

The Evolution of Cooperation Through Imitation¹

David K. Levine and Wolfgang Pesendorfer²

First version: September 29, 1999

This version: March 23, 2005

Abstract: We study evolutionarily stable outcomes for a class of games that admit cooperation and conflict as possible Nash equilibria. We make use of two ideas: existing strategies are more likely to be imitated than new strategies are to be introduced; players are able to identify opponents' behavior prior to interaction. The long-run evolutionary limit is efficient for the case of perfect recognition of opponents' behavior. For the case of imperfect recognition, efficiency is not achieved and long-run outcomes are more efficient the more accurate is the information. Strategies that emerge in the long-run are those where players reward opponents who are likely to play the same way, and punish opponents who are likely to play differently.

¹ We are grateful for the support of the National Science Foundation under Grants SES-99-86170, SES-0236882, SES-9911177. We would like to thank seminar participants at Stanford, the ITAM Theory Workshop, the Southern California Theory Conference, the Central Bank of Uruguay, and the Behavior, Evolution and Culture Series in the UCLA Anthropology Department. We owe particular debts of gratitude to Philip Johnson (who helped get us started on this line of research), to Pedro Dal Bo (for pointing out numerous mistakes and typos, and correcting an erroneous proof), and to Richard Scheelings (for careful proofreading).

² Departments of Economics, UCLA and Princeton.

1. Introduction

In this paper, we study games that admit both cooperative behavior and opportunistic behavior as possible Nash equilibria. We utilize a model of evolutionary learning to determine when one outcome emerges rather than the other.

Our evolutionary model assumes that players learn from other players through a process of imitation: Typically, when a player changes his behavior he will choose the best strategy used by other players. Sometimes the player will imitate some other player regardless of how successful this player is. In addition, there is a small probability that the player innovates, that is, introduces a strategy currently not in use by any other player. We assume that an individual is far more likely to imitate an existing, possibly sub-optimal, strategy than to adopt a strategy that is not in use by any other player.

Players are matched into pairs to play a symmetric game. We consider both global matching, and the local interaction model of Ellison [1993]. Unlike standard pairwise matching games, we assume that prior to choosing an action each player receives a signal that is informative about the opponent's strategy. A strategy for a player is a map from possible signals to actions. Players must choose a strategy prior to observing the signal and the chosen strategies determine the probability distribution over signals. An interpretation of this recognition-technology is as a model for the detection of 'rule-of-thumb' decision-making: players are committed to a particular rule-of-thumb because it is too costly to change behavior in any particular match, and furthermore, players can (perfectly or imperfectly) identify opponents' rules-of-thumb either because (say) past interactions are observable or else (say) because such rules manifest themselves observationally through language or involuntary gestures at the time of the match (blushing when telling a lie).

Consider the following example. Suppose the underlying game is a Prisoner's Dilemma. Further suppose that there is a signal that identifies when both players use the same strategy: that is, both players receive the signal "*same*" if and only if they choose the same strategy. Consider the strategy that takes a cooperative action if the player observes the "*same*" signal and a punishment action otherwise. Clearly, it is optimal for a player to adopt this strategy when he expects his opponents to use it. On the other hand, consider the strategy that defects regardless of signal. Again, this rule is optimal

whenever an opponent is expected to use it. With a sufficiently rich information structure these games typically have multiple equilibria that resemble the folk theorem of repeated games. Our evolutionary model of imitation allows us to characterize the behavior that emerges in the long-run.

When it is possible to identify the opponent's behavior without error, we show that the long-run equilibrium is efficient. If we observe the system through time, we may also observe brief bursts of conflict in which players attempt to maximize the difference between their own payoffs and that of players using an opposing strategy.

