Topics in Theoretical Economics

Volume 4, *Issue* 1 2004 *Article* 7

Imitation and Long Run Outcomes

Jayasri Dutta* Kislaya Prasad†

*University of Birmingham, j.dutta@bham.ac.uk

†Florida State University, kprasad@coss.fsu.edu

Imitation and Long Run Outcomes*

Jayasri Dutta and Kislaya Prasad

Abstract

In a number of evolutionary models the presence of mutations, or random components of choice, serve to refine predictions of long-run behavior. We analyze the effects of mutation rates that vary because of the presence of imitation. A full characterization of long-run outcomes is provided for familiar coordination and congestion games, and also a number of other games not previously considered within the evolutionary framework. Our results are often quite distinct from those in the literature. We apply these tools to a series of economic applications, including market games, where imitation can explain excess volatility of prices, and location games, where it leads to greater uniformity in choices.

KEYWORDS: Imitation, Evolutionary Games, Long Run Outcomes

*Jayasri Dutta Department of Economics, University of Birmingham, Edgbaston, Birmingham B15 2TT, U.K. Email: j.dutta@bham.ac.uk Kislaya Prasad Department of Economics, Florida State University, Tallahassee, FL 3206-2180, U.S.A. Email: kprasad@coss.fsu.edu

1 Introduction

The powerful predictions of economics are derived from the premise that rational individuals optimize. Real people often make mistakes—and of different kinds in different circumstances. Does it matter what mistakes they make? In this paper, we analyze the consequences of imitation as one sort of departure from rationality: every now and then, rational individuals choose the same sort of car, or invest in the same stocks, or live in the same area as others, even though rationality dictates otherwise. We evaluate the long run predictions of imitation, and find that these predictions are often very different from what we have learned to expect from random mistakes (as in Foster and Young, 1990, Young, 1993, and Kandori Mailath and Rob, 1993). And as with random mistakes, imitation matters even when it is infrequent. We derive the theory, and provide a full characterization of limit distributions of outcomes. Our methods are applicable to the familiar coordination and congestion games, where payoffs are a monotone function of the number of people who play the same strategy, as well as to non-monotone games. We apply these tools to a series of economic settings that illustrate the potential value of imitation as a positive theory.

The demonstration that small mistakes made by otherwise rational individuals can lead to powerful predictions about outcomes is central to a substantial body of research starting from Kandori, Mailath and Rob (1993) and Young (1993) (KMRY, from now on). Economic settings often have multiple equilibria, and some equilibrium outcomes are unlikely to persist if individuals make mistakes. Bergin and Lipman (1996) show that the long run persistence of one or another outcome is sensitive to the nature of mistakes. Their demonstration relies on "state dependent" errors, and they show that every invariant distribution of the noiseless process can be selected for a suitable specification of error probabilities. One way to proceed, in light of this difficulty, is to seek justification for a particular specification of noise in an economic or psychological theory about the sources of error (e.g. van Damme and Weibull, 2002, where it is costly to avoid mistakes). In our paper, we think of imitation as one sort of mistake, state dependent by construction, and find that it can lead to outcomes distinct from KMRY. In keeping with the literature, we consider the effects when imitation is rare. However, in many applications imitation may be non-negligible. So we deduce invariant distributions where possible, and examine the properties of long-run outcomes in the presence of significant imitation.

In our specification of the adjustment process, the likelihood that a strategy is used depends (i) on whether it is a best response to prevailing play in the population, and (ii) on its popularity. The relationship is stochastic: an unpopular strategy which is not a best response could still be used, but its probability would be small. In effect, we allow players to make two sorts of mistakes, randomization and imitation, and compare limit behavior as the frequency of one vanishes faster than the other. We show, in theory and in examples, when these limit outcomes are distinct. Surprisingly, the resulting dynamic always selects the same equilibria as KMRY in monotone games (these are the games usually studied in the evolutionary literature).

For non-monotone games, the introduction of imitation makes a genuine difference for the selection results. In some cases we get sharp selection results even with a positive probability of imitation, as the frequency of random mistakes vanishes, which makes it possible for us to assess the effects of adding "a little" imitation to a model with optimizing agents. Some equilibrium outcomes are simply ruled out by imitation, and these outcomes can be more efficient.[1]

Beyond questions of robustness, our analysis suggests that imitation may also be useful as a positive theory. Among other things, it can imply uniformity in behavior, but may also lead to non-degenerate limit distributions (even in games with a unique equilibrium). In one application of our ideas, to strategic market games, we show that small amounts of imitation can explain excess volatility of prices and trading volume. In another application, to a model of location with pricing, we show that imitation leads to everyone choosing to settle in the same location (when, in the absence of imitation, they would have chosen to spread out across locations). The economic applications are used not just to illustrate the theory we develop, but also to extend the range of applications of evolutionary models (ours, as well as those of others).

Our hypothesis is that it is human to imitate. Indeed, much recent research in neuro-science (Chaminade et. al., 2002) as well as animal behavior (Byrne & Russon, 1999) suggests that imitation is of greater consequence than previously believed. The propensity to imitate is crucial to social interaction, and experimental evidence suggests that human brains come equipped with the necessary tools. In some dimensions, such as language, it is virtually necessary to make the same choice as others. In others, imitation may be individually rational — an issue emphasized in the theory of "rational herding" (e.g. Bikhchandani, Hirshleifer and Welch, 1992). It may be an efficient rule of thumb in learning about uncertain payoffs (e.g. Schlag, 1998). Eshel, Samuelson and Shaked (1998) show, in a local interaction model, how imitating neighbors who earn high payoffs can lead to the evolution of altruism. Imitation is often useful – but this may build in a habit, and lead to imitative choices even when they are counterproductive. It does seem, after all, that we observe much greater conformity than would be warranted by its usefulness, in the choice of clothes or cars or first names of children.[2]

Much of the research on imitation (e.g. Apesteguia *et. al.*, 2003, Schlag, 1999, Vega–Redondo, 1997, in addition to papers cited above) assumes that other people's payoffs are observable, and their success enters into the decision to imitate. Experimental evidence from such games supports the claim that people imitate success. But in some contexts payoffs are unobservable, and imitation may be tied only to the popularity of a strategy. In this vein, Kirman (1993) develops a theory based entirely on random imitation. In his leading example, ants follow each other to one of two sources of food, and applications include noise trading, epidemics, and the diffusion

[1]For nonstrategic environments, in contrast, we would expect imitation of others to lead to efficiency so long as the probability that optimal strategies are used is large enough.

[2]Of course, people are sometimes driven by a desire for novelty as well. It would be interesting to examine the implication of such preferences, but this is beyond the scope of the present work.

of opinion. In this paper, we model imitation as unconditional—it does not depend upon the payoff of the people who are imitated. There is some experimental support for this kind of innovation in the context of herding and information cascades (see Celen and Kariv, 2004). If the environment is one where payoffs are observable, one may think of imitation as an occasional mistake made by otherwise rational individuals.

In our simplest model, N identical players must choose one of two actions. Payoffs depend on the action and on the number of others who choose it, and this effect can be non-monotone, allowing coordination and congestion effects. Individuals choose an action and get occasional opportunities to revise their decision. When they do, they choose best responses with high probability. We allow for two sorts of mistakes – random errors as in KMRY, and imitating a random draw from the rest of the population, as in Kirman (1993). This describes a stochastic process for the population profile of choices, and we identify the invariant distribution. Long-run, or stochastically stable, outcomes constitute the support of the limit of the invariant distribution as error probabilities go to zero. The long-run consequences of imitation can be described by studying limits as the probability of random errors vanishes at a rate different from that of imitation. The invariant distribution of the noiseless process which is selected as error probabilities go to zero depends on the rate at which the two kinds of errors vanish. When the two kinds of errors are present in very small amounts, their relative magnitude is significant in determining what the long run outcome is going to be.

The paper is organized as follows. In section 2 we describe the overall framework. We adopt a birth–death formalism. As we show, this matters little for the main theorem, and is convenient for deducing the limiting results. Section 3 contains the key results on long run outcomes and we provide some results on limit distributions when N becomes large. In section 4 we present four economic applications, and the implications of our results for the examples are worked out. Finally, in section 5, we consider a number of extensions of our results. We show that results remain valid for a wide class of imitation dynamics, and also when knowledge of the distribution of actions is incomplete. We extend some of our results to local interaction games where we show that increasing N has the same effect as decreasing the imitation probability.

2 Framework

We develop a model that allows us to examine long run outcomes of a process of social interaction in which the likelihood of a particular action being used depends on its optimality as well as its popularity. The specific context is one of N-player *symmetric* games with two actions (for each $i = 1, \ldots, N$ we have $a_i \in \{0, 1\}$). We let n_1 denote the number of people who choose action 1, and $n_0 = N - n_1$ the number who choose action 0. The payoff, or utility, of a player who chooses action j will be denoted by $V_j(n_1)$. A special case of this model is the population game where a finite number of players are matched at random to play a symmetric finite game. V_j then denotes the expected payoff from action j given the distribution of choices made by other players. But we are also interested in a number of other examples from economics in which payoffs depend on the distribution of actions in the population. Four such examples are discussed in section 4 below. We consider a market game (a monotone congestion game), a game of technology choice (a monotone coordination game), a location game with prices, and an example of clubs with congestion.

The framework is derived from previous work by Kandori, Mailath and Rob (1993), Young (1993), and Kirman (1993) on evolutionary selection. Our particular formulation is most closely related to Blume (1997). We specify a dynamic model of decision-making where, at each date t ($t = 1, 2, \ldots$), exactly one person gets the opportunity to make a choice from $\{0, 1\}$. This person is chosen at random, with probability $1/N$, and is then locked into this choice until the next time that a decision opportunity arises. The papers of Kandori, Mailath and Rob (1993) and Young (1993) assume the (myopic) *best response* dynamics — at a decision opportunity, a player takes the population distribution of choices as given and chooses a best response. Kirman (1993) specifies an *imitation* dynamic — a second player is drawn from the remaining pool and the decision-maker imitates her choice of action with probability π (as a result popular strategies are more likely to be used). All of these papers assume noise at the level of individual decision-making. In Kandori, Mailath and Rob (1993) and Young (1993) there is some probability of "mistakes," as players do not choose a best response. In Kirman (1993), a player will choose, with small probability ϵ, an action different from their previous choice (regardless of what the second player drawn happens to be doing).

In our paper the choice of strategy depends both on optimality and popularity. For simplicity, we assume the following dynamics. Absent random noise, at a decision opportunity, a player will imitate (some other player drawn at random from the rest of the population) with probability π. But with probability $1 - \pi$ she will optimize (i.e. choose a best response to the prevailing distribution of choices). The noise is modelled as in Kirman's paper — with probability ϵ the decision-maker chooses an action different from her previous choice. While the specification of the dynamics might seem special, the results are robust to generalization. We will show that a wide class of imitation rules lead to essentially the same results that we get with the dynamics just described.

