \hat{q}_l)$ then $\tilde{p}_{ss'} = 0$, where $P_\epsilon = \{\tilde{p}_{ss'}\}_{s,s' \in S}$: past actions are fixed, and an action j periods past at date t becomes $j+1$ periods past at date $t+1$. However, the matrix, P_ϵ is irreducible and has invariant distribution ξ_ϵ. For $n > l$, P_ϵ^n is a strictly positive matrix. In particular P_ϵ is aperiodic so that regardless of the initial distribution, $\bar{\xi}$, on states, $\bar{\xi} P_\epsilon^n \to \xi_\epsilon$. Let $S_E \subset S_A$ be the set of states $\{s_q\}$ such that q is an equilibrium. Similarly, S_S represents the set of states associated with strict Nash equilibria.

Definition 6 *A set of Nash equilibria Q^* is absorbing if $q \in Q^*$ and $\pi^i(\tilde{q}_i, q_{-i}) \geq \pi^i(q_i, q_{-i})$ imply that $(\tilde{q}_i, q_{-i}) \in Q^*$.*

In words, Q^* is absorbing if it is not possible to move out of the set Q^* with a sequence of unilateral weakly improving deviations: given $q \in Q^*$, any deviation by any agent that is non–disimproving must lead to a point in Q^*. Call an equilibrium absorbing if it is contained in a set of absorbing equilibria.[7] Thus, the set of Nash equilibria may be partitioned into three groups: (a) strict Nash equilibria, N_S, (b) weak Nash equilibria where all unilateral non–disimproving deviations lead to weak Nash equilibria, N_{WA}, and (c) weak Nash equilibria where there are unilateral non–disimproving deviations leading to non-equilibrium points, N_{WNA}. Thus, $NE = N_S \cup N_{WA} \cup N_{WNA}$ and the maximal set of absorbing Nash equilibria $= NE_M = N_S \cup N_{WA}$. Correspondingly, in the state space S, the associates states are denoted: S_E, S_S, S_{WA} and S_{WNA} and S_M.

Definition 7 *The system (π, C) is introspectively stable if for any q there is some equilibrium q^* in an absorbing set, and a sequence $q = q_o, \ldots, q_r = q^*$, such that for each $k = 0, \ldots r-1$, $\exists j$, $q_k^{-j} = q_{k+1}^{-j}$ and $\pi_j(q_k^j, q_k^{-j}) \leq \pi_j(q_{k+1}^j, q_k^{-j})$.*

In words, from *any* action vector there is a sequence of unilaterally improving variations that terminate at some absorbing equilibrium.

Remark 4 This condition is similar but weaker that the condition of weak acyclicity use in Young (1993), and serves a similar purpose. Without it, there may be cycles of non-Nash choices from which it is impossible to escape by unilateral improving deviations. From the stability perspective, such a loop has the same "attracting" properties as a strict Nash equilibrium. ∎

Theorem 1 *Suppose that the system is introspectively stable. Let ξ_ϵ be an invariant distribution of P_ϵ and let ξ be a limit of ξ_ϵ as $\epsilon \to 0$. Then the support of ξ is a subset of S_M, the set of states corresponding to strict or weak absorbing Nash equilibria.*

Thus, when agents learn from their own past experience, selecting for the current period their most successful choices in the past, the system under mutation converges to Nash behavior: with small rates of mutation the choices made in most periods constitute Nash equilibria *if* the system is introspectively stable. Introspective stability is thus the key condition that ensures convergence to Nash states. Section 6 characterizes a rich class of introspectively stable economic environments.

4 Imitative Dynamics.

In contrast to introspective dynamics, with imitative dynamics agents compare their payoff with those of others. In the present context, they may do so over some previous history of experience. We initially assume that $C^i = C^j$ and $\pi^i(\cdot) = \pi^j(\cdot)$, for all i and j, and discuss later how our analysis extends. Let $\pi_\tau^* = \max_j \pi_\tau^j$, and define

$$\hat{B}(\mathbf{q}(t)) = \{q \in C^i \mid \exists \tau,\ t-l \leq \tau \leq t,\ \exists j,\ q_\tau^j = q,\ \pi_\tau^j = \pi_\tau^*\}$$

to be the set of choices in C^i that for some agent at some time gave the highest relative payoff. By analogy with introspective dynamics:

[7] If Q^* is absorbing, no strict subset $\hat{Q} \subset Q^*$ can be absorbing. The collection of absorbing sets is uniquely defined: $\mathcal{Q}_A = \{Q(\alpha)\}_\alpha$. If an absorbing set contains just one point, $Q_\alpha = \{\bar{q}\}$, then that point, $\bar{q}$, is a strict Nash equilibrium.

Definition 8 *Agent i optimizes imitatively at state $s = \mathbf{q}(t)$, if the choice of action in period $t+1$ is drawn from a distribution $\hat{\gamma}^i(s)$ with support $\hat{B}^i_t(\mathbf{q}(t))$.*

Thus, agent i identifies the highest payoffs across players in the past, and selects from the actions chosen by those players whose payoffs were highest.

Remark 5 As with introspective dynamics, multiple models of plausible imitative behavior are possible when agents observe different payoffs from the same action in different periods.[8] This issue does not arise in standard dynamical models of behavior, because agents' action choices depend only on payoffs in the most recent period, but is present here if an agent recalls more than the most recent outcomes. As long as the model of imitative behavior has agents evaluate an action at a past date on the basis of its relative success at that date, our analysis is robust to the particular specification of imitative behavior. However, one could also contemplate models of behavior in which an agent compares their payoff in period τ from action q with the payoff of another agent taking action $\tilde{q}$ in some different period $\tilde{\tau}$. Such a model of behavior has both features of imitation and introspection, and since these give rise to very different dynamics, it should not be surprising to find that when these behavioral models are melded together, the dynamical system does not settle down in the long run. ■

Imitative optimization determines transition probabilities $\hat{p}_{ss'}$ and an associated transition matrix: $\hat{P}$. Modeling mutations as before: $\tilde{q}_i = q'_i$ with probability $(1-\epsilon) + \epsilon\theta^i_{q'_i}$ and $\tilde{q}_i = q''_i$ with probability $\epsilon\theta^i_{\hat{q}_i}$, $\hat{q}_i \neq q'_i$. Under mutation, the transition matrix is perturbed to a matrix $\hat{P}_\epsilon$. Denote the associated invariant distribution by $\hat{\xi}_\epsilon$.

Next we define a notion of stability for imitative systems, the content of which is to assert that from any point there is a path along which imitative behavior leads to a relative equilibrium (rather than cycle forever.) For $q \in C^i = C^j$, let e_q denote the vector $(q, q, \ldots, q)$, where each agents choice is the same. Given $q \in C = \times C^i$, define $F(\bar{q}) = \{e_{\tilde{q}} \mid \exists j,\ \tilde{q} = \bar{q}^j \text{ and } \pi^j \geq \pi^i,\ \forall i\}$.

Definition 9 *The system (π, C) is imitatively stable if for any $q \in C$, there is a strict relative equilibrium q^* such that $\exists$ a sequence $q = q_o, q_1, \ldots, q_r = q^*$, and for each $k = 0, \ldots, r-1$, $\exists \tilde{q} \in F(q_k)$, with $\tilde{q}$, q_{k+1} differing in one component, and for each k, if $\tilde{q}^i \neq q^i_{k+1}$, there is some $j \neq i$ such that $\pi^i(q_{k+1}) \geq \pi^j(q_{k+1})$.*

That is, at q_0, agents imitate the choice of the best performer, $\tilde{q}$. Then, with every player now making this choice, a deviation by i to q^i_1 raises i's payoff relative to some other agent (and hence *all* others from symmetry), leading other agents to imitate i. And so on.