We also study, albeit in a more restricted set of environments, the long-run outcome when identification of opposing strategies is imperfect. Here we show that in the long-run strategies that are informationally dominant emerge. Informational imperfections lead players to place too much weight on their own self-interest relative to both the common good and the punishment of opponents. In particular, if information is sufficiently poor, the static Nash equilibrium of the underlying game emerges as the unique long-run outcome. There are a variety of intermediate cases in which the intensity of cooperation and conflict depends on how reliably the opponent's strategy can be identified. In some circumstances, the unique long-run equilibrium may actually be worse than the static Nash equilibrium. In all cases, behavior may be interpreted as the maximization of an altruistic/spiteful utility function, where the degree of altruism or spite depends on how likely the opponent is to be the same type.

The idea that players may receive information about their opponent's behavior prior to choosing an action appears in Frank [1987]. There is a substantial literature on evolution with positive assortative matching, where players are more likely to be matched with others who are playing the same strategies (see Bergstrom [2002]).

Both Robson [1990] and Kim and Sobel [1995] consider a setting where players can send cheap-talk messages prior to choosing an action. Kim and Sobel [1995] show that in pure coordination games with pre-play communication only the efficient equilibrium emerges as the stable outcome of their evolutionary dynamic. Kim and Sobel [1995] also show that for general games there is a large set of evolutionarily stable outcomes. Models based on cheap-talk communication face the difficulty that defectors can mimic the message of a cooperative strategy at no cost. As a consequence, it becomes relatively easy to destabilize cooperative behavior. Binmore and Samuelson

[1990] analyze evolutionarily stable equilibria played by finite automata. In their setting, automata can use initial moves to identify each other. They provide an equilibrium refinement that ensures that efficient equilibria emerge in a repeated games setting. But as in the case of cheap talk communication, these initial moves of a cooperative machine are easy to imitate by selfish deviators. Therefore, in our evolutionary framework the approach of Binmore and Samuelson [1990] would run into the same difficulty as the cheap talk model.

Pre-play communication in our model is not cheap talk. Players cannot imitate the signals of other players without imitating their behavior. This makes it easier to sustain cooperative behavior in our model. As a result, we are able to pin down the long-run stable outcomes in more complicated games than, for example, the pure coordination games studied by Kim and Sobel [1995]. In section 6, we briefly consider the case where strategies can “masquerade” as cooperators while playing a selfish action. Whether such a free-riding strategy emerges depends on how costly it is to imitate the signals generated by other strategies.

Our evolutionary model stems from existing work on evolution in economic systems. In two influential papers Kandori, Mailath and Rob [1993] and Young [1993] showed how introducing random innovations (mutations) into a model of evolutionary adjustment enables predictions about which of several strict Nash equilibria will occur in the very long run. Key to this result is the possibility that strategies that perform poorly may be introduced into the population in sufficient numbers through innovation that they begin to perform well. Using this method, they and other researchers have been able to characterize when cooperation will emerge in coordination games using the criterion of risk dominance. Ellison’s [1993] local interaction model gives a plausible account of how this evolutionary adjustment can take place over sensible amounts of time.

In this paper, we give a different account of the spread of new strategies. Once a player introduces a new strategy, a process of imitation propagates the innovation. We think this is a plausible account of how new ideas spread. This modified propagation mechanism also makes it easier to find long-run equilibria. First, the long-run limit contains only pure strategies. Second, it is sufficient that a strategy profile beat all others in pair-wise contests. As we illustrate through examples, this is implied by, but weaker than, the criterion of $\frac{1}{2}$ -dominance proposed by Morris, Rob and Shin [1993].

There are several other papers that analyze models of imitation and mutation. Dutta and Prasad [2004] analyze a model that is similar to ours in that it considers both imitation and mutation of strategies.³ Dutta and Prasad assume that players choose an unconstrained best response (when they do not imitate or make errors) while our model assumes that players choose a best response from the set of strategies used by other players (a relative best response). More importantly, Dutta and Prasad focus on games with two strategies while our focus is on games with relatively complex strategy spaces. Vega-Redondo [1997] analyzes a stochastic process where players imitate the best strategy of opponents and occasionally mutate to a random strategy. Our model essentially takes the evolutionary process of Vega-Redondo [1997] and adds to it random imitation. The addition of random imitation makes it possible to find a simple characterization of the limiting distribution. Eshel, Samuelson and Shaked [1998] consider a model where players imitate neighboring players who earn a high payoff. They consider a game with two strategies and show that cooperative behavior will emerge in the long-run evolutionary equilibrium. In contrast to our model, players in their model do not play a relative best response and the result is sensitive to the form of interaction. Schlag [1998] considers a multi-armed bandit setting and asks which rule from a given class of imitation rules maximizes welfare.