This model of choice describes a stochastic dynamic process. The state variable is the profile of actions. In the binary action case the state is adequately represented by the number $m \in \{0, 1, \ldots, N\}$ of players who choose some specified action ($N - m$ people choose the other action). We adopt the following convention: $V_i(m)$ denotes the payoff to an agent who chooses i when a total of m players choose action 1. The process is a Markov chain and, whenever $\epsilon > 0$, is aperiodic, irreducible, reversible and ergodic. Consequently, there is a unique stationary distribution which will be denoted by $\mu(m; \pi, \epsilon)$. We follow the literature in examining the limit of this distribution as $\epsilon \to 0$ (we let $\mu^*(m; \pi) \equiv \lim_{\epsilon \to 0} \mu(m; \pi, \epsilon)$). States in the support of μ^* are said to be *stochastically stable* and will be identified as *long run outcomes* of the process. The stochastically stable states are a subset of the steady states of the unperturbed process (i.e. the process with $\epsilon = 0$) so that this notion serves as a powerful selection device. For small ϵ, the stochastic process spends most of its time in the set of stochastically stable states.[3]

We find that the nature of long run outcomes depends upon payoff differentials at the "edge" of the state space. Recall that the state space is $\{0, 1, \ldots, N\}$. The *edges* are $m = 0$ and $m = N$. We enumerate *edge properties.* Strategy 1 is said to *absorb* if $V_1(N) > V_0(N-1)$. It is said to *reflect* if $V_1(N) \leq V_0(N-1)$. A similar definition holds for strategy 0. The motivation for this terminology is that if i absorbs, and we somehow reach $n_i = N$, this state would persist since no decision-maker could increase her utility by choosing an action other than i. When i absorbs it is a symmetric (strict) Nash equilibrium of the game. If we somehow reach a state $n_i = N$, and i reflects, then decision-makers are no worse off choosing something else (in case the inequality is strict, they are better off choosing something else; if equality holds, we assume there is some chance that they choose something else). Each edge has all players playing some strategy i; the edge is said to absorb if the corresponding i absorbs.

We define the edge properties in terms of the payoff differential

$$\Delta(m) = V_1(m) - V_0(m-1).$$

Suppose m people choose action 1. Then each gets a payoff $V_1(m)$. If one of these people decided to choose action 0, the state would change (so that $n_1 = m - 1$ and $n_0 = N - m + 1$). $V_0(m-1)$ is the payoff of a person who decides to switch actions. $\Delta(m)$ measures the difference. There are four possible edge conditions:

AA Both edges absorb if $\Delta(1) < 0, \Delta(N) > 0$,

RR Both edges reflect if $\Delta(1) \geq 0, \Delta(N) \leq 0$,

RA The first edge reflects and the second absorbs if $\Delta(1) \geq 0, \Delta(N) > 0$,

AR The first edge absorbs and the second reflects if $\Delta(1) < 0, \Delta(N) \leq 0$.

We can classify a number of examples in terms of the properties of Δ and the edge properties. First, we have *linear* games where payoffs can be described by $\Delta(m) = d_0 + d_1 m$. Consider a population game, where a finite number of players are matched at random to play a symmetric game with two actions. $V_1(m)$ then denotes the expected payoff from action 1 when $m-1$ others choose the same action. The resulting $\Delta(m)$ will be linear in m. The second case is that of *monotone* games where $\Delta(n)$ is either increasing or decreasing. Within the class of monotone games, assuming the inequalities in the edge conditions are strict, the edge properties allow us to define:

[3] We are also interested in "small" π, which raises the possibility of discontinuities — $\lim_{\pi \to 0} \mu^*(m, \pi) \neq \mu^*(m, 0)$. We show robust properties true for $\pi > 0$.

(**AA**) Co-ordination game with $n_1 = N$ and $n_1 = 0$ as two equilibria.

(**RR**) Congestion game, with a unique equilibrium $n_1 = n^*$ where $0 < n^* < N$.

(**RA**) Dominance solvable, with $n_1 = N$ its unique equilibrium.

(**AR**) Dominance solvable, with $n_1 = 0$ its unique equilibrium.

Finally, we consider the case of *non-monotone* games, involving no restriction on the shape of Δ. The edge properties continue to identify equilibria for us, but there could be additional equilibria (e.g. $\Delta(m) \geq 0$, $\Delta(m+1) \leq 0$ for $m \notin \{0, N\}$). Even in this case, properties of the stationary distribution will depend only on edge properties. This despite the fact that these properties are sufficient for the characterization of equilibria only in monotone games.

3 Results

We now derive the stationary distribution of the dynamic model of choice specified in section 2. The transition probabilities (P_{ij}) of the Markov chain can be defined in terms of the values of the payoff differential $\Delta(\cdot)$. Recall that the *state* was defined above as the number of people (m) who choose action 1. Since only one player has a choice at any given date, the state can change by at most one ($P_{ij} = 0$ whenever $|i - j| > 1$). We call an increase of one a *birth,* and a decrease of one a *death.* It will be convenient to use the following notation for birth and death probabilities: $b(i) = P_{ii+1}$ and $d(i) = P_{ii-1}$.

A birth occurs only when a player whose current choice is action 0 gets a decision opportunity, and decides to play action 1 instead. In this case she may accidentally choose the opposite action (with probability ϵ) or may imitate someone who plays action 1 (which happens with probability $(1 - \epsilon)\pi m/(N - 1)$). Additionally (with probability $(1-\epsilon)(1-\pi)$) she may optimize, in which case there is a birth if $\Delta(m+1) \geq 0$. This leads to the following birth probabilities:

$$b(m) = \begin{cases} \frac{N-m}{N}(\epsilon + (1-\epsilon)\pi\frac{m}{N-1} + (1-\epsilon)(1-\pi)) & \text{if } \Delta(m+1) \geq 0 \\ \\ \frac{N-m}{N}(\epsilon + (1-\epsilon)\pi\frac{m}{N-1}) & \text{otherwise.} \end{cases} \tag{1}$$

Denote by $b^+(m)$ the value of $b(m)$ when $\Delta(m+1) \geq 0$, and by $b^-(m)$ the value when $\Delta(m+1) < 0$.[4]

A death occurs only when a player whose current choice is action 1 decides to do otherwise. A decision opportunity for such a player arises with probability m/N.

[4] It is possible here, and with $d(\cdot)$, to treat the case of $\Delta(m+1) = 0$ as distinct. We could allow either action to be played with positive probability. This would not affect our conclusions.

Once again she may choose action 0 because of the noise, from imitation, or because this is the optimal choice. The probabilities are given by:

$$d(m) = \begin{cases} \frac{m}{N}(\epsilon + (1-\epsilon)\pi\frac{N-m}{N-1}) & \text{if } \Delta(m) > 0 \\ \\ \frac{m}{N}(\epsilon + (1-\epsilon)\pi\frac{N-m}{N-1} + (1-\epsilon)(1-\pi)) & \text{otherwise.} \end{cases} \tag{2}$$

Denote by $d^+(m)$ the value of $d(m)$ when $\Delta(m) > 0$, and by $d^-(m)$ the value when $\Delta(m) \leq 0$.

Equations 1 and 2 describe the system completely (the only other entries to P that may be positive are P_{ii}, but their values can be inferred from the fact that rows of P sum to one). Two extreme cases are of interest. When $\pi = 0$ we have optimizing behavior only (the best response dynamics governs the evolution of the system). When $\pi = 1$ we have imitation only (a variant of Kirman's dynamic applies). In general we consider a mix of optimization and imitation. We call these the KMRY and Kirman limit respectively.

The process is aperiodic and irreducible, and so has a unique stationary distribution $\mu(\cdot;\pi,\epsilon)$. Since the process is reversible, we can use the *balance equations* to find this distribution. Suppressing the dependence of μ on ϵ and π, we have

$$\frac{\mu(m)}{\mu(m-1)} = \frac{b(m-1)}{d(m)}$$

for $m = 1, 2, \cdots, N$. Define, for $m = 1, \ldots, N$, the likelihood ratio functions

$$\ell(m) = \begin{cases} \frac{b^+(m-1)}{d^+(m)} & \text{if } \Delta(m) > 0 \\ \\ \frac{b^+(m-1)}{d^-(m)} & \text{if } \Delta(m) = 0 \\ \\ \frac{b^-(m-1)}{d^-(m)} & \text{if } \Delta(m) < 0 \end{cases} \tag{3}$$

and let $\ell(0) = 1$. The stationary distribution $\mu(m)$ then satisfies

$$\mu(m) = \frac{\prod_{k \leq m} \ell(k)}{\sum_{n=0}^{N} \prod_{k \leq n} \ell(n)} \tag{4}$$

for each $m = 0, 1, \ldots, N$.

We are interested in the behavior of μ as $\epsilon \to 0$. Define (*)-functions as limits, i.e. $\ell^*(n) = \lim_{\epsilon \to 0} \ell(n)$, similarly $\mu^*(m) = \lim_{\epsilon \to 0} \mu(m;\pi,\epsilon)$. Theorem 1 presents the limit properties of μ corresponding to each edge condition.

Theorem 1. *Let $0 < \pi < 1$. For each edge condition, the corresponding long-run distribution μ^* is as follows:*

(AA) *$\mu^*(m) > 0$ if and only if $m \in \{0, N\}$. Further,*

$$\mu^*(0) = (1 + \prod_{k=2}^{N-1} \ell^*(k))^{-1}.$$

(RR) *$\mu^*(m) \in (0,1)$ for each $m = 0, 1, \ldots, N$. Further,*

$$\mu^*(m) = \frac{\prod_{k \le m} \ell^*(k)}{\sum_{n=0}^{N} \prod_{k \le n} \ell^*(n)}.$$

(RA) *$\mu^*(N) = 1$ and $\mu^*(m) = 0$ for $m < N$.*

(AR) *$\mu^*(0) = 1$ and $\mu^*(m) = 0$ for $m > 1$.*

Proof. From equation 3 above, when $\Delta(m) > 0$,

$$\ell(m) = \left(\frac{N-m+1}{m}\right)\left(\frac{\epsilon + (1-\epsilon)\pi\frac{m-1}{N-1} + (1-\epsilon)(1-\pi)}{\epsilon + (1-\epsilon)\pi\frac{N-m}{N-1}}\right).$$

When $\Delta(m) = 0$ we have

$$\ell(m) = \left(\frac{N-m+1}{m}\right)\left(\frac{\epsilon + (1-\epsilon)\pi\frac{m-1}{N-1} + (1-\epsilon)(1-\pi)}{\epsilon + (1-\epsilon)\pi\frac{N-m}{N-1} + (1-\epsilon)(1-\pi)}\right).$$

And in case $\Delta(m) < 0$, we have

$$\ell(m) = \left(\frac{N-m+1}{m}\right)\left(\frac{\epsilon + (1-\epsilon)\pi\frac{m-1}{N-1}}{\epsilon + (1-\epsilon)\pi\frac{N-m}{N-1} + (1-\epsilon)(1-\pi)}\right).$$

For $m = 2, 3, \ldots, N-1$ the values of $\ell^*(m) \equiv \lim_{\epsilon \to 0} \ell(m)$ are as follows:

$$\ell^*(m) = \left(\frac{N-m+1}{m}\right)\left(\frac{\pi\frac{m-1}{N-1} + (1-\pi)}{\pi\frac{N-m}{N-1}}\right) \quad \text{if } \Delta(m) > 0,$$

$$\ell^*(m) = \left(\frac{N-m+1}{m}\right)\left(\frac{\pi\frac{m-1}{N-1} + (1-\pi)}{\pi\frac{N-m}{N-1} + (1-\pi)}\right) \quad \text{if } \Delta(m) = 0,$$

$$\ell^*(m) = \left(\frac{N-m+1}{m}\right)\left(\frac{\pi\frac{m-1}{N-1}}{\pi\frac{N-m}{N-1} + (1-\pi)}\right) \quad \text{if } \Delta(m) < 0.$$

In particular, $\ell^*(m)$ is strictly positive and finite whenever $m = 2, 3, \ldots, N-1$.