Identify those states in S_A that correspond to strict relative equilibria: $S_{SRE} = \{s \in S_A \mid s = s_q, q \in C, q \text{ a strict relative equilibrium}\}$. With this terminology:

Theorem 2 *Suppose that the system is imitatively stable, a relative equilibrium exists, and all relative equilibria are strict. Let $\hat{\xi}_\epsilon \to \hat{\xi}$. Then the support of $\hat{\xi}$ is a subset of S_{SRE}, the set of states associated with strict relative equilibria.*

---

[8] One might also suppose that rather than imitating the most successful actions, agents place positive probability on imitating more successful actions. Our results are not sensitive to this modification.

Remark 6 One can modify the analysis to allow for the possibility that agents differ in their payoff functions, feasible action choices or information sets. In such an instance, even though an agent's payoffs depend on the actions of all agents, it only makes sense for an agent i to compare the payoff from his action with those of similarly–situated agents. Then one can define disjoint comparison groups, $\zeta^A, \zeta^B, ...\zeta^M$, where each comparison group contains at least two agents, each agent is in a comparison group, and each agent in a comparison group ζ^J compares his actions and payoffs only with those actions and payoffs of other agents in his comparison group. This formulation permits the possibility that the payoff functions of agents in ζ^J are different from those of agents in other comparison groups. Say that an action vector q and a vector of comparison groups ζ is a *heterogeneous relative equilibrium* if for each comparison group ζ^J and each $i, j \in \zeta^J$,

$$\pi^i(q) = \pi^j(q)$$

and $\forall \tilde{q}^i \in C^i$, there exists a $j \in \zeta^J$ such that

$$\pi^i(\tilde{q}^i, q^{-i}) \leq \pi^j(q^j, \tilde{q}^i, q^{-\{i,j\}}).$$

In this framework, imitative stability leads to convergence to relative equilibrium within each comparative group occurs, although behavior in different groups may differ. Our analysis extends straightforwardly to this environment. ■

5 Intuition.

The central ideas differentiating these two models of comparative dynamics may easily be explained. The intuition is most easily given from the continuous perspective with differential payoff functions. Suppose that a dynamic system is at rest at some point $q = (q^1, q^2, \ldots, q^n)$: from period to period each agent chooses the same action. In this context, consider the impact of a perturbation (mutation) in the choice of action by agent i, of size Δ. Thus, i's action moves to $q^i + \Delta$. The impact on agent i's payoff is $\pi^i(q^{-i}, q^i + \Delta) \approx \pi^i(q) + \frac{\partial \pi^i(q)}{\partial q^i}\Delta$. The impact on some other j's payoff is $\pi^j(q^{-\{i,j\}}, q^i + \Delta, q^j) \approx \pi^j(q) + \frac{\partial \pi^j(q)}{\partial q^i}\Delta$.

With an introspective dynamic, a player compares his payoff following the mutation with his prior payoff: $\pi^i(q^{-i}, q^i + \Delta) - \pi^i(q) \approx \frac{\partial \pi^i(q)}{\partial q^i}\Delta$. So, if at the initial state, $\frac{\partial \pi^i(q)}{\partial q^i} > 0$, the impact of the mutation is to move the state to a new state — the system will not return to the initial state. With an introspective dynamic, the mutation, in effect, allows the agent to calculate the gradient of the profit function in the direction of the mutation — unless this is non–positive, the agent will move to the action determined by the mutation. Hence, with introspective dynamics, a rest point of the system must satisfy $\frac{\partial \pi^i(q)}{\partial q^i} = 0$ (modulo details concerning the grid size for feasible actions). That is, each rest point must be a Nash equilibrium: at a rest point, each agent's action maximizes his payoffs given the actions of the other agents.

With imitative dynamics, an agent compares his or her payoff with those of others before and after the mutation, and agents tend to switch to the new action if and only if it leads to greater *relative* payoffs. Following a mutation, $\pi^i(q^{-i}, q^i+\Delta) - \pi^j(q^{-\{i,j\}}, q^i+\Delta, q^j) \approx [\pi^i(q) - \pi^j(q)] + [\frac{\partial \pi^i(q)}{\partial q^i} - \frac{\partial \pi^j(q)}{\partial q^i}]\Delta$. Mutations aside, for q to be a rest point of the system, it must be that $[\pi^i(q) - \pi^j(q)] = 0$ for each i and

j. From such a point, $\pi^i(q^{-i}, q^i + \Delta) - \pi^j(q^{-\{i,j\}}, q^i + \Delta, q^j) \approx [\frac{\partial \pi^i(q)}{\partial q^i} - \frac{\partial \pi^j(q)}{\partial q^i}]\Delta$. Thus, at a rest point of the system, following a mutation in i's action, for i to not switch to the action reached by mutation, it must be that the action of any agent j is not more attractive (comparing payoffs following the mutation), $[\frac{\partial \pi^i(q)}{\partial q^i} - \frac{\partial \pi^j(q)}{\partial q^i}]\Delta \leq 0$ for all Δ. Thus, with an imitative dynamic, a characterizing feature of a rest point is that $[\frac{\partial \pi^i(q)}{\partial q^i} - \frac{\partial \pi^j(q)}{\partial q^i}] = 0$ (again, modulo details about grid selection). That is, the rest point is *not* a Nash equilibrium: instead, at the rest point, each agent maximizes the *difference* between his payoffs and those of other agents.

To understand why introspective and imitative models of behavior lead to different outcomes consider the homogeneous good oligopoly environment in which each firm is initially producing the competitive equilibrium level of output. Suppose one firm is hit by a mutation and produces marginally less output. After the mutation, the profits of this firm rise (since price is now above marginal cost), but the profits of other firms rise by even more (since they are now producing more than the mutating firm). With an introspective dynamic, the mutating firm compares its profit with past profits and concludes that the experiment was a success. With an imitative dynamic, the mutating firm compares its profit with those of other firms, and since its profit is lower than those of other firms, the mutating firm concludes that the experiment was a failure.

It is worth noting that in large populations, Nash equilibrium behavior always obtains. At a rest point, two types of conditions must be satisfied: a level condition, requiring that payoffs are equal across players and a gradient condition requiring that a variation in one agent's action has the same impact on that player's payoffs as on the payoff of other players. If, for $j \neq i$, $\frac{\partial \pi^j}{\partial q^i} = 0$, the gradient condition reduces to the first order condition for i, so a rest point is a Nash equilibrium. Consequently, in a model where each agent is negligible in the population, imitative dynamics generate predictions of Nash behavior.

In comparative dynamics generally, the role played by mutations admits a broader interpretation than that with best response dynamics. With best response dynamics, agents optimize against the current population distribution so that rest points must be Nash equilibria. When there are many Nash equilibria there are many rest points. In particular, every strict Nash equilibrium is a rest point. The dynamic alone does not guide the selection from within the set of rest points. Rather, the formulation of the mutation serves this crucial role by implicitly providing a measure of the size of the basins of attraction of the rest points of the dynamic system: a basin of attraction is larger if it requires more simultaneous mutations to escape it. This gives the best response dynamics model its cutting power — for example — to select among strict Nash equilibria, or select the risk dominant equilibrium. In contrast, with the comparative dynamic, apart from measuring the relative size of basins of attraction, the impact of mutations is, in essence, to provide the agent with the gradient on the payoff function, comparing payoffs before and after mutation. Here, the number of simultaneous mutations required to move from most absorbing states is one. If the mutation leads to a favorable payoff comparison relative to historical action choices or the choices of others, the mutation becomes the status quo.

6 Strategic Complements and Substitutes.

In this section we discuss an important class of games — where agent i's own action, q_i and the aggregate action of other agents, $Q_{-i} = \sum_{j \neq i} q_j$, are either strategic complements or substitutes. The specific structure of these games allows us to identify conditions under which introspective and imitative stability hold. We assume that payoff functions take the form $\pi_i(q_i, \mathbf{q}_{-i}) = \pi_i(q_i, Q_{-i})$, where π_i is strictly concave in q_i. (Concavity is not necessary for the results — see the discussion following theorem 3 below. Implicitly we assume that the finite grid of action-payoff combinations is derived from or can be embedded in a continuous game so that notions such as concavity are to be understood in terms of the continuous game whose discretization is the discrete game.)