Dealing with stochastic evolutionary models more broadly, Bergin and Lipman [1994] show that the relative probabilities of different types of noise can make an enormous difference in long-run equilibrium; here our process of imitation gives a particular theory of how those relative probabilities are determined. Van Damme and Weibull [1998] study a model in which it is costly to reduce errors, and show that the standard 2x2 results on risk dominance go through. Johnson, Pesendorfer and Levine [2000] show how the standard theory can predict the emergence of cooperation in a trading game with information systems of the type introduced by Kandori [1992b]. By way of contrast, the modified model of evolutionary learning presented here allows us to study more complex games. Finally, Kandori and Rob [1993] have a model in which winning pair-wise contests is sufficient for a strategy to be the long-run outcome.

³ Kirman [1993] also considers a model of random imitation and mutation. Kirman's model differs from other evolutionary models discussed here in that intentional behavior plays no role.

However, in their model winning all pair-wise contests implies 1/2-dominance, which is not true in our applications.

Our model of imitation has a great deal of similarity to the replicator dynamic in the sense that strategies that are used are more likely to be imitated than those that are not. In particular, when there is no noise in the system, all pure strategy profiles are absorbing, just as they are in the replicator dynamic. For this reason also our results have even more of the flavor of Evolutionary Stable Strategies than is the case with the usual stochastic model of mutation. A good resource for results on the replicator dynamics and Evolutionary Stable Strategies, including the relevant case of cheap talk, is Weibull [1995].

2. The Model and Basic Characterization of Long Run Outcomes

In this section, we develop the basic characterization of long-run outcomes. The subsequent section applies this result to the game discussed in the introduction.

We study a normal form game played by m players. All players have access to the same set of pure strategies S^0 with generic element s . We assume that S^0 is finite and denote a profile of pure strategies by $\sigma \in (S^0)^m$. We assume that individual players play pure strategies from S^0 . The utility of player i depends on his strategy s^i and the profile σ and is written $u^i(s^i, \sigma)$.

The support of the profile σ , denoted $\text{supp}(\sigma)$, is the set of all those strategies which are used by at least one player (and thus excluding those strategies not used by anyone in the population). We write $\sigma(s)$ for the profile in which all players use the strategy s and σ_t for the profile at time t . Starting with an initial profile σ_0 , the profile σ_t is determined from σ_{t-1} according to the following process of imitation and innovation.

- 1) One player i is chosen at random from the population of m players. This player only will be changing his strategy.
- 2) With probability $C\varepsilon$ player i chooses from $\text{supp}(\sigma_{t-1})$ randomly with probabilities $p^i(s | \sigma_{t-1})$. This is called *imitation*.
- 3) With probability ε^n player i chooses each strategy from S^0 with probabilities $q(s) > 0$. This is called *innovation*: strategies are chosen probabilistically regardless of how widely used they are (if used at all), or how successful they are.

- 4) With probability $1 - C\varepsilon - \varepsilon^n$ player i randomizes with equal probability⁴ among the strategies that solve

$$\max_{s \in \text{supp}(\sigma_{t-1})} u^i(s, \sigma_{t-1}).$$

This is called a *relative best response*: it is the best response among those strategies that are actually used by the particular population.

When $C = 0$ our evolutionary model coincides with the model of Vega-Redondo (1997). (In that paper the ‘relative best response’ is called ‘imitation’.) Our model is also closely related to the model of Dutta and Prasad [2004]. The difference is in the specification of the best response (item (4)). We assume that players choose a relative best response while their model assumes that players choose an unconstrained best response.