For $m \in \{1, N\}$ note that

$$\ell(1) = \begin{cases} N\frac{\epsilon+(1-\epsilon)(1-\pi)}{\epsilon+\pi(1-\epsilon)} & \text{if } \Delta(1) > 0 \\ N(\epsilon + (1-\epsilon)(1-\pi)) & \text{if } \Delta(1) = 0 \\ N\epsilon & \text{otherwise.} \end{cases}$$

$$\ell(N) = \begin{cases} \frac{1}{N}\frac{\epsilon+(1-\epsilon)\pi}{\epsilon+(1-\pi)(1-\epsilon)} & \text{if } \Delta(N) < 0 \\ \frac{1}{N(\epsilon+(1-\epsilon)(1-\pi))} & \text{if } \Delta(N) = 0 \\ \frac{1}{N\epsilon} & \text{otherwise.} \end{cases}$$

Define $x(m) = \prod_{k \le m} \ell(m)$ and $X = \sum_{m=0}^{N} x(m)$ so that equation 4 can be rewritten as $\mu(m) = x(m)/X$.

Suppose edge condition **AA** is satisfied. In this case $x(0) = \ell(0) = 1$ and, since $\ell^*(1) = 0$, $x^*(m) = 0$ for $m = 1, 2, \ldots N-1$. Now consider $x^*(N)$. Since

$$x(N) = \ell(1)\ell(N) \prod_{k=2}^{N-1} \ell(k) = N\epsilon \cdot \frac{1}{N\epsilon} \prod_{k=2}^{N-1} \ell(k),$$

$x^*(N)$ is finite and equals $\prod_{k=2}^{N-1} \ell^*(k)$. The stationary distribution has support $\{0, 1\}$ with probability $\mu^*(0) = (1 + \prod_{k=2}^{N-1} \ell^*(k))^{-1}$ and $\mu^*(N) = 1 - \mu^*(0)$.

Suppose edge condition **RR** is satisfied. Now $\ell^*(1)$ and $\ell^*(N)$ will be finite and positive. As a consequence, $\mu^*(m) > 0$ for each $m \in \{0, 1, \ldots, N\}$. Further,

$$\mu^*(m) = \frac{\prod_{k \le m} \ell^*(k)}{\sum_{n=0}^{N} \prod_{k \le n} \ell^*(n)}.$$

Suppose edge condition **AR** is satisfied (so that $\Delta(1) < 0$ and $\Delta(N) \le 0$. Since $\Delta(1) < 0$, $\ell(1) = N\epsilon$. When $\Delta(N) \le 0$, there are two cases – $\Delta(N) < 0$ and $\Delta(N) = 0$. In either case, $\ell^*(N)$ converges to a positive and finite value. Since $\ell^*(1) = 0$, we have $x^*(0) = 1$ and $x^*(m) = 0$ for $m > 0$. So $\mu^*(0) = 1$ and $\mu^*(m) = 0$ for $m > 0$. The argument for edge condition **RA** is symmetric. □

Remarks. For the best response dynamics, when payoffs are linear, there is a celebrated 1/2–dominance result according to which the long run outcome depends upon payoff differentials at the state $N/2$ (see, for instance, Ellison 2000). The important insight from Theorem 1 is that, with even a small probability of imitation, the limit behavior of the process responds to edge rather than median properties of payoffs. This has some interesting implications. The **AA** case behaves "like" a coordination game. When payoffs are monotone, the edge condition **AA** implies that we have a genuine coordination games. In this case the limit distribution gives positive weight to both edges (and not just to the risk dominant equilibrium). Somewhat surprising is that this continues to be the case even when payoffs are *not* monotone, and we have other mixed equilibia that may even be selected by the pure best-response dynamics. We return to this point in Counterexample 1 of section 3.1.

The edge conditions **AR** and **RA** would arise if one of the actions were a dominant strategy, in which case the limit distribution would place all its mass at the edge where all players play this strategy. But these conditions could arise in other ways as well. Suppose, for instance, that for some N (possibly quite large) payoffs imply the following values for Δ: $\Delta(1) < 0$, $\Delta(N) < 0$, and $\Delta(m) > 0$ for all other m. Theorem 1 indicates the outcome $\mu^*(0) = 1$. But consider the pure best response dynamics. From all states $m > 1$ the best responses takes us in the direction of $m^* = N - 1$ (at $m = 1$ a player choosing either action would switch). Comparing the sizes of the basin of attraction under the best response dynamics has now become quite futile. Increasing N, and simultaneously the basin of attraction of $m^* = N - 1$, does not help at all.

The final case is **RR**, of congestion at the edges. This is necessary, as well as sufficient, for the support of the limit distribution to be all of $\{0, 1, \ldots, N\}$.

3.1 Kirman and KMRY Limits

Two special cases are of interest. First, when $\pi = 1$ we have imitation only. We call this the Kirman limit. Next is the case of the best response dynamics only, with $\pi = 0$, which we call this the KMRY limit. We also examine the limit as $\pi \to 0$. Is it the case that for small amounts of imitation the KMRY limit is a good approximation? The answer depends on the properties of Δ, specifically whether the game is monotone or not. Instead of general monotone games we focus on finite symmetric games (where the additional restriction is linearity) and then go on to consider non-monotone games.

The following corollary is immediate from the previous definitions of $\ell(\cdot)$.

Corollary 1 (Kirman Limit)**.** *Let* $\pi = 1$. *Then* $\mu^*(0) = \mu^*(N) = 1/2$, *and* $\mu^*(m) = 0$ *for* $1 < m < N$.

Proof. When $\pi = 1$

$$\ell(m) = \left(\frac{N-m+1}{m}\right)\left(\frac{\epsilon + (1-\epsilon)\frac{m-1}{N-1}}{\epsilon + (1-\epsilon)\frac{N-m}{N-1}}\right)$$

regardless of the value of Δ. For $m = 2, \ldots, N-1$,

$$\ell^*(m) = \frac{N-m+1}{m} \cdot \frac{m-1}{N-m},$$

which is finite and positive. Now $\ell(1) = N\epsilon$ and $\ell(N) = \frac{1}{N\epsilon}$. So $x^*(N) = \prod_{m=2}^{N-1} \ell^*(m) = 1$. In addition, $x^*(0) = 1$ and $x^*(m) = 0$ for $m = 2, \ldots, N-1$. This implies that $\mu^*(0) = \mu^*(N) = 1/2$, and $\mu^*(m) = 0$ for $1 < m < N$. □

The KMRY limit is the case of optimization only ($\pi = 0$). Now the outcome *will* depend upon the value of $\Delta(\cdot)$ over $\{1, 2, \ldots, N\}$. A number of papers have examined outcomes for this case with additional assumptions about payoffs (e.g. symmetric population games and general coordination games). We have imposed no restrictions on the shape of Δ as yet, though we will consider several cases later. The next result shows that it is still possible to reach some conclusions about the long run outcomes that can arise. We note that if $\Delta(1) \le 0$ ($\Delta(N) \ge 0$) then action 0 (action 1) is a symmetric Nash equilibrium. But there could be "mixed" Nash equilibria as well – this occurs if $\Delta(m) \ge 0$ and $\Delta(m+1) \le 0$ for some m. When m players choose action 1 and $N-m$ choose action 0, a player in neither group can gain from deviation.

Corollary 2 (KMRY Limit)**.** *Let* $\pi = 0$. *Then*

1. $\mu^*(0) > 0$ *only if* $\Delta(1) \le 0$.
2. $\mu^*(N) > 0$ *only if* $\Delta(N) \ge 0$.
3. $\mu^*(m) > 0$ *for* $1 < m < N$ *only if* $\Delta(m) \ge 0$ *and* $\Delta(m+1) \le 0$.

Proof. When $\pi = 0$, $\ell(m) = (N - m + 1)(\epsilon)^z/m$ where $z = -1$ if $\Delta(m) > 0$, $z = 0$ if $\Delta(m) = 0$, and $z = 1$ if $\Delta(m) < 0$.

(1) Suppose $\mu^*(0) > 0$ and $\Delta(1) > 0$. $\Delta(1) > 0$ implies that $x^*(1) = \ell^*(1) = \infty$. Now $\mu^*(0) = 1/X^*$, so that $X^* < \infty$. This implies that $\mu^*(1) = \infty$, a contradiction.

(2) Suppose $\mu^*(N) > 0$ and $\Delta(N) < 0$. The latter assertion implies that $\ell(N) = \epsilon/N$. From the balance conditions $\ell(N) = \mu^\epsilon(N)/\mu^\epsilon(N-1)$. For ϵ small, $\mu^\epsilon \gg 0$ so that the convergence of $\ell(N)$ to zero requires $\mu^\epsilon(N-1) \to \infty$, a contradiction.

(3) Suppose $\mu^*(m) > 0$ and $\Delta(m) < 0$ for some $1 < m < N$. The latter implies $\ell(m) = (N-m+1)\epsilon/m$ which converges to zero as $\epsilon \to 0$. Since $\ell(m) = \mu^\epsilon(m)/\mu^\epsilon(m-1)$ and $\mu^\epsilon(m) > 0$, we require $\mu^\epsilon(m-1) \to \infty$, a contradiction. The other case to consider is $\mu^*(m) > 0$ and $\Delta(m+1) > 0$. The latter assertion implies that $\ell(m+1) = (N-m)/((m+1)\epsilon)$. Since $\ell(m+1) = \mu^\epsilon(m+1)/\mu^\epsilon(m)$ and $\mu^\epsilon(m) > 0$, for $\ell(m+1) \to \infty$ we require $\mu^\epsilon(m+1) \to \infty$ which is impossible. □

If we impose further restrictions on Δ, we can characterize the distribution μ^* fully. An important case is that of symmetric strategic form games played by a population of players. The players are assumed to be randomly matched, and a player's utility is the expected payoff given the population distribution of strategies. There are four cases to consider: (1) coordination games (this gives rise to edge condition **AA**), (2) games with asymmetric strict Nash equilibria, e.g. "hawk–dove," (giving rise to edge condition **RR**), (3) dominance solvable games (with edge condition **AR**), and (4) dominance solvable games with edge condition **RA**. An application of Theorem 1 gives us the corresponding long run outcomes. For comparison, we state the KMRY result. The proof is omitted (the case of **AR** and **RA** is obvious, **RR** follows from the previous corollary upon noting that there is only one mixed Nash equilibrium here, and the argument for **AA** is well known).