In a model with a continuous action space, strategic complements and substitutes are defined by monotonicity in the reaction functions. Let $r_i(Q_{-i}) = \arg\max \pi_i(q_i, \mathbf{q}_{-i})$. Say that the payoff functions satisfy strategic complements if $Q_{-i} > Q'_{-i}$ implies $r_i(Q_{-i}) > r_i(Q'_{-i})$ and strategic substitutes if $Q_{-i} > Q'_{-i}$ implies $r_i(Q_{-i}) < r_i(Q'_{-i})$. For a discrete action space, the best response, $r_i(Q_{-i})$, need not be single–valued, as adjacent actions, q_i and $q_i + \Delta$, may provide agent i the same payoff (strict concavity of π_i in q_i ensures that only adjacent actions can be best responses). However, if actions are strategic complements then $Q_{-i} \geq Q'_{-i}$ implies $r_i(Q_{-i}) \geq r_i(Q'_{-i})$; and if actions are strategic substitutes then $Q_{-i} \geq Q'_{-i}$ implies $r_i(Q_{-i}) \leq r_i(Q'_{-i})$.

In what follows we give conditions under which introspective and imitative stability are satisfied. We begin with imitative stability.

Proposition 1 *If payoff functions satisfy strategic complements then the system is introspectively stable.*

Thus, if payoff functions satisfy strategic complements, introspective dynamics lead to Nash equilibria. It is straightforward to show that when there are only two agents (so that $Q_{-i} = q_{-i}$), that if actions are strategic substitutes then the system is introspectively stable — with two agents, the actions can be reformulated so they enter as strategic complements. When there are more than two agents our proof that if actions are strategic substitutes then the system is introspectively stable requires another condition:

Definition 10 *Payoff functions satisfy damping if*

$$\pi_i(Q_{-i}, q_i) \leq \pi_i(Q_{-i}, q_i + \Delta)$$

then

$$\pi_i(Q_{-i} + \Delta, q_i) > \pi_i(Q_{-i} + \Delta, q_i - \Delta).$$

In terms of reaction functions, the condition requires that if $r_i(Q_{-i}) \geq q_i + \Delta$, then $r_i(Q_{-i} + \Delta) > q_i$. The damping condition is innocuous. The continuous analog, which implies this, is that $\frac{dr_i(Q_{-i})}{dQ_{-i}} \geq -1$. That is, a one unit increase in the aggregate action of other agents lowers agent i's best response by less than one unit. To understand when it is satisfied, note that in an oligopoly context, if firm i's inverse demand function is given by $p_i(Q_{-i}, q_i) = a - bQ_{-i} - dq_i$, where $a, d > 0$, then for any cost function, the damping condition is satisfied as long as $b \leq d$, i.e. as long as i's demand is more sensitive to his output than the total output of other firms. More generally, if agent i's payoffs can be written in the form,

$\pi_i(Q_{-i}, q_i) = \hat{\pi}^i(aQ_{-i} + bq_i, q_i)$, $0 < a \leq b$ (e.g. homogeneous demand), where $\hat{\pi}$ is strictly concave, then the damping condition on payoffs is satisfied. In the substitutes case, damping is sufficient for introspective stability:

Theorem 3 *If payoff functions satisfy strategic substitutes and damping, then the system is introspectively stable.*

In view of theorem 1, strict Nash equilibria will always be in the introspectively stable set. The difficulty is in showing that for some initial states the system cannot cycle around forever among non–Nash equilibria, or cannot cycle in and out of weak Nash equilibria to non–Nash equilibria. The constructive proof argues that suppose the system did cycle forever. Then somewhere in that cycle there is a minimum aggregate action state. Starting from any such minimum aggregate action state, we show that we can construct a sequence of weakly improving deviations that has the property that each agent's action is either monotonically increasing or declining. Given the finite action space, this implies that the sequence must terminate, and we show that it terminates at a Nash equilibrium. The key to constructing these monotone sequences is first to exploit the incremental improving property of the payoff function in own action so that if $\pi_i(Q_{-i}, q_i) < \pi_i(Q_{-i}, q_i + j\Delta)$, $j > 1$, then $\pi_i(Q_{-i}, q_i) < \pi_i(Q_{-i}, q_i + \Delta)$.[9] Consequently, we can restrict attention to increases in actions of one grid unit. Second, we exploit the fact that at the initial minimum aggregate action state, any reduction in action by agents must lower their payoffs (else violating the assumption that we are at a minimum aggregate action). Then continuing we exploit inductively the damping condition so a one unit increase in the action by one agent, lowers the best response by another agent by at most one unit; and following a one unit reduction in an action by one agent, no other agent wants to reduce his action. Next, we consider imitative stability.

The relative profit function (where i's payoff is compared with that of j) is written $r^{ij}(q^1, \ldots, q^n) = \pi^i(q^1, \ldots, q^n) - \pi^j(q^1, \ldots, q^n)$, and the marginal relative benefit to a variation by agent i is given by: $\frac{\partial r^{ij}(q^1,\ldots,q^n)}{\partial q^i} = \frac{\partial \pi^i(q^1,\ldots,q^n)}{\partial q^i} - \frac{\partial \pi^j(q^1,\ldots,q^n)}{\partial q^i}$. Define $Q^r = \{q^* \mid \frac{\partial r^{ij}(q^*,\ldots,q^*)}{\partial q^i} = 0\}$. For any $q^* \in Q^r$, the profile $(q^*, q^*, \ldots, q^*)$ is a relative equilibrium. Assume that every such q^* is in the grid. If there is a unique relative equilibrium, denote it by q^r. Finally, we assume that $\phi^{ij}(q) = \frac{\partial}{\partial q}\{\frac{\partial r^{ij}(q,\ldots,q)}{\partial q^i}\} < 0$. This concavity condition is satisfied is satisfied in standard economic environments, including all those that we consider in section 8.

Theorem 4 *Suppose that the relative profit function is strictly concave in own action. Then the system is imitatively stable. When ϕ^{ij} is negative for all q, the symmetric relative equilibrium is unique.*

At a symmetric choice level, q, if a small change (say increase) in action by some player raises the player's relative payoff, then others imitate. For a sufficiently small change, at the higher action level, no player's relative payoff can be raised by a reduction in their choice. Thus, for an initial point, there is a sequence of steps increasing (or decreasing) to a relative equilibrium.

[9] The incremental improving condition that our proof employs does not require that payoffs be strictly concave in own action, but also follows given various forms of quasiconcavity.

7 Long Run Welfare.

When actions are payoff substitutes (higher action levels for players other that i reduces i's payoff), actions are greater and welfare lower under imitative dynamics; and when actions are payoff complements, the converse is true.

To see this, note that at a Nash equilibrium, $\frac{\partial \pi^i}{\partial q^i} = 0$, whereas in a relative equilibrium, $\frac{\partial r^{ij}}{\partial q^i} = \frac{\partial \pi^i}{\partial q^i} - \frac{\partial \pi^j}{\partial q^i} = 0$. When actions are payoff substitutes $\frac{\partial \pi^j}{\partial q^i} < 0$, so that $\frac{\partial r^{ij}}{\partial q^i} > 0$. Thus, i will raise his action and others will follow via imitation. Under imitation dynamics with mutation, action choices are higher than in Nash equilibrium. With the payoff substitutes assumption, payoffs are lower. The converse holds when actions are payoff complements.

With introspective dynamics, agents take into account how their actions affect payoffs of all agents commonly; but with imitative dynamics, agents do not. For example, in a homogeneous good oligopoly game, introspecting agents take into account how their output affects price; but with an imitative dynamic, agents do not consider the effects of their output on price. Hence, if actions are payoff and strategic substitutes, imitative dynamics lead agents to select more aggressive actions than those chosen by introspecting agents. If, instead, actions are payoff and strategic complements, imitative dynamics lead agents to select less aggressive actions than those chosen by introspecting agents.