The process defined above gives rise to a Markov process M on the state space $(S^0)^m$. Because innovation has strictly positive probability, the process M is positively recurrent, and so has a unique invariant distribution μ^ε . We analyze this process as ε goes to zero. Note that when $\varepsilon \rightarrow 0$ both the probability of innovation and the probability of imitation converge to zero. However, for $n > 1$ the probability of innovation converges to zero at a faster rate than the probability of imitation.

We make two assumptions about imitation. First, we assume that n is larger than m , the number of players.

Unlikely Innovation: $n > m$.

The second assumption considers profiles in which players are using different strategies. It requires that for any strategy s in a profile there is a strictly positive chance of a player who was not playing s in that profile imitating s .

Connected Imitation *If $s \in \text{supp}(\sigma)$ and $\text{supp}(\sigma)$ is not a singleton then there is a player j such that $s^j \neq s$ and $p^j(s | \sigma) > 0$.*

⁴ This assumption is to ensure that in degenerate games in which ties are generic all maximizers have a chance of being played. With a finite population and generic utility, ties do not occur, and the support will generally be a singleton.

Consider a strategy profile σ with the strategy s and at least one other strategy in its support. Connected imitation implies that there is some agent who does not play s and who has a strictly positive probability of switching to s . The probability of such a switch is at least $C'\varepsilon$ for some strictly positive constant C' . Repeated application of the assumption of connected imitation therefore implies that the probability that *all* players who do not play s switch to s is at least $(C'\varepsilon)^{m-1}$. Since $n > m$ it follows that for small ε this is much more likely than the event that one single player innovates.

These assumptions are maintained throughout the paper.

We view as plausible the idea that imitation is more likely than innovation. The specific assumption of unlikely innovation is designed to yield sharp theoretical results, and not an assumption we would defend on empirical grounds. The assumption of connected imitation requires comment. Different players, although having access to the same set of strategies, have different utility functions, and may feel quite differently about those strategies. Hence, connected imitation in effect says that these differences are small enough so that imitation remains plausible. In our application below, the underlying symmetry of the situation means that a strategy that makes sense for one player makes sense for someone else. Further note that strategies may condition on a player's role, for example, cooperate if player 1; defect if player 2. Hence the same strategy may be sensible to imitate for players with different roles.

The possibility of imitation and the fact that imitation is much more likely than innovation distinguishes our model from the model of Kandori, Mailath and Rob [1993], or Young [1993]. Results similar to the ones obtained by those authors would hold in our model when $C = 0$ and when the population is large.⁵

We first establish a basic result about the limit distribution as $\varepsilon \rightarrow 0$. Unlikely innovation implies that heterogeneity is less stable than homogeneity. A heterogeneous population can evolve into a homogeneous population purely through imitation, while a homogeneous population cannot evolve at all without at least one innovation. Theorem 1 confirms this intuition by showing that the limit invariant distribution places weight only on homogeneous populations.

⁵ For a large population the difference between the relative best response analyzed here and a true best response is typically insignificant since a bounded number of innovations ensures that all strategies are played. Whenever $\text{supp}(\sigma_{t-1}) = S^0$ the relative best response is of course a true best response.

Theorem 1: The limit $\mu = \lim_{\varepsilon \rightarrow 0} \mu^\varepsilon$ exists and $\mu(\sigma) > 0$ implies that $\sigma = \sigma(s)$ for some $s \in S$.⁶

Proof of this, and all results, can be found in the Appendix. We refer to states σ with $\mu(\sigma) > 0$ as stochastically stable states. Theorem 1 says that all players use the same strategy in every stochastically stable state. Hence, our model of imitation leads to uniformity in the behavior of players. Although imitation is random, the combination of random imitation and relative best response ensures uniform behavior in all states that are stochastically stable.