Corollary 3 (Strategic Form Games)**.** *Suppose* $\pi = 0$*. Then,*

(AA) μ^* *assigns probability one to the state where all players choose the risk-dominant action,*

(RR) μ^* *assigns probability one to the mixed Nash equilibrium,*

(RA) $\mu^*(N) = 1$

(AR) $\mu^*(0) = 1$.

The significant difference here is with **AA** and **RR**. In the case of **AA**, when $\pi \in (0, 1)$, Theorem 1 assigns strictly positive probability to the states where the pure strategy Nash equilibria are played. It is possible to show, however, that as $\pi \to 0$ we have convergence to the KMRY limit. In the case of **RR**, the result of Theorem 1 indicates positive probability for all states, whereas the KMRY result yields a degenerate distribution. Next, we examine the limit, as $\pi \to 0$, of the distribution $\mu^*(\cdot; \pi)$ (first in the strategic game case, and then more generally). This limit is well behaved for monotone games, but there may be a discontinuity for non-monotone games.

Corollary 4. *For any strategic form game, with* $\pi > 0$*, the stationary distributions* $\mu^*(m; \pi)$ *satisfy* $\lim_{\pi \to 0} \mu^*(m; \pi) = \mu^*(m; 0)$.

Proof. This is clearly true in case **AR** or **RA** is satisfied. Now suppose **AA** is satisfied, and action 1 is risk-dominant. Then there exists $M^* < N/2$ such that

$$m \geq M^* \Rightarrow \Delta(m) > 0.$$

Now for $\pi > 0$ Theorem 1 yields $\mu^*(0) = (1 + \prod_{m=2}^{N-1} \ell^*(m))^{-1}$ and $\mu^*(N) = 1 - \mu^*(0)$. Some routine calculations show that

$$\prod_{m=2}^{N-1} \ell^*(m) = \prod_{m=M^*}^{N-M^*-1} \frac{\pi(m-1) + (1-\pi)(N-1)}{\pi(N-m)}.$$

Each term in the product diverges since $(1-\pi)/\pi \to \infty$ as $\pi \to 0$. Consequently $\mu^*(0) \to 0$ and $\mu^*(N) \to 1$.

When **RR** is satisfied ("hawk-dove") there is a unique mixed equilibrium (in addition to two asymmetric strict equilibria). From Theorem 1

$$\mu^*(m) = x^*(m) / \sum_{m=0}^{N} x^*(m).$$

Further, there exists some m^* such that $\Delta(m) > 0$ for $m \leq m^*$ and $\Delta(m) < 0$ for $m > m^*$. For $m \leq m^*$

$$x^*(m) = \prod_{k \leq m} \ell^*(k) = \prod_{k \leq m} \left(\frac{N-k+1}{k} \right) \left(\frac{\pi(k-1)/(N-1) + (1-\pi)}{\pi(N-k)/(N-1)} \right).$$

Observe that these terms are $O(\pi^{-m})$ for $m \leq m^*$. A similar argument shows that for $m > m^*$ the $x(m)$ terms, when $m = m^* + i$ are $O(\pi^{-m^*+i})$. So X^* is a sum of $x^*(m^*)$ and terms of $O(\pi^{-r})$ where $r < m^*$. So

$$\mu^*(m^*) = x^*(m^*)/X^* \to 1 \text{ as } \pi \to 0.$$

This is precisely the KMRY limit $\mu^*(\cdot; 0)$. □

It will be evident from an inspection of the proof that the conclusion of Corollary 4 holds for any symmetric monotone game. We state, without proving,

Corollary 5. *For any monotone game,* $\lim_{\pi \to 0} \mu^*(m; \pi) = \mu^*(m; 0)$.

The results on monotone games rely crucially on the fact that the edge properties determine global properties of payoffs. There are a number of interesting economic examples where Δ is non-monotone. Results can then be quite different. This is illustrated next with a simple counterexample.

Counterexample 1. *Suppose* $N = 6$, $\Delta(1) < 0$, $\Delta(2) > 0$, $\Delta(3) > 0$, $\Delta(4) < 0$, $\Delta(5) < 0$, *and* $\Delta(6) > 0$. *Then*

$$\lim_{\pi \to 0} \mu^*(m; \pi) \neq \mu^*(m; 0).$$

Proof. For all $\pi > 0$, since **AA** holds, Theorem 1 imples that the support of $\mu^*(\cdot; \pi)$ is $\{0, N\}$. We compute $\mu^*(\cdot; 0)$. Note that $\ell(0) = 1$, $\ell(1) = 6\epsilon$, $\ell(6) = 1/\ell(1)$, $\ell(2) = 5/(2\epsilon)$, $\ell(5) = 1/\ell(2)$, $\ell(3) = 4/(3\epsilon)$, and $\ell(4) = 1/\ell(3)$. Routine calculations show that $x(3) = 20/\epsilon$ and $X = 32 + 12\epsilon + 20/\epsilon$. Then

$$\mu^\epsilon(3; 0) = \frac{20}{32\epsilon + 12\epsilon^2 + 20}.$$

We have

$$\mu^*(3; 0) \equiv \lim_{\epsilon \to 0} \mu^\epsilon(3; 0) = 1.$$

□

We note finally that there is nothing pathological about the pattern of payoffs here (which can be generated by quasi-concave utility functions). This discontinuity can arise in the context of natural economic examples.

3.2 Distributions for Large N

Having studied long run properties of the dynamic model of choice for a fixed N, we are now interested in determining how the process evolves when N becomes large. Our approach so far has been to compute the stationary distribution μ, and its limit μ^*. Here, we focus on **RR** games, assume $\epsilon = 0$, take $\pi = N/(N+q)$, and study the asymptotic form of μ as $N \to \infty$. In the **RR** case, so long as there is positive probability of playing best responses, the process is irreducible and aperiodic even when $\epsilon = 0$. So there is a unique stationary distribution. Its limit distribution can be approximated by a continuous function, assumed to have a density $f(\cdot)$.

Let $x = m/N$, and note that

$$\ell(x) = \frac{f(x)}{f(x - \frac{1}{N})},$$

where $f(x)$ is the density of the limit distribution $\mu(\cdot)$ on $(0,1]$. From this,

$$N(\ell(x) - 1) = \frac{(f(x) - f(x - \frac{1}{N})}{\frac{1}{N}} \frac{1}{f(x - \frac{1}{N})}$$

so that

$$\frac{\partial \ln f(x)}{\partial x} = \frac{f'(x)}{f(x)} = \lim_{N \to \infty} (N\ell(x) - 1)$$

whenever the density is differentiable at x. Suppose, for some positive constant q, that $\pi = \frac{N}{N+q}$, i.e. $\pi = 1 - O(\frac{1}{N})$. We can rewrite:

$$\ell^+(x) = \frac{1 - x + \frac{1}{N}}{x} \frac{\pi(x - \frac{1}{N}) + (1-\pi)(1 - \frac{1}{N})}{\pi(1-x)};$$

$$\ell^0(x) = \frac{1 - x + \frac{1}{N}}{x} \frac{\pi(x - \frac{1}{N}) + (1-\pi)(1 - \frac{1}{N})}{\pi(1-x) + (1-\pi)(1 - \frac{1}{N})};$$

$$\ell^-(x) = \frac{1 - x + \frac{1}{N}}{x} \frac{\pi(x - \frac{1}{N})}{\pi(1-x) + (1-\pi)(1 - \frac{1}{N})};$$

and note that

$$\begin{aligned} \ell(x) &= \ell^+(x) \quad \text{if} \quad \Delta(x) > 0 \\ \ell(x) &= \ell^0(x) \quad \text{if} \quad \Delta(x) = 0 \\ \ell(x) &= \ell^-(x) \quad \text{if} \quad \Delta(x) < 0. \end{aligned}$$

Define $g^+(x;q) = \lim_{N \to \infty} N(\ell^+(x) - 1)$ with $\pi = \frac{N}{N+q}$. We obtain, similarly, $g^0(x;q)$ and $g^-(x,q)$. Now observe that

$$g^{+}(x,q) = \lim_{N}[\frac{N(1-\pi)}{\pi x} + \frac{1}{\pi(1-x)} - \frac{1}{x} - \frac{1}{N\pi x(1-x)}$$

implying

$$g^{+}(x,q) = \frac{q-1}{x} + \frac{1}{1-x}.$$

Similar calculations imply

$$g^{0}(x,q) = \frac{q-1}{x} - \frac{q-1}{1-x}$$

$$g^{-}(x,q) = -\frac{1}{x} + \frac{q-1}{1-x}.$$

We obtain:

$$f(x) \propto x^{q-1}(1-x)^{-1} \quad \text{if} \quad \Delta(x) > 0;$$
$$f(x) \propto x^{q-1}(1-x)^{q-1} \quad \text{if} \quad \Delta(x) = 0;$$
$$f(x) \propto x^{-1}(1-x)^{q-1} \quad \text{if} \quad \Delta(x) < 0.$$

We enumerate some properties of the density.

1. Integrating constants can be chosen to ensure continuity of the density function $f(x)$, and thus differentiability of the associated distribution function $F(x) = \int_0^x f(t)dt$ defined over $(0, 1]$.

2. This distribution is related to the Beta distribution, which has density: $b(x; a, b) \propto x^{a-1}(1-x)^{b-1}$ for $a > 0, b > 0$; the distribution F obtains as a mixture of $B(0, q)$, $B(q, q)$ and $B(q, 0)$.

3. The distribution is Beta(q, q) whenever $\Delta(x) = 0$ for all x, as in Kirman (*op. cit.*)

4. In a symmetric RR game, with $\Delta(1/2) = 0$, we obtain a symmetric distribution.

5. The distribution is unimodal (with tails thinner than Beta) in the symmetric *monotone* **RR** case.

4 Economic Applications

For ease of exposition, we consider simple parametric cases but it will be evident that each example is at the start of a very extensive literature. There are two examples of monotone games (a coordination game of technology choice and a market game with congestion). The other two examples are of non-monotone games (a location game with both spillovers and crowding, and a club economy).

4.1 Coordinated technology choice

Coordination games have been extensively studied in the evolutionary game theory literature, with the example of network externality in technology choice being a motivating example. A very general treatment of this problem is by Kandori and Rob (1998) who consider bandwagon effects and competing technology choice when there are more than two competing technologies and various degrees of partial compatibilities between them. Our example is in the same spirit, although it is closer in its explicit formulation to spatial increasing return models in which workers choose firms or locations in response to wage differentials (e.g. Fujita, Krugman, and Venables, 1999). In these models, as in other compelling papers by Diamond (1982) and Cooper and John (1988), there are multiple equilibria and the possibility of coordination failure.