This is an ironic outcome. Underlying the imitative dynamic is an informational structure in which each agent observes the actions and payoffs of other agents and mimics the actions of agents with higher payoffs, presumably banking on the thought that imitating best current practice would increase payoffs. In the long run, they are confounded; introspective dynamics lead to more profitable outcomes. An informational environment in which agents cannot observe what everyone does, but do track how their own experimentation affect profits, leads them to a better outcome.

8 Examples.

In this section we discuss a number of key economic environments for which our analysis is relevant. In particular, with appropriate slope or shape restrictions, for the models discussed subsequently convergence under imitative or introspective dynamics occurs. This permits us to characterize better the properties of the different dynamics.

Homogeneous good oligopoly. Let firms face inverse demand $P(\cdot)$, $P(0) > 0, P' < 0$, and have a common cost function, $c(\cdot)$, that satisfies $c(0) = 0, c' > 0, c'' \geq 0$. Suppose that in the continuous formulation of the game there is a unique symmetric Nash equilibrium in which each agent plays q^* that is contained in the set of feasible actions, $C = \{a, a + \Delta, \ldots, k\Delta\}$. If the inverse demand function is strictly convex, then the introspectively stable set is the singleton set, $\{(q^*, \ldots q^*)\}$. So, too, this is the outcome if the inverse demand function is linear and the cost function $c(\cdot)$ is strictly convex.

If inverse demand and cost functions are both linear, then discretization of the action grid introduces additional weak Nash equilibria, in which j agents play $q^* + \Delta$, j agents play $q^* - \Delta$ and the remaining $n - 2j \geq 0$ agents play q^*. This is because $\pi((n-1)q^* + \Delta, q^* - \Delta) = \pi((n-1)q^* + \Delta, q^*)$, and $\pi((n-1)q^* - \Delta, q^* + \Delta) = \pi((n-1)q^* - \Delta, q^*)$. However, these weak Nash equilibria are not in the introspectively stable set, since there exists a sequence of one–step non–disimproving deviations from any of these weak Nash in the direction of the strict Nash equilibrium action. For example, have one of the j agents taking action $q^* - \Delta$ raise his action to q^*. Then since at the new action vector,

$\pi((n-1)q^*+\Delta, q^*) < \pi((n-1)q^*, q^*)$, a deviation by one of the j agents taking action $q^*+\Delta$ to q^* will raise his payoff. Hence, again, the introspectively stable set is again the singleton set, $\{(q^*, \ldots q^*)\}$. Finally, if the demand function is strictly concave, and the cost function linear, then discretization of the action grid introduces additional strict Nash equilibria, in which j agents play $q^*+\Delta$, j agents play $q^*-\Delta$ and the remaining $n-2j \geq 0$ agents play q^*. This is because $\pi((n-1)q^*+\Delta, q_*-\Delta) > \pi((n-1)q^*+\Delta, q_*)$, and $\pi((n-1)q^*-\Delta, q^*+\Delta) > \pi((n-1)q^*-\Delta, q^*)$. In this case, the introspectively stable set consists of all such action vectors. Of course, as $\Delta \to 0$, all elements of the introspectively stable set approach the unique continuous Nash equilibrium outcome.

Heterogeneous good, n–firm oligopoly. Suppose that each firm i faces demand

$$P(Q_{-i}, q_i) = a - bQ_{-i} - dq_i, \quad a, d > 0; \frac{-d}{n-1} < b \leq d,$$

and there are no costs of production. If $b > 0$, then the goods are payoff and strategic substitutes, and if $b < 0$, the goods are complements. Then, as $\Delta \to 0$, introspective dynamics lead to

$$q^i = \frac{a}{(n-1)b + 2d}.$$

In the case of perfect substitutes,

$$q^i = \frac{a}{(n+1)b}.$$

More generally, the smaller is b, the more an agent produces; and if d is positive, the output of the other firms causes each firm to raise production to exploit the complementarities. Imitative dynamics lead to

$$q^r = \frac{a}{(n-2)b + 2d}.$$

In the case of perfect substitutes,

$$q^r = \frac{a}{nb},$$

yielding the Walrasian, competitive, zero profit, outcome, as Vega–Redondo (1996) finds. More generally, if the goods are substitutes ($b > 0$), then firms over–produce with imitative dynamics, $q^r > q^i$, earning correspondingly lower profits; and if the goods are complements ($b < 0$), then firms under–produce with imitative dynamics, $q^r < q^i$, earning correspondingly lower profits as they fail to exploit complementarities as fully.

R&D spillovers. Consider a two stage game in which n firms first make capital investments and then hire labor to produce output in a competitive economy. There are R&D spillovers so that the capital investments of other firms reduce firm i's costs. Firm i's profits are given by

$$\pi(K_{-i}, k_i) = max_{l_i} K^\alpha k_i^\beta l_i^\gamma - rk_i - wl_i,$$

where $K = \sum_j k_j$, $r > 0$ is the unit cost of capital, $w > 0$ is the prevailing wage rate, $\alpha + \beta + \gamma < 1$, and $\alpha, \beta, \gamma > 0$. The same formulation permits an analysis of pollution externalities, whereby the total capital employed pollutes the river, reducing each firm's output, $\alpha < 0, \alpha + \beta > 0$.

With introspective dynamics, as $\Delta \to 0$, equilibrium is characterized by

$$\beta(nk_i)^\alpha k_i^{\beta-1} l_i^\gamma + \alpha(nk_i)^{\alpha-1} k_i^\beta l_i^\gamma = r; \quad \gamma(nk_i)^\alpha k_i^\beta l_i^{\gamma-1} = w,$$

while with imitative dynamics agents do not consider how their actions affect the aggregate, so equilibrium is characterized by

$$\beta(nk_i)^\alpha k_i^{\beta-1} l_i^\gamma = r; \quad \gamma[nk_i]^\alpha k_i^\beta l_i^{\gamma-1} = w.$$

Tragedy of the commons. Consider the choice by each family in a village of how many sheep should graze on the common village pasture. Let $V(G)$ be the value derived by each sheep from grazing when G sheep graze: $V(0) > 0; V' < 0; V'' < 0$, and $G = \sum_j g_j$. The cost to a family i of raising g_i sheep is given by $c(g_i)$. Then the profits of each family as a function of the number of sheep that graze is given by $\pi(G_{-i}, g_i) = g_i V(G_{-i} + g_i) - c(g_i)$.

We could alternatively pose the tragedy of the commons as an over–fishing story, where the number of fish caught by other fisherman, F_{-i}, raise the cost to each fisherman i: $c(F_{-i}, F_i)$. Fish can be sold for price p so fisherman i's profits are given by $\pi(F_{-i}, F_i) = pF_i - c(F_i, F_{-i})$. These payoff functions are structurally identical to those of firms in the homogeneous good oligopoly game, so the same predictions obtain.

Private provision of public good. Each agent i's utility is a strictly increasing function of both the aggregate level of public good, x, and his consumption of a private good y: $U(\sum_j x_j, y_i)$. Each individual has I dollars that he can devote to purchases of the two goods. Then, as $\Delta \to 0$ with introspective dynamics, the standard public goods Nash equilibrium obtains in which agents equate personal marginal rates of substitution between the public and private good to price ratios; while with imitative dynamics, each agent devotes his entire income toward purchases of the private good.

Team production. Consider a team of n workers, whose total output, $Q(\sum_i e_i)$, is an increasing function of total worker effort. Each worker's wage is an increasing function of the team's output. As worker i exerts more effort, the disutility he incurs, $v(e_i)$, rises. Thus, worker i's utility, $U(e_i, \sum_{j \neq i} e_j) = u(w(Q(\sum_j e_j))) - v(e_i)$, takes a form similar to that with the private provision of a public good.