To get some intuition for Theorem 1 note that by random imitation the Markov process will occasionally arrive at states where all players choose the same strategy. Without innovation such states are absorbing because neither imitation nor the relative best response can introduce a new strategy to the population. Our assumption “*Unlikely Imitation*” ensures that new strategies are introduced sufficiently infrequently so that only states where all players choose the same strategy can be stochastically stable.

This greatly simplifies the analysis and is the key to our characterization results below.

3. Pair-wise Contests in Matching Games

We now specialize to consider games with an underlying symmetric structure induced by matching players. We allow both for global matching procedures, and for local interaction of the type described by Ellison [1993].

Specifically, we suppose an underlying two person utility function $u(s^i, s^j)$. In the global matching model, the normal form utility function is

$$u^i(s^i, \sigma) = (1/m) \sum_{j=1}^m u(s^i, s^j).$$

This is the single population matching model that is common in the literature,⁷ and like most of the literature, we simplify by assuming that a player is as likely to be matched with himself as with another player.⁸ Notice that all players are *a priori* identical, so the

⁶ A similar result in the context of genetic algorithms may be found in Dawid [1999].

⁷ See Hahn [1995] for an extension of the standard model to multiple populations. Friedman [1998] considers a model in which players are sometimes matched with opponents from the same population and sometimes with opponents from a different population.

⁸ Our results would also hold if we assumed that players cannot be matched with themselves. In that case, we would assume an odd number of players.

assumption of connected imitation makes sense. We refer to this model as the *global matching model*. In the global matching model, we assume that m is an even number.

We also consider Ellison's [1993] *local interaction model*. Here it is assumed that players are arranged on the circle, so that player $m + 1$ is player 1, and that each player is matched with k neighbors on either side. That is,

$$u^i(s^i, \sigma) = (1/k) \sum_{j=i-k}^{i-1} u(s^i, s^j) + (1/k) \sum_{j=i+1}^{i+k} u(s^i, s^j).$$

In the local interaction model, we generally expect the population to be relatively large compared to the size of neighborhoods. It simplifies proofs considerably, if make an explicit assumption to this effect, and hereafter we assume $m > (2k + 1)^2$.

Next, we define what it means for one strategy to *beat* another strategy. For $0 \leq \alpha \leq 1$, the mixed strategy that plays s with probability α and s' with probability $1 - \alpha$ is denoted by $\alpha s + (1 - \alpha)s'$.

Definition 1: *The strategy s beats the strategy s' iff*

$$u(s, \alpha s + (1 - \alpha)s') - u(s', \alpha s + (1 - \alpha)s') > 0$$

for all $1/2 \leq \alpha < 1$.

Thus, a strategy s beats s' if s yields higher utility against any combination of s and s' that puts more weight on s than on s' . Our main characterization result (Theorem 2) will show that a strategy is the unique stochastically stable state if it beats all other strategies.

The next definition weakens Definition 1 to allow for ties.

Definition 2: *The strategy s weakly beats s' iff*

$$u(s, \alpha s + (1 - \alpha)s') - u(s', \alpha s + (1 - \alpha)s') > 0$$

for all $1/2 < \alpha < 1$ and $u(s, \frac{1}{2}s + \frac{1}{2}s') - u(s', \frac{1}{2}s + \frac{1}{2}s') = 0$. *The strategy s is tied with s' iff*

$$u(s, \alpha s + (1 - \alpha)s') - u(s', \alpha s + (1 - \alpha)s') = 0$$

for all $0 \leq \alpha \leq 1$.

We say that a strategy “*beats the field*” if it beats every other strategy.

Definition 3: *If s beats all $s' \neq s$ we say that s beats the field. If for all $s' \neq s$ either s weakly beats s' or is tied with s' we say that s weakly beats the field.*

Theorem 2 shows that a strategy that beats the field is the unique stochastically stable state. Moreover, if a strategy weakly beats the field then μ must place strictly positive probability on that strategy.