There are N workers, each with one unit of inelastically supplied labor. There are two firms who use different technologies to produce a single consumption good. Firm 0 operates a technology with constant returns to scale: if it has n_0 workers, the total output is given by

$$X_0(n_0) = \sum_{i=1}^{n_0} \ell_i = n_0.$$

Firm 1 operates a technology with increasing returns to scale. Workers must spend a fixed amount of labor, $c < 1$, to access the technology (say because it requires a continuous investment in skills) and can use the rest in production.

$$X_1(n_1) = (\sum_{i=1}^{n_1} (\ell_i - c))^{1+\alpha} = (1-c)^{1+\alpha} n_1^{\alpha}; \quad \alpha > 1.$$

Output is divided equally among workers. Individuals choose to join one or another firm, work and produce output, and consume their share. Payoffs are

$$V_0(N - n_0) = \frac{X_0}{n_0} = 1$$

$$V_1(n_1) = \frac{X_1}{n_1} = (1-c)^{1+\alpha} n_1^{\alpha}.$$

We obtain

$$\Delta(n) = V_1(n) - V_0(n-1) = (1-c)^{1+\alpha} n^{\alpha} - 1,$$

illustrated in Figure 1 with $N = 50$, $c = 0.85$, and $\alpha = 1.25$. This is a monotone co-ordination game ($\Delta(n+1) > \Delta(n)$) with property **AA** whenever $N > 1/(1-c)^{\frac{1+\alpha}{\alpha}}$, assumed true from now on. Note that

$$\Delta(n) \geq 0 \Leftrightarrow n \geq n^* \equiv \frac{1}{(1-c)^{\frac{1+\alpha}{\alpha}}}.$$

This game has two equilibria, $n = 0$ and $n = N$. The latter is efficient, and maximizes the sum of payoffs $(N-n)V_0(n) + nV_1(n) = (N-n) + (n(1-c))^{1+\alpha}$.

In this example we have a monotone coordination game which satisfies property **AA**. $\Delta(\cdot)$ increases with n. The long run outcome in the KMRY limit can be found by comparing the cardinality of basins of attraction. The basin of $n = 0$, the constant returns to scale technology, is the set of states for which $\Delta(\cdot) < 0$. The basin of $n = N$, the increasing returns technology, is the set of states for which $\Delta(\cdot) > 0$. The latter is the efficient state. In the KMRY case, the distribution $\mu^*(\cdot; 0)$ assigns probability one to the state with the larger basin of attraction. The size of the basin depends upon values of the parameters. For instance, when $N = 10$, $\alpha = 0.5$, and $c = 0.1$, the long run outcome is $n = N$. If we increase c to 0.5, the long run outcome will be $n = 0$ (so that KMRY can pick the inefficient technology). By comparison, when $\pi > 0$ we have strictly positive probability for both equilibria but get the KMRY limit as $\pi \to 0$.

4.2 Clubs with congestion

In models of club formation, agents form groups so as to benefit from externalities they create for one another. In Buchanan's (1965) pioneering model, the benefit arises because members are able to share the costs of an excludable public good. Potential gains are limited by the fact that crowding is undesirable. Increasing club membership reduces the per capita costs of the public good, but increases crowding costs (leading typically to a finite optimal size). There is an extensive literature on clubs, reviewed by Scotchmer (2001).

N individuals choose whether to subscribe to a club that produces an excludable public good x. Consumption suffers from congestion. If a group has n_1 members, and collects subscription revenues of R, the total production of x is $x(n_1) = AR + \frac{c}{n_1}$. We assume a subscription fee of 1 per member so that $R = n_1$. Each individual, i, chooses whether to join the club ($a_i = 1$) or not ($a_i = 0$). Payoffs are

$$V_1(n_1) = An_1 + \frac{c}{n_1} - 1;$$

and

$$V_0(N - n_0) = 0.$$

From this, we have

$$\Delta(n) = \frac{1}{n}(An^2 - n + c).$$

This game has property **RA** if $c > 1 - A$ and $AN^2 - N + c > 0$. The game is non-monotone, and

$$\Delta(n) < 0 \Leftrightarrow n \in (n_L, n_H)$$

where

$$n_L = \frac{1 - \sqrt{1 - 4Ac}}{2A} \quad \text{and} \quad n_H = \frac{1 + \sqrt{1 - 4Ac}}{2A}).$$

This Δ is illustrated in Figure 2. This game has three equilibria, with $n = n_L$, $n = n_H$, and $n = N$. For N large enough, the third equilibrium, with full subscription, is efficient.

In this model the game has property **RA** whenever $c > 1-A$ and $AN^2-N+c > 0$. From Theorem 1, $\mu^*(N;\pi) = 1$ for all $\pi > 0$: the full subscription equilibrium is the long run outcome. By contrast, KMRY can pick the interior (possibly inefficient) outcome. Figure 2 illustrates this possibility. The full subscription equilibrium ($n = N$) is labelled C, and will be the long run outcome when $\pi > 0$. The KMRY outcome is $n = n_L$, labelled D. The basin of attraction of D is all of $\{0, 1, \ldots, n_H\}$, a very large proportion of the state space.

4.3 A market game

In strategic market game models traders simultaneously communicate bids and offers of supply of goods to the market (Shapley and Shubik, 1977, and an extensive recent literature). The ratio of bids to supplies then determines the terms of trade. In this example we examine how prices emerge from the buying and selling decisions of individuals, and explore the properties of price distributions in the presence of optimization and imitation.

There are N players and two goods, 0 and 1. All players are endowed with one unit of each good. Their utility from consuming (x_0, x_1) is

$$U(x_0, x_1) = x_0 + \theta x_1.$$

Players can bid one unit of either of these goods, to buy the other. The amount they get in exchange depends on the price, determined to balance demand and supply (goods are divisible, bids are not). Strategies $a_j \in \{0, 1\}$ correspond to bidding a unit of good j. Let $n_1 = \sum_i a_i$, and $n_0 = \sum_i (1 - a_i)$ be the number of bids on each side. This results in the price $p(n_0, n_1) = n_1/n_0$. Bidders get p units of good 1 or $\frac{1}{p}$ units of good 0.

Payoffs are:

$$V_0(n) = \theta(1 + p(n)) = \theta \frac{N}{N-n};$$

$$V_1(n) = 1 + \frac{1}{p(n)} = \frac{N}{n}.$$

We obtain

$$\Delta(n) = N(\frac{1}{n} - \frac{\theta}{N-n+1}),$$

which is illustrated in Figure 3 with $N = 50$ and $\theta = 1$. This a monotone congestion game ($\Delta(n+1) < \Delta(n)$) and has property **RR** whenever $\frac{1}{N} < \theta < N$. We assume this to be true, and note that

$$\Delta(n) \geq 0 \Leftrightarrow n \leq n^* \equiv \frac{N+1}{1+\theta}$$

$$\Leftrightarrow p \leq p^* \equiv \frac{N+1}{N\theta - 1}.$$

This game has a unique equilibrium in pure strategies whenever n^* is an integer. (The unique mixed strategy equilibrium has support on integers contiguous to n^* otherwise). We note in passing that every outcome $n \in \{1, \cdots, N-1\}$ is efficient, because the sum of payoffs, $(N-n)V_0(n) + nV_1(n) = N(1+\theta)$ is invariant to N.

The market game is a monotone congestion game and satisfies **RR**. From Theorem 1 we note that the support of μ^* is all of $\{0, 1, \ldots, N\}$. Let $n^* = (N+1)/(1+\theta)$ be the integer solution of $\Delta(n) = 0$ (the non-integer solution case is similar). The KMRY limit assigns probability one to n^*, in which case the corresponding price (number of units that bidders for good 1 get) is $p = p^* \equiv (N+1)/(N\theta - 1)$. When $\pi > 0$, the price p ranges from 0 to ∞, with the different prices arising with the probability of the corresponding $0 \leq n \leq N$, i.e. with distribution given by $\mu^*(n)$ defined in Theorem 1. The model displays price volatility. The volatility decreases with π and, as indicated by Corollary 5, we get convergence to a degenerate distribution which assigns probability one to p^*.

In Figure 4 we plot the price distributions from a simulation of the market game (for $N = 10$ and $\theta = 0.5$).[5] We report results for a number of values of π (the support of the distribution is truncated at $N-1$ to exclude the infinite price when $n = N$). For $\pi = 0$, we have an approximation of the KMRY price distribution. The distributions for $\pi = 0.2$ and $\pi = 0.5$ are also graphed. As discussed in section 3.2, we can obtain the price distribution directly in the limit ($\epsilon = 0$) if we take $\pi = N/(N+q)$ and $N \to \infty$ for some positive constant q. This distribution, with the support truncated as before, is illustrated in Figure 5.

We graph the conditional expectation of the price at time $t+1$, p_{t+1}, given the time t price p_t. In Figure 6, the bold line is $\mathbb{E}(p_{t+1}|p_t)$; also drawn is the dashed 45° line. A very interesting feature of this graph is that if we increase π, the expectation function becomes closer to, and eventually coincides with, the 45° line. In other words, it begins to look like a random walk. We present a plot of the corresponding conditional variance in Figure 7.

4.4 Location with pricing

Economic geography is concerned with interactions between the location of economic activity and concentrations of population (Krugman, 1995, and Fujita, Krugman and Venables, 1999). These authors model the cumulative process by which spatial concentration arises and reinforces itself using an evolutionary dynamic story. In their model, workers migrate to locations that offer higher wage rates and, employing an evolutionary stability criterion (specifically, the replicator dynamics), they are able to characterize some equilibria as stable.

[5]We use 20,000 iterations and $\epsilon = 0.01$. In each iteration exactly one person gets a decision revision opportunity.

In our model, there are two locations 0 and 1, and N individuals who choose to move to one of these locations. In this analysis, we want to think of N as a large number, and evaluate decisions and outcomes for large N. Let $x = n_0/N$ be the proportion of the population at location 0, and $1 - x$ the proportion at location 1. Individuals prefer to move to a location with a larger population, reflecting positive spillovers; however, they must pay for land at their preferred location. The price of land at each location increases with congestion. Specifically,

$$V_0(1-x) = Ax - p(x).$$

We assume that

$$p(x) = 3x^2 - 2x^3 \quad \text{for} \quad 0 \le x \le 1.$$

This choice drives our results via its effect on the shape of Δ — it is a convenient way of obtaining the properties of Δ that arise in a natural way from more complicated models (see, for instance, chapter 5 of Fujita *et. al.*, 1999). Similarly,

$$V_1(x) = A(1-x) - p(1-x).$$

Note that the two locations are *ex-ante* identical, and the implied game is symmetric. Efficient outcomes maximize

$$\bar{V}(x) = xV_0(x) + (1-x)V_1(x).$$

In this problem, $x = 1/2$ is efficient if and only if

$$V(1/2) \ge \max\{V_0(1-x), V_1(0)\} \;\Leftrightarrow A < 1.$$

If $A \ge 1$, spillovers overwhelm congestion, and symmetric outcomes $x = 1$ or $x = 0$ are efficient: the entire population should live at the same location. We note that

$$\begin{aligned}\Delta(x) &= V_0(1-x-\frac{1}{N}) - V_1(1-x) \\ &\propto V_0(1-x) - V_1(1-x)\end{aligned}$$

for N large. We obtain

$$\Delta(0) < 0 \Leftrightarrow A > 1 \Leftrightarrow \Delta(1) > 0.$$

In other words, the game is **AA** whenever $A > 1$ (from symmetry, either is efficient) and is **RR** otherwise. The **RR** case is like the market game and we focus on $A > 1$.

$$\Delta(x) = A(2x-1) - 3(x^2 - (1-x)^2) + 2(x^3 - (1-x)^3)$$

is non-monotone, and $\Delta(x) = 0$ has three distinct solutions whenever $1 < A < 3/2$ (see Figure 8). In addition to the outcomes $x = 0$ and $x = 1$, $x = 1/2$ is also an equilibrium. With our large N approximation, the two other zeros of Δ are also equilibria (but not so for finite N). As it turns out, they are never stable so that the approximation assumption is innocuous.