9 Concluding Comments

This paper investigates evolutionary dynamics when agents make choices on the basis of relative performance criteria. We identify two distinct classes of learning behavior: *imitative* dynamics and *introspective* dynamics. With imitative dynamics agents behave as if other agents' experience is relevant for them, tending to imitate the actions of more successful agents. In contrast, with introspective dynamics, agents behave as if their own past experience is relevant for them, tending to choose past successful actions. We find that these different models of learning behavior predict fundamentally different outcomes: Introspective dynamics lead to Nash equilibria, whereas imitative dynamics lead to an outcome in which no agent can increase the *difference* between his payoffs and those of other agents. Paradoxically, comparing outcomes across dynamics, agent payoffs are lower for imitative than introspective dynamics — mimicking best practice turns out to be counterproductive.

Appendix

A *graph* (S, E) is a (nonempty) collection of nodes or vertices S and a (un)ordered collection of edges, E, where an *edge*, e, is a set containing two elements of S; thus $e = \{s, s'\} = \{s', s\}$. When the edges are ordered, called *arcs*, the graph is called a *digraph* (directed graph) — in this case write an arc $e = (s, s')$. Call s the *initial node* of e, denoted $i(e)$, and s' the *final node*, denoted $f(e)$. A *path* (with no loops) from s' to s'', denoted $\mathcal{P}_{s's''}$, is a collection of arcs $\{(s_i, s_{i+1})\}_{i=1}^{k}$, such that $s' = s_1$ and $s'' = s_{k+1}$, and the initial nodes of all arcs are distinct. Given a state space S, and transition matrix M on S, define the set of arcs as $E_M = \{(s, s') \mid m_{ss'} > 0\}$.

Definition 11 *Given $s \in S$, an s-graph is a graph (S, E) such that*

1. *$s' \neq s$ implies $\exists$ a unique $\tilde{s}$, $(s', \tilde{s}) \in E$.*
2. *$\nexists s^*$, $s* \neq s$ and $(s, s^*) \in E$.*
3. *If $s' \neq s$, there is a path from s' to s.*

Let $\mathcal{G}_s$ be the set of all s-graphs Given a graph G, let $E(G)$ be the set of arcs in G. Each arc is identified as a pair of states, $e = (s', s'')$, with associated transition probability $m_e = m_{s's''}$. Let $\beta_s = \sum_{G \in \mathcal{G}_s} \prod_{e \in E(G)} m_e$ and $\beta = \sum_{s \in S} \beta_s$.

Lemma 1 *(Markov chain tree theorem.) If M is irreducible there is a unique invariant distribution $\{\pi\}_{s \in S}$, where $\pi_s = \frac{\beta_s}{\beta}$.*

In the context of mutations, identify the matrix M as P_ϵ. Note that because every $s_{\mathbf{q}} \in S$ is a minimal absorbing state, there are at least $\#C$ minimal absorbing states. Because every s-graph has all but one element of S as the initial point of some arc, for $G \in \mathcal{G}_s$, in the expression $\prod_{e \in E(G)} m_e$ there are at least $\#C - 1$ terms of order no larger than ϵ. Thus, for any s, $\beta_s \leq 0(\epsilon^{\#C-1})$. Since β_s is the sum of terms of varying orders of magnitude, the order of magnitude of β_s is determined by the term of largest order in the sum. If this order of magnitude is smaller than a dominant term in some $\beta_{s'}$, then as mutations go to zero, π_s will converge to 0.

Theorem 1 *Suppose that the system is introspectively stable. Let ξ_ϵ be an invariant distribution of P_ϵ. Let ξ be a limit of ξ_ϵ as $\epsilon \to 0$. The support of ξ is a subset of S_M, the set of states corresponding to strict or weak absorbing Nash equilibria.*

Proof: Call a collection of states $A \subseteq S$ an *absorbing set* of states if there is zero probability of leaving that collection: if $s \in A$ and $s' \in S \setminus A$, then $p_{ss'} = 0$. Call such a collection a *minimal absorbing set* if there is no strict subset that is absorbing. Thus, each singleton set $\{s_q\}$, $q \in A$ is a minimal absorbing set. For each $s \in S$, identify those states from which s can be reached with positive probability under P:

$$H^1(s) = \{s' \in S \mid p_{s's} > 0\}$$

and define $H^j(s)$ inductively for $j \geq 2$:

$$H^j(s) = H^{j-1}(s) \cup \{s' \mid p_{s's''} > 0, \text{ for some } s'' \in H^{j-1}(s)\}$$

Since $\{H^j(s)\}$ is an increasing collection of sets and S is finite, $H^j(s)$ is constant after a finite number number of steps and equal to a set $H^*(s)$. Then $\tilde{s} \in H^*(s)$ if $\exists s^1, \ldots s^k$, such that $\tilde{s} = s^1$, $s^k = s$ and

$p_{s^j s^{j+1}} > 0$, $j = 1, \ldots, k-1$. Let $H_A = \cup_{s \in S_A} H^*(s)$, the set of states under introspective dynamics with positive probability of ending up at some state in S_A.

Partition S according to $S = H_A \cup H_o$ where $H_o = S \setminus H_A$. By construction, H_o is absorbing. In fact, $H_o = \emptyset$, and all minimal absorbing sets are singletons. To see this,consider any $\mathbf{q}(t)$. Starting from this state, let $R(\mathbf{q}(t))$ be the set of states reachable with positive probability from $\mathbf{q}(t)$. There is some $\mathbf{q}' \in R(\mathbf{q}(t))$ such that the payoff to some player j at some vector in the state $\mathbf{q}'$ is the highest payoff at any action vector in any state in $R(\mathbf{q}(t))$. Under the dynamics, when this state is reached (and it has positive probability of being reached) player j has positive probability of then making that choice $l+1$ times (and more) in succession, and so with positive probability, from $\mathbf{q}(t)$ a state where j's history is constant has positive probability . Starting from that state, the argument may be repeated for other agents: so there is positive probability of reaching a state where each agents history is a constant vector.

It is sufficient to show that for any $s \notin S_M$, $\exists s^* \in S_M$ and $\epsilon \cdot O(\beta_{s^*}) \geq O(\beta_s)$, where $O(x)$ denotes the order of magnitude of x and the inequality "$O(x) \geq O(y)$" means that x is at least as large in order of magnitude as y. Furthermore, since the support of any limiting invariant distribution will lie in the set $H_o \cup S_A$, since $H_o = \emptyset$ it is sufficient to consider only states in S_A.

Let $G \in \mathcal{G}_s$ and let E_s be the set of edges in G. Suppose that $s \in S_A$, but $s \notin S_M$. If $s \in S_A$, then from introspective stability, there is some $s^* \in S_N$ and a path $\mathcal{P}_{ss^*} = \{(s_j, s_{j+1}\}_{j=1}^{k-1}$, from s to s^* ($s = s_1$, $s^* = s_k$), whose edges consist of nodes distance 1 apart. Let $s = (q, q, \ldots, q) = s_q$. From introspective stability there is some i and $\tilde{q}^i$ such that $\pi^i(q^{-i}, q^i) \leq \pi^i(q^{-i}, \tilde{q}^i)$. Let $\tilde{q} = (q^{-i}, \tilde{q}^i)$, so $d(q, \tilde{q}) = 1$, where $d(q, \tilde{q}) = \#\{i \mid q^i \neq \tilde{q}^i\}$.

Thus, one mutation moves the state to $s' = (q, q, \ldots, q, \tilde{q})$ and with positive probability after l additional periods the state $\tilde{s} = (\tilde{q}, \tilde{q}, \ldots, \tilde{q}) = s_{\tilde{q}}$ is reached. If $\tilde{q}$ is not an equilibrium, then again from introspective stability, there is some j and $\hat{q}^j$ such that $\pi^j(\tilde{q}^{-j}, \tilde{q}^j) \leq \pi^j(\tilde{q}^{-j}, \hat{q}^j)$. Let $\hat{q} = (\tilde{q}^{-j}, \hat{q}^j)$, so one mutation moves the system to $s'' = (\tilde{q}, \ldots, \tilde{q}, \hat{q})$ and with positive probability the state $\hat{s} = (\hat{q}, \ldots, \hat{q}) = \hat{q}$ is reached after l additional periods. Using the path provided from the introspective stability condition gives a path of zero and one mutations that take the system from $s = s_q$, to $s^* = s_{q^*}$. ¿From G form an s^*-graph, G', as follows. Consider first the case where $q^* \in N_S$, a strict Nash equilibrium.