Theorem 2: *In the global matching or local interaction model, if s beats the field then $\mu(\sigma(s)) = 1$. If s weakly beats the field then $\mu(\sigma(s)) > 0$. Moreover, if $\mu(\sigma(s')) > 0$ then $u(s, \frac{1}{2}s + \frac{1}{2}s') = u(s', \frac{1}{2}s + \frac{1}{2}s')$.*

To gain an intuition for Theorem 2, consider a pair of strategies, s, s' . For small ε we can give an approximate upper bound for the probability that the Markov process moves between $\sigma(s)$ and $\sigma(s')$. Since players choose a relative best response with probability close to one, exactly one innovation and enough imitations are required so that the system will move to $\sigma(s')$ in the relative best response dynamic. Otherwise, the Markov process simply returns to $\sigma(s)$. Suppose that one innovation and at least r imitations are required to move from $\sigma(s)$ to $\sigma(s')$ and suppose that one innovation and no more than $r' < r$ imitations are required to move from $\sigma(s')$ to $\sigma(s)$. Thus, for small ε , the approximate relative likelihood of a transition from $\sigma(s)$ to $\sigma(s')$ is

$$\varepsilon^n (C\varepsilon)^r$$

Similarly, the approximate relative likelihood of a transition from $\sigma(s')$ to $\sigma(s)$ is

$$\varepsilon^n (C\varepsilon)^{r'}$$

Thus, the ratio of the two probabilities is

$$\frac{\varepsilon^n (C\varepsilon)^r}{\varepsilon^n (C\varepsilon)^{r'}} = (C\varepsilon)^{r-r'}$$

Hence, when $r' < r$ and ε is small, we conclude that a transition from $\sigma(s)$ to $\sigma(s')$ is far less likely than a transition from $\sigma(s')$ to $\sigma(s)$.⁹

⁹ It is known that these types of arguments can not generally be used to prove theorems about the stochastically stable state.

The role of the hypothesis that s beats the field is to insure that $r > r'$ for all s' . The reason that this is true is different in the two different global and local matching cases. In the global matching case, since m is an even number, from the definition of beating the field, it follows that $r \geq m/2$, while $r' \leq m/2 - 1$, from which it is apparent that $r > r'$. In the local matching case, suppose the configuration is such that all players $\ell = L, L + 1, \dots, L + 2k$ are all playing s . In this case we say that σ has *contiguous* s and refer to the players $\ell = L, L + 1, \dots, L + 2k$ as the contiguous set of s . When s beats the field and σ has contiguous s then σ converges to $\sigma(s)$ in the relative best response dynamic. This follows because for players in a contiguous set of s it is never a relative best response to switch away from s , and for players on the border of a contiguous set it is always a relative best response to switch to s . Starting from $\sigma(s')$ it is therefore sufficient to have one innovation to s followed by $2k$ imitations to s to reach $\sigma(s)$ and hence $r' \leq 2k$. Starting from $\sigma(s)$ we must eliminate all contiguous sets of s to move to $\sigma(s')$. This requires one innovation to s' and at least $m/(2k + 1) - 1$ imitations to s' . Therefore, $r \geq (m/(2k + 1)) - 1$. Since we are assuming that $m > (2k + 1)^2$, it follows again that $r' > r$.

Notice that the amount of time it takes to get from s' to s is inversely related to the probability. In the local interaction case, this is of order ε^{-2k} ; in the global case this is of order $\varepsilon^{-(m/2)-1}$. If the population size is large and ε is small then convergence in the global case is very slow. By contrast, regardless of the population size, convergence in the local case will be relatively rapid if the number of opponents interacted with is small. It is for this reason that Ellison [1993] argues that local interaction is a more plausible model of the spread of ideas through a population.