For the location problem the edge condition **AA** is satisfied whenever $A > 1$ (but, as Figure 8 indicates, this is not a monotone game). Suppose $3/2 > A > 1$. The solution $x = 1/2$ is stable in the KMRY case whenever $\Delta(1/4) > 0$ and $\Delta(3/4) < 0$. This is because the basin of attraction, B, of $x = 1/2$ is more than twice the size of its complement in $[0, 1]$. It will always require fewer mutations to enter B than to escape it. But in the presence of imitation, since the game satisfies **AA**, Theorem 1 tells us that the support of the limit distribution will be $\{0, N\}$. The KMRY outcome is not robust to the presence of imitation. Small amounts of imitation lead to a great deal of uniformity in choices. In the long run we get crowding as the population is either all at location 0 or location 1.

When $A < 1$, the location game satisfies **RR**. The analysis in this case is very similar to that of the market game, and is omitted. However we include a picture, Figure 9, of the distribution of location choices in the limit ($\epsilon = 0$).

5 Extensions

The dynamics of choice specified in section 2 and 3 relies upon a model of imitation derived from Kirman. In section 5.1, we show that Theorem 1 remains valid for a very wide class of imitation dynamics. The only requirement is that, at each date, any agent can be imitated by another with positive probability. We extend the results in two other directions, relaxing the assumption that the full distribution of choices is available to agents. Motivated by Young (1993), we describe a model where agents sample from the recent history of choices and, following Ellison (1993), we describe a local interaction model. In the latter case, we focus on 2×2 coordination games, and are able to obtain convergence results when the population becomes large.

5.1 General Imitation Rules

The only property of Kirman's imitation model that plays a role in Theorem 1 is that a decision maker needs to have positive probability of being influenced by everyone else. This is satisfied by a number of other plausible imitation rules. For instance, suppose a decision maker draws a random sample of size k and imitates according to a probability distribution which is the sample frequency of choices. We get the same results as before. As another example, suppose that there is some notion of neighborhood or distance between agents and agents who are "close" exert a greater influence. But so long as "distant" agents have a strictly positive probability of being imitated the conclusion of Theorem 1 holds exactly. The defining property is:

Property A. For each date t, and every agent j, there is positive probability of j's action being imitated by player $i \neq j$ chosen to be the decision-maker at $t + 1$.

As before, we assume there is positive probability that the decision-maker at date $t+1$ chooses a best response.

Theorem 2. *Suppose* $0 < \pi < 1$, *and Property A is satisfied. Then,*

(AA) *The support of* μ^* *is* $\{0, N\}$.

(RR) *The support of* μ^* *is* $\{0, 1, \dots, N\}$.

(RA) *The support of* μ^* *is* $\{N\}$.

(AR) *The support of* μ^* *is* $\{0\}$.

Further, if $\pi = 1$ *then the support of the stationary distribution is* $\{0, N\}$.

5.2 Limited Information

Our results above assume that the population distribution of choices is known. This information requirement can be substantially weakened. Suppose the population distribution of choices cannot be observed. However there is a record of the T most recent choices made: $y = (y_1, \dots, y_T)$. Y denotes the set of binary vectors of length T.

The decision maker observes a random sample of length k from y, which is the basis for her decision. We first consider the case where she imitates according to a probability distribution which is the sample frequency of choices. Let y^k be the k-vector observed by the decision maker, and let $\theta(y^k)$ denote her choice. Denote the empirical frequency of action 1 by $q(y^k) \equiv \sum_{i=1}^k y_i^k / k$. Then

$$\theta(y^k) = \begin{cases} 1 & \text{with probability } q(y^k) \\ \\ 0 & \text{with probability } 1 - q(y^k) \end{cases}$$

This will be called the *proportional* imitation rule. We can combine this with optimization relative to the agent's sample y^k. I.e. we define $\Delta(\cdot)$ directly on Y^k, the set of binary k-vectors. In case $\Delta(y^k) \geq 0$ the optimal decision is to choose 1 whereas if $\Delta(y^k) < 0$ the best thing is to choose 0. We perturb the process by introducing mistakes as before. Let $z \in Y$ be the zero vector $(0, 0, \dots, 0)$ and let $u \in Y$ denote the vector $(1, 1, \dots, 1)$.

Theorem 3. (1) *With the proportional imitation rule, when* $\pi = 1$, $\lim_{\varepsilon \to 0} \mu_x^\varepsilon > 0$ *if and only if* $x \in \{z, u\}$. (2) *Suppose* $0 < \pi < 1$. *Corresponding to each edge condition, the support of the stationary distribution is as follows:* **(AA)** $\{z, u\}$, **(RR)** Y, **(RA)** $\{u\}$, **(AR)** $\{z\}$.

5.3 Local Interaction

We consider 2×2 coordination games as in Ellison (1993). The two actions are labelled A and B, where A is risk-dominant. There are N players located on a circle. Ellison considered $2k$ neighborhoods of i: the neighbors of i are $\{i-k, i-k+1, \ldots, i-1, i+1, \ldots, i+k-1, i+k\}$. We limit attention to $k = 1$ where the neighbors of i are the adjacent players $i-1$ and $i+1$ (note that $i = 1$ and $j = N$ are neighbors). The state of the dynamic system will be denoted by a string of length N denoting the choices of the players, e.g. $x = AABAB \ldots AB$ with $x(i)$ denoting the choice of the i-th player.

With the best-response dynamics, for transition $(x \to x')$, $x'(i) = A$ if at least one of $x(i-1)$ and $x(i+1)$ equals A. In this case there are two possible limiting (or absorbing) states $\mathbf{A} \equiv AAA \ldots A$ and $\mathbf{B} \equiv BBB \ldots B$. When N is even there is also a limit cycle (denoted by $\mathbf{C}$) where the system alternates between $ABAB \ldots AB$ and $BABA \ldots BA$. When we consider a perturbed version of this process, with a probability of error ε, the long run distribution puts all of its mass on $\mathbf{A}$ (Ellison, 1993).

The imitation process is defined as follows. At each date, each player, imitates one of her neighbors with equal probability. With the imitation process, the states $\mathbf{A}$ and $\mathbf{B}$ and the cycle $\mathbf{C}$ are again the only possible outcomes. However, perturbing the process does not narrow down the set of possibilities – the limit distribution gives positive probability to each of the three outcomes (we show this below). We can combine imitation and optimization so that each player has probability π of imitation and probability $(1-\pi)$ of choosing a best response. Again $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ are all in the support of the limit distribution.

Theorem 4. *Suppose $0 < \pi \leq 1$. Then the support of μ^* is $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$.*

We find an interesting property of this system in simulations. As N becomes large, the uniform state $\mathbf{A}$ becomes much more likely. We do not yet have a theoretical characterization of this property.

6 Conclusion

The presence of imitation, even infrequent imitation, can significantly affect the qualitative conclusions of evolutionary models. In this paper we model imitation as a sort of departure from rationality, whereby individuals make the mistake of imitating someone, even though rationality dictates otherwise. There are two slightly different ways of thinking about imitation and its effects here. First, we may view imitation as a cause of state-dependent variation in mutation rates — individuals still make random mistakes, but they may also make the mistake of imitating. We then obtain results on long-run outcomes as both types of error vanish. In a second approach, we can think of imitation as a phenomenon that modifies the unperturbed dynamics of the system. Here, we examine long-run outcomes as the probability of random mistakes vanishes (but allow positive amounts of imitation to persist). Surprisingly, this is often sufficient for selection of a unique equilibrium. This is evident from Theorem 1 — whenever a game has exactly one absorbing edge, we get convergence to a distribution that places all its mass on the absorbing edge. If both edges absorb, both have positive probability (and all equilibria outside of the set of edges have zero probability). A monotone game where both edges absorb is a coordination game, in which case we get the 1/2-dominance result as the imitation probability vanishes. For games in which neither edge absorbs, we obtain a limit distribution that puts positive probability on all states, even when there is a unique equilibrium. Here also, in the monotone case, we get convergence to a distribution that puts all of its mass on the unique equilibrium as the mistake probability goes to zero. A striking result is that, often, the limit as the imitation probability vanishes will be quite different from the limit of the no imitation (KMRY) case. The location game and club economy illustrate this, and the phenomenon occurs quite broadly for non-monotone games.

References

[1] Apsteguia, J., S. Huck and J. Oechssler (2003) Imitation – Theory and Experimental Evidence, Bonn Econ Discussion Papers (available at www.wiwi.uni-bonn.de/with2/oechssler).

[2] Bergin, J. and B. Lipman (1996) Evolution with state-dependent mutations, *Econometrica,* 64, 943–956.

[3] Blume, L. E. (1997) Population Games, In: The Economy as a Complex Evolving System II, W. B. Arthur, S. Durlauf, and D. Lane, eds., Reading, MA: Addison Wesley.

[4] Buchanan, J. (1965) An Economic Theory of Clubs, *Economica,* XXXII, 1-14.

[5] Byrne, R. W. and A.E. Russon (1999) Learning by Imitation: a hierarchical approach, *Behavioral and Brain Science*, 21,667–684.

[6] Celen, B. and S. Kariv (2004) Distinguishing Informational Cascades from Herd Behavior in the Laboratory, *American Economic Review,* 94, 484–498.

[7] Chaminade T., J. Grezes, J. Meltzoff (2002) A PET exploration of neural mechanisms involved in reciprocal imitation, *NeuroImage,* 15:265-72.

[8] Cooper, R. and A. John (1988) Coordinating coordination failures in Keynesian models, *The Quarterly Journal of Economics,* CIII, 441–464.

[9] Diamond, P. (1982) Aggregate Demand Management in Search Equilibrium, *Journal of Political Economy,* 90, 881–94.

[10] Ellison, G. (1993) Learning, Local Interaction and Coordination, *Econometrica,* 61, 1047–1072.