Retain an arc $e = (s', s'')$ in G, if (a) $s' \neq s^*$ and $\nexists \tilde{s} \in S$, $(s', \tilde{s}) \in \mathcal{P}_{ss^*}$, or (b) $e = (s', s'') \in \mathcal{P}_{ss^*}$. Break the remaining arcs, R, in G. Thus, arc $e = (s', s'')$ is broken if $s' = s^*$ or if s' is the initial point of some edge in $\mathcal{P}_{ss^*}$, but $e \notin \mathcal{P}_{ss^*}$. If $e \in R$ then there is a unique $e' \in \mathcal{P}_{ss^*}$, $i(e) = i(e')$, but $e \notin \mathcal{P}_{ss^*}$. Break each such arc e and replace it with the arc e'. Finally, add the arc $e = (s, s_2) = (s_1, s_2)$. Call the new graph G'.

Under this construction, the new graph G' is an s^*-graph, where node s the initial node of an edge in G' and there is no edge in G' with initial node s^*, whereas in G the converse holds. Edges outside R are unchanged, and for an initial node $\hat{s}$, of an edge $e \in R$, if s' is the successor in G, and s'' the successor in G', then $O(\tilde{p}_{\hat{s}s'}) \geq O(\tilde{p}_{\hat{s}s''})$. Finally, since the edge in G' with initial node s requires one mutation to move to its successor whereas the node s^* in G required at least 2 mutations to move to its successor, say $\bar{s}$, $O(\tilde{p}_{ss_1}) \geq \epsilon \cdot O(\tilde{p}_{s^*\bar{s}})$. Thus,

$$O\Big(\prod_{e \in E_s} m_e\Big) = \epsilon \cdot O\Big(\prod_{e \in E_{s^*}} m_e\Big)$$

Next consider the case where $q^* \in N_{WA}$. Define $s(q, \tilde{q}, r) = (q, \ldots, q, \tilde{q}, \ldots, \tilde{q})$ — q occurs in the first

$l+1-r$ positions and $\tilde{q}$ occurs in the last the r positions. From the introspective stability condition, let $\tilde{q}$ differ from q in one coordinate, weakly improving for that agent. Adjoin s_q to the state $\tilde{s} = (q, \ldots, q, \tilde{q})$. From $\tilde{s}$ the successor state in G is $\tilde{s} = (q, \ldots, q, \tilde{q}, \tilde{q}) = s(q, \tilde{q}, 2)$ or some other state. In the latter case, an additional mutation has occurred. If so, break the arc there, and add the arc $((q, \ldots, q, \tilde{q}), (q, \ldots, q, \tilde{q}, \tilde{q}))$. This new arc has positive probability in the unperturbed system. If in G, the successor to $s(q, \tilde{q}, 2)$ is $s(q, \tilde{q}, 3)$ retain that arc. If the successor to $s(q, \tilde{q}, 2)$ is not $s(q, \tilde{q}, 3)$ (again a mutation has occurred), break the arc and add a new arc $(s(q, \tilde{q}, 2), s(q, \tilde{q}, 3))$. Proceed, breaking arcs where a successor does not extend the sequence: if the successor to $s(q, \tilde{q}, r)$ is $s' \neq s(q, \tilde{q}, r+1)$, break that arc and add the arc $(s(q, \tilde{q}, r), s(q, \tilde{q}, r+1))$. This process creates a path from s_q to $s_{\tilde{q}}$. Each arc in G is replaced with one requiring the same or fewer mutations.

Proceed in this manner, constructing a path to s_{q^*}. From s_q^*, in G the successor either arises from just one weakly improving deviation, or just one strictly dis-improving deviation, or there is more than one choice changed in the successor state. In the second case, break the arc, and reverse it — replacing a one-mutation deviation with a zero cost arc in the unperturbed system, and form a tree with root s_{q^*}. In the third case, break the arc (which has an associated cost of two mutations) and again form a tree with root s_{q^*}. In either of these two cases, the new tree has weight at least an ϵ order of magnitude larger than G. In the first case, there is a weakly improving deviation from s^* to some other Nash equilibrium state $s_{\hat{q}}$. Break arcs in G where necessary to form a (zero cost) path from s_{q^*} to $s_{\hat{q}}$, apart from the arc $((q^*, q^*, \ldots, q^*), (q^*, \ldots, q^*, \hat{q}))$. As in the previous discussion, each arc in G is replaced with one requiring the same or fewer mutations. At $s_{\hat{q}}$, the possibilities are the same as at s_q^*: one or two mutations are required to leave $s_{\hat{q}}$. In the case where two are required, form the tree with root $s_{\hat{q}}$, or else connect $s_{\hat{q}}$ to its successor determined by introspective stability. This can only happen a finite number of times — eventually, a Nash state is reached where two mutations are required to leave. Break that arc and form the tree rooted at that state. In all possible cases, the "weight" of the new graph, G', is an ϵ order of magnitude larger that of G.

∎

Theorem 2 *Suppose that the system is imitatively stable and all relative equilibria are strict. Let $\hat{\xi}_\epsilon \to \hat{\xi}$. Then $\hat{\xi} \in S_{SRE}$, the set of states associated with strict relative equilibria.*

Proof: In this case, all absorbing sets are in S_A. In fact, all absorbing sets are in the subset of S_A where the choices of all agents are the same. To see this not that for any state s, with imitation dynamics, there is positive probability that every agent will make the same choice next period (since they all observe the same "best performers" in previous periods). In this case, each player receives the same payoff which may be higher or lower than that choice yielded earlier (when not all were choosing it). In either case, the choice is still top ranked in relative payoff terms, and so will be chosen again with positive probability by each agent. This continues — so there is positive probability of reaching a state where each agent has made the same choice in the previous $l+1$ periods.

Consider a action profile q that is not a strict relative equilibrium and let G be a tree rooted at s_q. Let $\{q_0, \ldots, q_r\}$ be the sequence specified by imitative stability, with $q = q_0$ and $q_r = q^*$. Locate s_{q^*} in G and break the arc from s_{q^*}. Since q^* is a strict relative equilibrium, two mutations are required to leave state s_{q^*}. And along the sequence specified by imitative stability, one mutation initiates a move from q^r to q^{r+1} along a path requiring no mutations. Every arc with an initial point s^k, $k = 1, \ldots, r$ requires at least one mutation. Create a new tree, G', rooted at s_{q^*} If an arc is broken in G an arc is added in the construction of G' to create a path from s_{q^r} to $s_{q^{r+1}}$, and the arc added in G' involves no more mutations. At s_{q^*}, break the arc in G with s_{q^*} as the initial point and let s_{q^*} be the root of the new tree, G'. In G, at least two mutations are required to leave s_{q^*}. So, the arcs in tree G' involve at least one fewer mutations than in G. The produce of terms in the G' tree are an ϵ order of magnitude smaller than the product of terms in G. Thus, if q is not a relative strict equilibrium, s_q has 0 probability in the limiting distribution.