The hypothesis, that a strategy ‘beats the field’ relates to the idea of $1/2$ -dominance introduced by Morris, Rob and Shin [1993]. The concept of $1/2$ -dominance says that when half or more of the population is playing s against any other combination of strategies, it is a best response to play s . The concept here is weaker in two respects: first, s must only beat homogeneous opponents, not arbitrary combinations of strategies. Second, s must win only in the sense of being a relative best-response, it need not actually be a best-response; a third strategy may actually do better than s , and this is significant as we will see in examples below. On the other hand, $1/2$ -dominance clearly implies winning all pair-wise contests. So if there is a $1/2$ -dominant strategy, we know from Morris, Rob and

Shin [1993] that it is stochastically stable with respect to the usual evolutionary dynamic, and that it is also stochastically stable when innovation is unlikely. It should also be noted that unlike Morris, Rob and Shin [1993], we consider the local case as well as the global case.

4. Matching Games with Behavioral Types

In this section we apply the just-obtained characterization results to a pair-wise matching game. Every period, players are matched into pairs to play a symmetric normal form game. Prior to choosing an action, each player receives a “signal” containing information about how his opponent will behave in the game. We examine how the long-run outcome depends upon the amount of information contained in the signals.

The underlying game played in each match is symmetric. The action space for both players is A and the payoff of a player who takes action a and whose opponent takes action a' is $U(a, a')$.

Prior to meeting an opponent, players simultaneously choose a strategy from a finite set. Strategies serve two roles. First, they influence the information that is generated about the player and his opponent; and second, they govern behavior as a function of the generated information. Formally, let Y denote a finite set of signals. A strategy for a player is a map $s : Y \rightarrow A$ and S denotes the set of strategies. If a player chooses $s \in S$ and his opponent chooses $s' \in S$ then the player receives signal y with probability $\pi(y | s, s')$. Signals are private information. In our interpretation, a signal reflects what the opponent can learn about the player’s behavior prior to the interaction.

We motivate this choice of ‘recognition technology’ as two ways. Strategies govern the behavior of agents over many matches. Players are committed to a particular strategy because it is too costly to change behavior in any particular match. If a player could observe the past interactions of an upcoming opponent he, would be able to form a prediction of how that opponent would behave during their own upcoming match. Alternatively, it could be assumed that strategies are directly observable. For example, an individual who rarely lies may blush whenever he is dishonest. Seeing an opponent blush would indicate that he would be unlikely to be dishonest in future interactions. (This example is due to Frank (1987)).

As an example, suppose that $Y = \{0,1\}$. Further assume that $\pi(y = 0 \mid s', s) = 1$ if $s' = s$ and $\pi(y = 1 \mid s', s) = 1$ if $s' \neq s$. Thus, if two players meet who use the same strategy then both receive the signal 0 whereas when two players meet who use different strategies then both receive the signal 1. In other words, players recognize if their opponents use the same or a different strategy prior to play. This example is important, because it turns out that strategies that recognize themselves are likely to emerge in the long-run equilibrium.

It bears emphasis that the space of signals is necessarily smaller than the set of strategies: the cardinality of the space of strategies is at least A^Y , which is greater than that of Y provided that there are at least two actions. This relative coarseness of the signal space means that it is not possible that a signal could reveal the precise strategy of an opponent for every possible strategy profile.

If player i uses strategy s^i and his opponent uses strategy s^j then the expected payoff of player i is given by

$$u(s^i, s^j) = \sum_{y \in Y} \sum_{y' \in Y} U(s^i(y), s^j(y')) \pi(y \mid s^i, s^j) \pi(y' \mid s^j, s^i).$$

We consider three scenarios. In the first – analyzed in section 4.1 – a player is able to determine with certainty whether his opponents in each match use the same strategy he does. For this case, we show for a general class of utility functions that an efficient equilibrium will emerge as the long run outcome. In the second scenario – analyzed in section 4.2 – we consider the case of noisy identification of opponents' behavior for a restricted class of utility functions. In that restricted environment, we relate the degree of cooperation among agents to the ability of agents to identify types who use similar rules. The model of section 4.2 supposes a great deal of symmetry in the signaling mechanism. The third scenario – analyzed in section 4.3 – replaces the symmetry assumption with an informational dominance condition.

4.1 Perfect Identification

The first scenario assumes that each player can identify with certainty whether an opponent is using the same strategy.