[11] Ellison, G. (2000) Basins of attraction, long-run stochastic stability, and the speed of step-by-step evolution, *Review of Economic Studies*, 67, 17-45.

[12] Eshel, I., L. Samuelson, and A. Shaked (1998) Altruists, Egoists, and Hooligans in a Local Interaction Model, 88, 157–179.

[13] Foster, D. and H. Peyton Young (1990) Stochastic Evolutionary Game Dynamics, *Theoretical Population Biology,* 38, 219–232.

[14] Fujita, M., P. Krugman, A. J. Venables (1999) The Spatial Economy, Cities, Regions, and International Trade, Cambridge, MA: MIT Press.

[15] Kandori, M., G. Mailath and R. Rob (1993) Learning, Mutation, and Long Run Equilibria in Games, *Econometrica,* 61, 29–56.

[16] Kandori, M. and R. Rob (1998) Bandwagon Effects and Long Run Technology Choice, *Games and Economic Behavior,* 22, 30–60.

[17] Kirman, A. (1993) Ants, Rationality, and Recruitment, *The Quarterly Journal of Economics,* CVIII, 137–156.

[18] Krugman, P. (1995) Development, Geography, and Economic Theory, Cambridge, MA: MIT press.

[19] Morris, S. (2000) Contagion, *Review of Economic Studies,* 67, 57–78.

[20] Sclag, K. H. (1998) Why Imitate, and If So, How? A Bounded Rationality Approach to Multi-armed Bandits, *Journal of Economic Theory,* 78, 130-156.

[21] Sclag, K. H. (1999) Which One Should I Imitate?, *Journal of Mathematical Economics,* 31, 493–522.

[22] Scotchmer, S. (2001) Local Public Goods and Clubs, forthcoming in Handbook of Public Economics Vol IV, 2002, A. Auerbach and M. Feldstein, eds, (North-Holland).

[23] Shapley L. and M. Shubik (1977) Trade Using One Commodity as a Means of Payment, *Journal of Political Economy*, 85, 937–968.

[24] van Damme, E. and J. Weibull (2002) Evolution in Games with Endogenous Mistake Probabilities, *Journal of Economic Theory,* 106, 296–315.

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

[26] Young, H. P. (1993) The Evolution of Conventions, *Econometrica,* 61, 57-84.

[27] Young, H. P. (1998) Individual strategy and social structure: An evolutionary theory of institutions. Princeton, NJ: Princeton University Press.

Appendix

We first gather together results of Young (1993) used in proofs of theorems. Let the unperturbed process, i.e. the process with $\varepsilon = 0$, be denoted by P^0. P^0_{xy} is then the probability of a transition from x to y ($x \to y$) in the absence of mistakes. We consider *regular* perturbations of P^0 (induced via mistakes by individual decision-makers):

1. P^ε is aperiodic and irreducible for all $\varepsilon \in (0, a]$;
2. $\lim_{\varepsilon \to 0} P^\varepsilon_{xy} = P^0_{xy}$;
3. $P^\varepsilon_{xy} > 0$ for some ε implies $\exists r \geq 0$ such that $0 < \lim_{\varepsilon \to 0} \varepsilon^{-r} P^\varepsilon_{xy} < \infty$.

Condition (3) says that a transition $x \to y$ is either impossible in P^ε or P^ε_{xy} is of order ε^r as ε becomes small. We call $r(x,y) = r$ the *resistance* of $x \to y$. The following observation is crucial:

$$r(x,y) = 0 \quad \text{if and only if} \quad P^0_{xy} > 0.$$

Young (1993) characterizes the stochastically stable states in terms of the graph G, whose vertex set is the state space X. For any vertex z of G a $z-$*tree* T is a spanning tree in G such that, for every state $x \neq z$, there exists a unique directed path from x to z. Let $\mathcal{T}_z$ be the set of all z-trees in G and define

$$\gamma(z) = \min_{\mathcal{T}_z} \sum_{(x,y) \in T} r(x,y).$$

This leads to the following result (Young, 1993, Appendix):

Lemma 1 (Young)**.** *Let P^ε be a regular perturbation of P^0 and let μ^ε be its stationary distribution. Then $\lim_{\varepsilon \to 0} \mu^\varepsilon = \mu^*$ exists and μ^* is a stationary distribution of P^0. Moreover, $\mu^*_x > 0$ if and only if $\gamma(x) \leq \gamma(y)$ for all y in X.*

We let $\gamma^* \equiv \min_z \gamma(z)$ so that $\lim_{\varepsilon \to 0} \mu^\varepsilon_z > 0$ if $\gamma(z) = \gamma^*$ and $\mu^\varepsilon_z = 0$ otherwise.

Proof of Theorem 2.

Proof. The state space is $X = \{0, 1, \ldots, N\}$ where $m \in X$ is the number of people who choose action 1 (and the rest choose 2). In the unperturbed process p^0, i.e. the process with $\varepsilon = 0$, there is positive probability that an agent who gets to revise her decision will imitate the choice of someone else in the population. As a consequence, whenever $m \neq 0$ (there is at least one person in the population who chooses 1) $P^0(m, m+1) > 0$ – there is positive probability that someone who does not choose 1 gets to make a choice, and positive probability that they imitate someone who does choose 1. Similarly, whenever $m \neq N$ we have $P^0(m, m-1) > 0$.

As a result, in the perturbed process, $r(m, m+1) = r(m, m-1) = 0$ whenever $m \notin \{0, N\}$ (irrespective of the value of $\Delta(\cdot)$). The two remaining transitions are

$r(0,1)$ and $r(N,N-1)$ and these will depend on $\Delta(\cdot)$ unless $\pi = 1$ (transitions $u \to v$ where $|u - v| > 1$ have zero probability in P^{ε}).

If $\pi = 1$, $r(0,1) = r(N,N-1) = 1$ since only one event with probability of order ε is needed for the transition. So $\gamma(0) = \gamma(N) = \gamma^* = 1$. For any $0 < m < N$, $\gamma(m) = 2$. The stationary distribution μ^* of P^0 has $\mu^*(m) > 0$ if and only if $m \in \{0, N\}$.

When $0 < \pi < 1$ the values of $\Delta(1)$ and $\Delta(N)$ will determine the values of $\gamma(\cdot)$. If $\Delta(1) < 0$ then $r(0,1) = 1$, whereas if $\Delta(1) \geq 0$ then $r(0,1) = 0$. Similarly, if $\Delta(N) < 0$ then $r(N,N-1) = 0$ whereas if $\Delta(N) \geq 0$ then $r(N,N-1) = 1$. Now there are four cases to consider:

(AA) Here $r(0,1) = 1$ and $r(N,N-1) = 1$ and $\gamma(0) = \gamma(N) = \gamma^* = 1$. The support of the stationary distribution μ^* is $\{0, N\}$.

(AR) Here $r(0,1) = 1$ and $r(N,N-1) = 0$. The support of the stationary distribution μ^* is $\{0\}$.

(RA) The support of μ^* is $\{N\}$.

(RR) Now $\gamma(m) = 0$ for all m and the support of μ^* is all of X.

□

The history y is ordered such that y_1 is the most recent choice and y_T is the choice made T periods ago. The value $y_i = 1$ indicates that a player decided to choose 1 (i periods ago); $y_i = 0$ if the player decided to choose 0. The set Y of all binary vectors of length T is the set of possible histories, and will be the state space. The only possible transitions are of the form $y \to y'$ where $y' = (0, y_1, \ldots, y_{T-1})$ or $y' = (1, y_1, \ldots, y_{T-1})$. In the unperturbed process, the random variable $\theta(y)$ is the decision in state y, leading to the transition $y \to y'$ where $y' = (\theta(y), y_1, \ldots, y_{T-1})$. Note that $\theta(y)$ is related to $\theta(y^k)$, defined in section 5.2, by the fact that y^k is a random sample of length k from y.

Proof of Theorem 3.

Proof. (1) Let $y = (y_1, \ldots, y_T)$ denote the history and y^k a random sample of length k from y. If there exists a y^k such that $q(y^k) > 0$ there is a path from y to u with zero resistance (the sum of resistance of all edges in the path is zero). If there exists y^k such that $q(y^k) < 1$, there is a path from y to z with zero resistance.

Starting at z, we need at least one transitions of $O(\varepsilon)$ before we reach a state from which there is a zero resistance path to u. Consequently, any u-tree must have resistance at least one. In fact, we can show that $\gamma(u) = 1$. We construct a u-tree as follows: from any $x \neq u$ build a path to u consisting of transitions $x \to x'$ where $x' = (1, x_1, \ldots, x_{T-1})$. The total resistance of this u-tree is one. By a similar argument, $\gamma(z) = 1$.

Now consider some $x \notin \{z, u\}$. Any path originating from z must have resistance at least one. Any path originating from u must have resistance at least one. Any x-tree must have resistance at least two.

We have $\gamma(u) = \gamma(z) = \gamma^* = 1$, and an application of Young's lemma proves the result.

The proof of (2) is similar to that of (1), and we outline it briefly. Adding the best response dynamic cannot make a difference to the resistance of any transition $x \to x'$ where $x \notin \{z, u\}$. The best response dynamic can change the resistance of transitions $z \to (1, 0, \ldots, 0)$ and $u \to (0, 1, \ldots, 1)$. Then only the edge conditions matter, and lead to the familiar results. □

Proof of Theorem 4.

Proof. $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ are absorbing states of both the imitation and best response processes. If $(x \to x')$ has positive probability in the (unperturbed) best response process, it also has positive probability in the (unperturbed) imitation process. However, additional transitions also have positive probability in the imitation process. So we can focus exclusively on the imitation process. We will apply Lemma 1 from the Appendix.