∎

Proposition 1 *If the payoff functions satisfy strategic complements then the system is introspectively stable.*

Proof: Suppose that the system is not introspectively stable so that there is some $\mathbf{q}$ such that there is no sequence of one step weak (or strict) best responses leading from q to an equilibrium. In this case, given $q^0 = q$, $\exists q^1$, differing from q in one component, say i and $\pi_i(q^0) < \pi_i(q^1)$, where the i-th component is the best response by i to q^0. Because q^1 is not an equilibrium, $\exists\ q^2$ differing from q^1 in one component, determined by a best response of some player j. Proceed in this manner to obtain a sequence $\{q^k\}$ containing no equilibrium. Because the set of possible q-profiles is finite, the sequence must eventually form a loop. If $\underline{Q}$ is the lowest sum in the loop, then it must be the case that at the profile $\underline{q}$ achieving $\underline{Q}$, the successor to $\underline{q}$ is some $\tilde{q}$ differing from $\underline{q}$ in one component which is larger. In this case, since $\tilde{q}$ is not an equilibrium, some other agent wishes to change their output level. Because of the strategic complements assumption, the best response must be to increase output. Proceed in this way to find a monotone increasing sequence of output choices for agents that make quantity changes. Monotonicity implies that this sequence converges — to some q^*. This must be an equilibrium since no agent wishes to change output, contradicting the assumption that each quantity is part of a cycle completing a loop.

∎

Theorem 3 *If payoff functions satisfy strategic substitutes and damping, then the system is introspectively stable.*

Proof: Suppose not, so there is some q, not an equilibrium and such that every sequence q^r, $q = q^1$, with the property that q^r and q^{r+1} differ in one component, say $q^r_j \neq q^{r+1}_j$ and q^{r+1}_j weakly improving for j relative to q^r_j does not pass through an equilibrium point. Let $\underline{q}$ be the vector with the lowest sum, $\underline{Q}$, reachable by any such sequence. The following algorithm provides a path to an equilibrium via one step non-disimproving deviations, so that introspective stability is satisfied.

Stage 1.

1. (**increase**) Starting at $Q^1 = \underline{Q}$, there is at least one agent for which a unit increase is a strictly improving deviation. Pick one such agent and increase their output by 1, thus raising aggregate output to $Q^1 + 1$.
2. (**decrease**) At $Q^1 + 1$, identify the set of agents that wish to reduce output. If there are any with a strict preference, select one and reduce that agents output by 1. If no agent with a strict preference

exists but some agent has a weak preference to reduce output select one such agent and reduce that agents output by 1. If no agent has a weak or strict preference to reduce output, go to stage 2.

At the end of stage 1, beginning of stage 2:

(a) There is no agent that strictly or weakly prefers to reduce output.

(b) Aggregate output at the beginning of stage 2, Q^2, is either higher (by one unit) or at the initial level: $Q^2 \geq Q^1$. If $Q^2 = Q^1$, some agent lowered output. If $Q^2 = Q^1 + 1$ then no one had even a weak preference to reduce output. In either case, no agent wishes to reduce output.

(c) Finally, if one agent has increased output and one has decreased output, these must be different agents.

Stage 2.

1. (**increase**) Starting at Q^2, if there is at least one agent for which a unit increase is a strictly improving deviation, pick one such agent and increase their output by 1. If no such agent exists, but there is an agent with a weak preference to increase output, raise the output of one such agent. In either of these cases, aggregate output is now $Q^2 + 1$. If neither of these cases obtains (no agent has even a weak preference to increase output), terminate the algorithm. (Output increased in stage 1, unless the initial increase was followed by a reduction. If a reduction occurred, then the aggregate level is Q^1, and no further reductions can occur. If no reduction occurred, and no increase occurs at the beginning of stage 2, no agent wishes to increase or decrease output.) Otherwise, proceed to the next step.
2. (**decrease**) At Q^2+1, identify the set of agents that wish to reduce output. If there are any with a strict preference, select one and reduce that agents output by 1. If no agent with a strict preference exists, but some agent has a weak preference to reduce output, select one such agent and reduce that agents output by 1.

At the end of stage 2, beginning of stage 3:

(a) There is no agent that strictly or weakly prefers to reduce output.

(b) Aggregate output at the beginning of stage 3, Q^3, is either higher (by one unit) or at the initial level of Q^2. If a reduction occurred, the output level is the same, Q^2, as at the beginning of stage 2 (end of stage 1), where no agent wanted to reduce output, or else it is one unit higher and no agent wants to reduce output.

(c) If i increased output in stage 1, i will not reduce output in stage 2. If i increased output in stage 1, $q_i^2 = q_i^1 + 1$, and, depending on whether there was a reduction or not in that stage, Q_{-i}^2 is either Q_{-i}^1 or $Q_{-i}^1 - 1$. In stage 2, if some agent (other that i) raises output, then the aggregate (excluding i) becomes $Q_{-i}^1 + 1$ or Q_{-i}^1. Since at stage 1, agent i raised output, from q_i^1 to $q_i^2 = q_i^1 + 1$, $\pi_i(Q_{-i}^1, q_i^1) < \pi_i(Q_{-i}^1, q_i^1 + 1)$. introspective stability implies $\pi_i(Q_{-i}^1 + 1, q_i^1 + 1) > \pi_i(Q_{-i}^1 + 1, q_i^1)$ (since $\pi_i(Q_{-i}^1, q_i^1) \leq \pi_i(Q_{-i}^1, q_i^1 + 1)$), so in this case i won't reduce output. From the substitutes property, this implies that if q_i^2 is a better response than q_i^1 when $Q_{-i}^2 = Q_{-i}^1$ instead of $Q_{-i}^1 + 1$. So, in both cases, i will not reduce output.

(d) If a reduction by j occurred in period 1, then at that point $\pi_j(Q_{-j}^1 + 1, q_j^1 - 1) \geq \pi_j(Q_{-j}^1 + 1, q_j^1)$, At next stage, j could only weakly prefer to raise output; if no other agent strictly prefers to raise output then the state is an equilibrium since no agent has even a weak preference to lower output.

Suppose that a sequence of output vectors, $q^1, \ldots, q^t$ are generated by this algorithm and have the properties:

(i) $Q^\tau \leq Q^{\tau+1}$, $\tau = 1, \ldots, t-1$, and

(ii) for each i, q_i^τ is monotone increasing or monotone decreasing.

(iii) For each τ, for each i, $\pi_i(Q_{-i}^\tau, q_i^\tau) > \pi_i(Q_{-i}^\tau, q_i^\tau - 1)$; at the end of period τ, no agent weakly prefers to decrease output.

<u>*Proceed to stage t.*</u>

1 If there is no agent that wishes to raise output, either strictly or weakly, then no increase is made, and the process stops. Otherwise, increase the output of one such agent, giving precedence to agents with strict preference. Then go to step 2.

2 If there is some agent that strictly or weakly prefers to reduce their action, then reduce the action of the agent,

If in 1, the selected agent has never reduced output in a previous period and in 2 the selected agent has never increased output in a previous period, then proceed to the next stage.

Otherwise, there are two possibilities: (c1) an increase by someone who previously reduced output or (c2) a reduction by someone who previously reduced output.

Consider the first case (c1): i who previously reduced output at date $\tau < t$ (so τ is the most recent stage at which i moved), has a strict or weakly improving increase. At the beginning of stage τ, total output is $Q_{-i}^\tau + q_i^\tau$. The output increase in that stage by some player $j \neq i$, leads to an output of $(Q_{-i}^\tau + 1) + q_i^\tau$, followed by a drop to $(Q_{-i}^\tau + 1) + (q_i^\tau - 1)$, when i reduces output. Since output is nondecreasing from stage to stage and there is no further change by i until stage t, the output of agents other than i, at the beginning of stage $t-1$ must be at least $Q_{-i}^\tau + 1$. If it is higher, $Q_{-i}^{t-1} > Q_{-i}^\tau + 1$, then since i chose down when output excluding i was lower (at $Q_{-i}^\tau + 1$), it cannot be that i would choose to increase output when facing higher outputs of others (Q_{-i}^{t-1}).

Otherwise, $Q_{-i}^{t-1} = Q_{-i}^\tau + 1$. At $Q_{-i}^\tau + 1$, i weakly prefers $q_i^\tau - 1$ to q_i^τ, so the only possibility for i to now increase output is if i is indifferent between $q_i^\tau - 1$ and q_i^τ at $Q_{-i}^{t-1} = Q_{-i}^\tau + 1$. In this case, there is no agent with a strict preference to increase or a weak preference to reduce output, and so the state is an equilibrium. Thus, the first case leads to an equilibrium.