Now consider the perturbed process. The following *paths* can be traversed with one $O(\varepsilon)$ event: $\mathbf{A} \to \mathbf{C}$, $\mathbf{B} \to \mathbf{C}$, $\mathbf{C} \to \mathbf{A}$ and $\mathbf{C} \to \mathbf{B}$. Paths $\mathbf{A} \to \mathbf{B}$ and $\mathbf{B} \to \mathbf{A}$ can be traversed with two $O(\varepsilon)$ events. Since there are zero resistance paths from any $x \notin \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$ to one of the absorbing states we must have $\gamma(\mathbf{A}) = \gamma(\mathbf{B}) = \gamma(\mathbf{C}) = 2$. Further, for $x \notin \{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$, $\gamma(x) \geq 3$, so that $\gamma^* = 2$. The support of μ^* is $\{\mathbf{A}, \mathbf{B}, \mathbf{C}\}$. □

Figure 1. Payoff differences in the coordination game

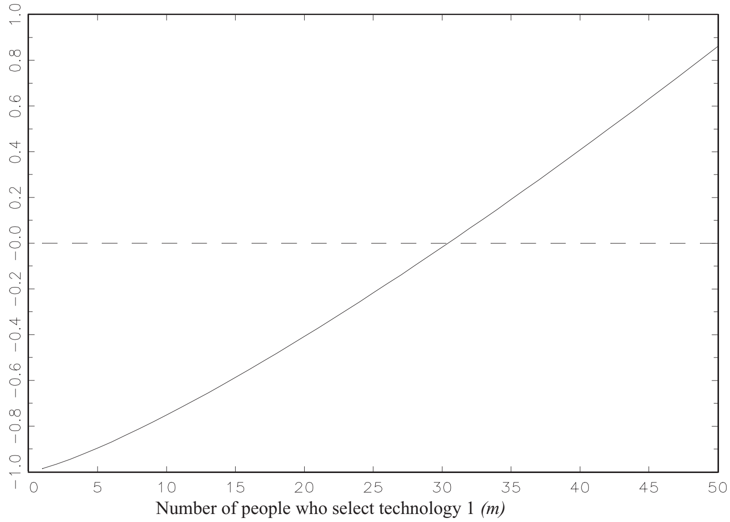

Figure 2. Payoff differences in the club model

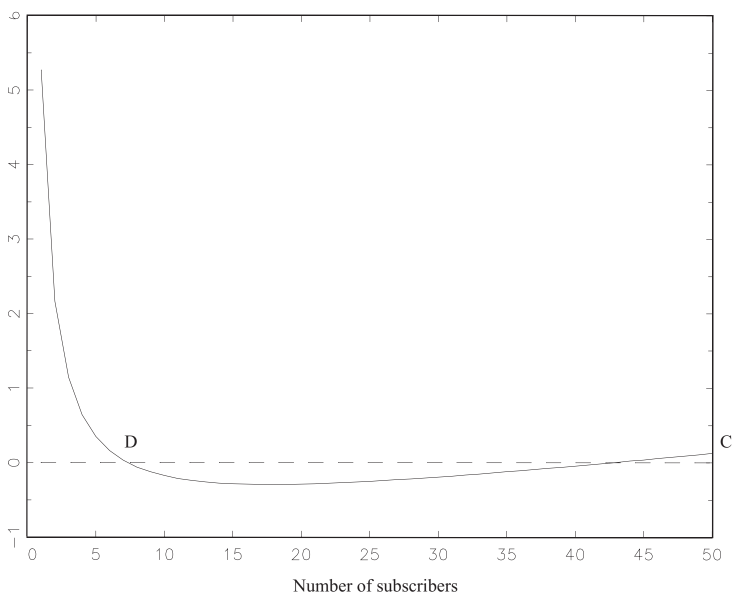

Figure 3. Payoff differences for the market game

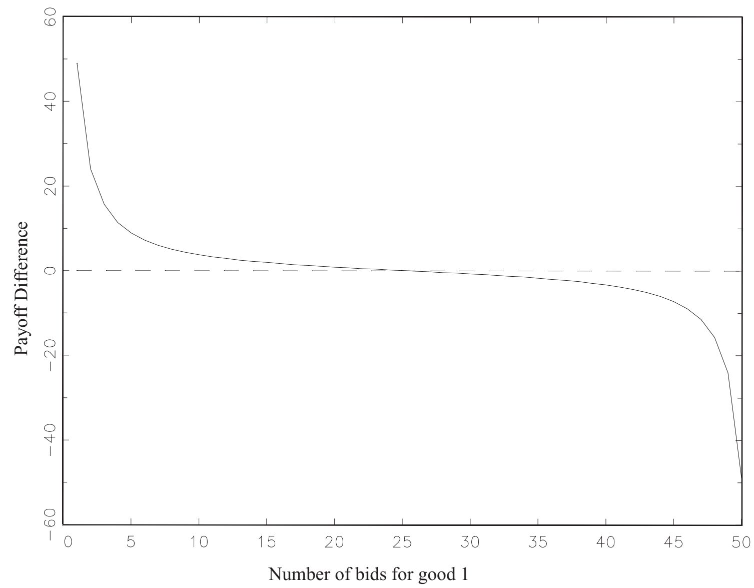

Figure 4. Market Game: Price Distribution

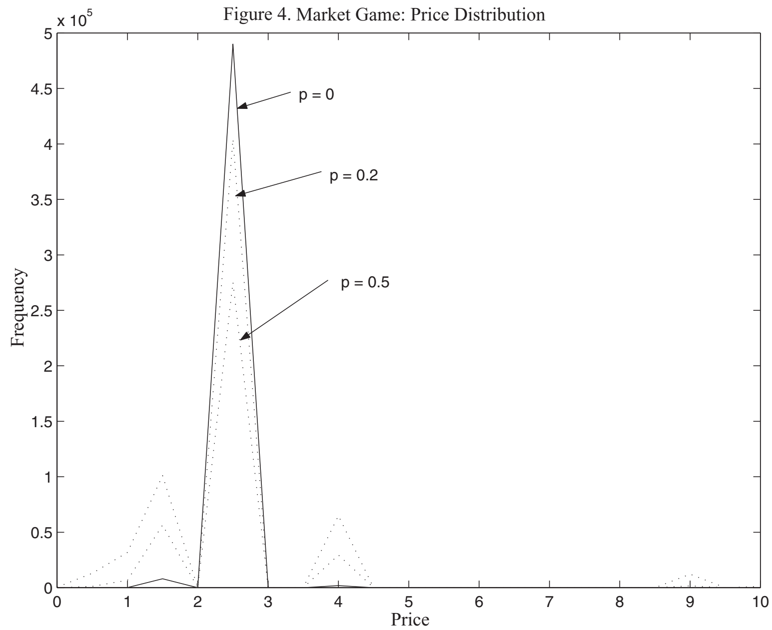

Figure 5. Price distribution in the limit (large N)

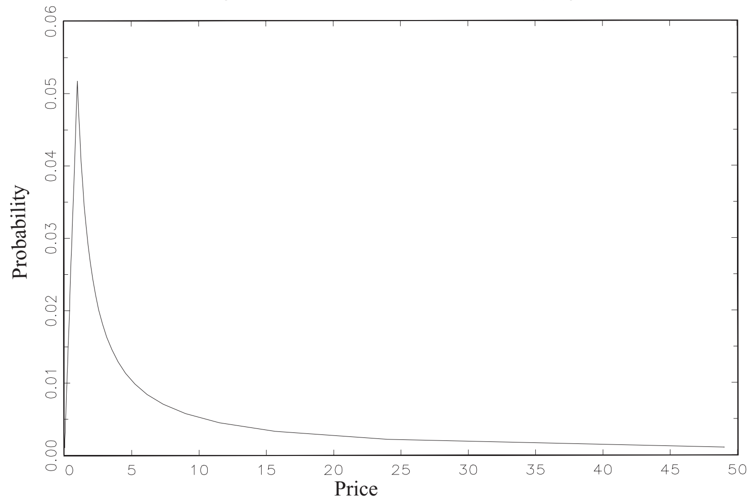

Figure 6. Conditional expectation of price

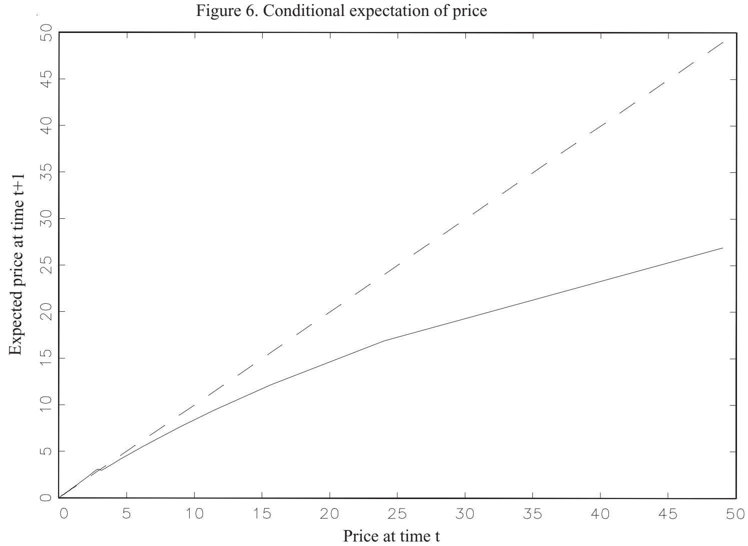

Figure 7. Conditional Variance

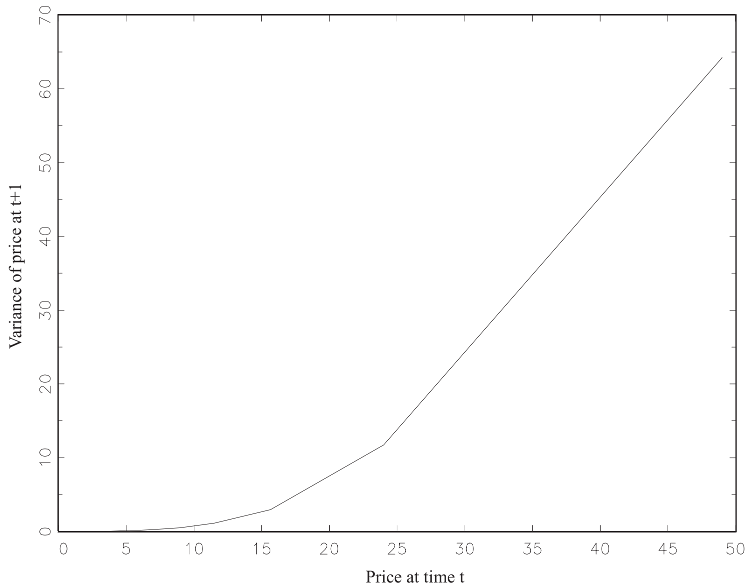

Figure 8. Payoff differences in model of location with pricing

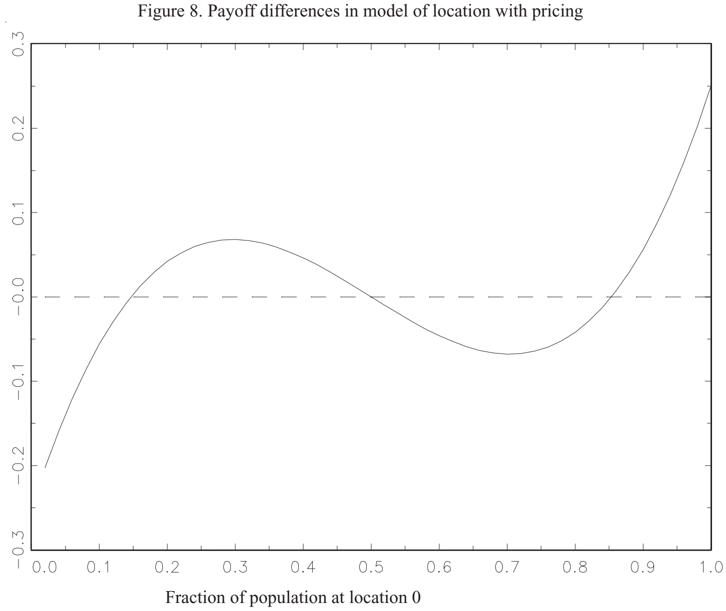

Figure 9. Distribution of location choices