Next, consider the second case (c2): some agent reduces output and has a history only of output increases prior to t, with the most recent occurring at stage τ. So, at date t, following an increase by some agent k, agent i reduces output by one unit τ, and at stage τ there was an increase by i (matched by a reduction in that stage by some player j). Modify the history up to period $t-1$ by replacing the increase of i at stage τ with an increase for k at that stage. For agents other than i and j, there is no change. For k, the impact of deleting i's increase at τ and replacing it with an increase by k is that at all stages between τ and $t-1$, k faces an aggregate output of other agents that is no larger than Q_{-k}^{t-1}, against which k has a best response to increase. At period t aggregate output excluding i is now one larger than initially (where i wished to reduce output). So, following an increase by some agent r (the algorithm terminates if no one wishes to increase): (d1) i may again wish to reduce output, or (d2) some agent, $j \neq i$, who has increased in the past may wish to reduce output, or (d3) some agent with no past increases may wish to reduce.

In case d1 the algorithm proceeds as before. Because aggregate output of agents other than i is higher at $t-1$ under the new sequence than the old, i will never choose to raise output at this level. If after the modification of i's output, a cycle now occurs for j (d2), the aggregate output facing i at stage $t-1$ is unchanged and one higher than initially. So, after a first reduction of i's output, no subsequent increase ever occurs for i.

Since there are only a finite number of agents and quantities, from stage $t-1$ only a finite number of cycle reductions can occur. The algorithm will iterate at stage t a finite number of times before either converging at this stage to a state where no agent wishes to increase or decrease output (strictly), or else the algorithm will move to stage $t+1$, with all quantities being monotone sequences.

∎

Theorem 4 *Suppose that the relative profit function is strictly concave in own action. Then the system is imitatively stable. When ϕ^{ij} is negative for all q, the symmetric relative equilibrium is unique.*

Proof: Consider a point $q \in C = \times C^i$. For some j, $\pi^j(q) \geq \pi^i(q)$, $\forall i$. Thus, under imitation, there is positive probability that all agents will choose q^j subsequently: under imitation $(q^j, \ldots, q^j)$ is reached with positive probability. So, assume that in q each agent makes the same choice: $q^j = q^*$, $\forall j$.

Let the relative profit between i and j be $r^{ij}(q) = \pi^i(q) - \pi^j(q)$, $q \in C$. Since in q each agent makes the same choice, $r^{ij}(q) = 0$ or $\pi^i(q) = \pi^j(q)$. Suppose that q is not a relative equilibrium. So, $\exists \tilde{q}^i \in C^i$, $r^{ij}(\tilde{q}^i, q^{-i}) = \pi^i(\tilde{q}^i, q^{-i}) - \pi^j(\tilde{q}^i, q^{-i}) > 0$. Let $\hat{q}^i$ maximize $r^{ij}(\cdot, q^{-i})$. and suppose (without loss of generality) that $\hat{q}^i > q^*$, where q^* is the common choice of other agents. Since r^{ij} is strictly concave in q^i, $r^{ij}(q^i + \Delta, q^{-i}) > 0$. Subtracting $\pi^i(q) - \pi^j(q)$ gives $0 < \pi^i(q^i + \Delta, q^{-i}) - \pi^j(q^i + \Delta, q^{-i}) - \pi^i(q) - \pi^j(q)$ or $0 < [\pi^i(q^i + \Delta, q^{-i}) - \pi^i(q)] - [\pi^j(q^i + \Delta, q^{-i}) - \pi^j(q)]$. Dividing by Δ gives, approximately,

$$\frac{\partial \pi^i}{\partial q^i} - \frac{\partial \pi^j}{\partial q^i} > 0$$

Let $q + \underline{\Delta}$ be the vector q incremented by Δ in each coordinate. Since at this choice vector, everyone makes the same choice, $r^{ij}(q + \underline{\Delta}) = 0$. At $q + \underline{\Delta}$, consider a downward mutation — from $q^i + \Delta$ to q^i.

$$\begin{aligned} r^{ij}(q^i, q^{-i} + \underline{\Delta}) &= r^{ij}(q^i, q^{-i} + \underline{\Delta}) - r^{ij}(q + \underline{\Delta}) \\ &= [\pi^i(q^i, q^{-i} + \underline{\Delta}) - \pi^i(q + \underline{\Delta})] - [\pi^j(q^i, q^{-i} + \underline{\Delta}) - \pi^j(q + \underline{\Delta})] \end{aligned}$$

Dividing by Δ, this is approximately

$$-[\frac{\partial \pi^i}{\partial q^i} - \frac{\partial \pi^j}{\partial q^i}]$$

Thus, if initially some agent had an incentive to make a higher choice, at the new level when others have adjusted, no agent wishes to decrease their choice. Repeating this calculation leads to a path from q up to the point where an agents relative payoff is not increased by either raising or lowering their choice — a relative equilibrium.

Next, if the relative profit function is strictly concave in own action (r^{ij} strictly concave in q^i), and $\phi^{ij}(q) = \frac{\partial}{\partial q}\{\frac{\partial r^{ij}(q,\ldots,q)}{\partial q^i}\} < 0$, then there is a unique symmetric relative equilibrium. To see this, note that since $\phi^{ij}(q) = \frac{\partial}{\partial q}\{\frac{\partial r^{ij}(q,\ldots,q)}{\partial q^i}\} < 0$, for low symmetric action levels, $q < q^r$, each agent's relative payoff is increasing in own action, but as the symmetric level increases, the marginal relative benefit to increasing own action declines to zero at q^r and then becomes negative. For example, in the case of oligopoly, $r^{ij}(q^1,\ldots,q^n) = P(\sum q^k)q^i - c(q^i) - [P(\sum q^k)q^j - c(q^j)]$. Thus, $\frac{\partial r^{ij}(q^1,\ldots,q^n)}{\partial q^i} = P(\sum q^k) + P'(\sum q^k)q^i - c'(q^i) - [P'(\sum q^k)q^j]$ and $\frac{\partial r^{ij}(q,\ldots,q)}{\partial q^i} = P(nq) + P'(nq)q - c'(q) - [P'(nq)q] = P(nq) - c'(q)$. The derivative of this with respect to q is $nP'(q) - c''(q) < 0$. (In the linear case, $r^{ij}(q^1,\ldots,q^n) = ([(a-c) - b[q^i + Q^{-i}])(q^i - q^j)$, so that $\frac{\partial r^{ij}(q,\ldots,q)}{\partial q^i} = (a-c) - 2bq^i - bQ^{-i} + bq^j = (a-c) - bQ + b(q^j - q^j)$. At the symmetric level q, this becomes $(a-c) - bnq$, and the derivative with respect to q is $-nb$.)

∎

REFERENCES

[1] Aumann, R, (1961): "Almost Strictly Competitive Games", *Journal of the Society for Industrial and Applied Mathematics*, Vol 9, No. 4, 544-550.

[2] Borgers, T. and R. Sarin (1997): "Learning Through Reinforcement and Replicator Dynamics", *Journal of Economic Theory*, 77, 1-14.

[3] Kandori, M., G. Mailath, and R. Rob (1993): "Learning, Mutations and Long-Run Equilibria in Games", *Econometrica*, 61, 29-56.

[4] Karandikar, R,, Mookherjee, D., Ray., D, and F. Vega-Redondo (1998), "Evolving Aspiration and Cooperation", *Journal of Economic Theory*, 80, 292-331.

[5] Kats, A. and J.-F. Thisse (1992): "Unilaterally Competitive Games", *International Journal of Game Theory*, 21, 291-299.

[6] Schlag, K.H., (1998), "Why Imitate, and If So, How", Journal of Economic Theory, 78, 130-156.

[7] Vega-Redondo, F (1997): "The Evolution of Walrasian Behavior", *Econometrica*, 65, 375-384.

[8] Young, P. (1993): "The Evolution of Conventions", *Econometrica*, 61, 57-84.