Assumption 1: *There is a $y_0 \in Y$ such that $\pi(y_0 \mid s, s) = 1$ for every $s \in S$ and $\pi(y_0 \mid s, s') = 0$ for $s \neq s'$.*

The signal y_0 is generated if and only if the strategies of the two players coincide and therefore players can identify those opponents who behave in the same way as they do.

Before analyzing the evolutionarily stable outcomes, it is instructive to consider the pure strategy Nash equilibria of the static two-player game. In that game, each player simultaneously chooses a strategy $s \in S$ and payoffs are given by the payoff function u defined above. Let a_* denote the pure min-max action of the underlying game

$$\max_{a' \in A} U(a', a'') \geq \max_{a' \in A} U(a', a_*) \equiv u_*, \forall a'' \in A$$

Let $a \in A$ be any other action that satisfies

$$U(a, a) \geq u_*$$

Let s_{aa_*} be defined as follows:

$$s_{aa_*}(y) = \begin{cases} a & \text{if } y = y_0 \\ a_* & \text{if } y \neq y_0 \end{cases}$$

The strategy s_{aa_*} plays the action a if the opponent has chosen s_{aa_*} and the minmax action a_* if the opponent has chosen a strategy different from s_{aa_*} . It is immediate that the profile (s_{aa_*}, s_{aa_*}) constitutes a Nash equilibrium for every action $a \in A$ with $U(a, a) \geq u_*$. Hence, we have the following ‘‘folk theorem’’ for this benchmark game.

Benchmark Theorem: *If $v = U(a, a)$ for some $a \in A$ and $v \geq u_*$ then v is a Nash equilibrium payoff.*

Although the game has many Nash equilibria we will show that under our evolutionary dynamics the outcome is efficient. To facilitate the analysis, we make two assumptions on the utility function U of the underlying game. Assumption 2 requires that there be a (weakly) Pareto optimal symmetric profile. Let \bar{a} be the best symmetric outcome, i.e.,

$$\bar{a} \in \arg \max_{a \in A} U(a, a)$$

Assumption 2 says that there is no asymmetric profile that is strictly better than (\bar{a}, \bar{a}) for both players.

Assumption 2: $U(a, a') > U(\bar{a}, \bar{a}) \Rightarrow U(a', a) \leq U(\bar{a}, \bar{a})$

If there is a public randomization device then Assumption 2 is always satisfied if we include actions that may depend on the outcome of the public randomization. In that case, we can use a coin flip to decide which player is the row player and which player is the column player. Once roles are assigned, players choose the Pareto optimal actions.

Assumption 3 requires that there is an action $\underline{a} \in A$ that ensures that the player gets a payoff that is at least as large as the payoff of his opponent.

Assumption 3: *There is an $\underline{a} \in A$ such that $U(\underline{a}, a) - U(a, \underline{a}) \geq 0$ for all $a \in A$.*

Note that the payoff difference $U(a', a) - U(a, a')$ defines a symmetric zero-sum game and hence has a (possibly mixed) minmax strategy. Assumption 3 says that this minmax strategy is an element of A , that is, the game defined by the payoff differences has a pure minmax strategy. Assumption 3 is always satisfied if we include the possibly mixed minmax action as one of the elements of A .

Let $s_0 := s_{\bar{a}\underline{a}}$. Hence, the strategy s_0 is defined as

$$s_0(y) = \begin{cases} \bar{a} & \text{if } y = y_0 \\ \underline{a} & \text{if } y \neq y_0 \end{cases}$$

When the opponent plays s_0 , the strategy s_0 takes the Pareto efficient symmetric action. When the opponent does not play s_0 the strategy s_0 chooses the punishment action \underline{a} . Note that the punishment action maximizes the minimum difference between the player's payoff and his opponent's payoff. Assumption 2 implies that for every $a \in A$ $U(\bar{a}, \bar{a}) - U(a, \underline{a}) \geq 0$ and therefore action \underline{a} is indeed a punishment.

Theorem 3 shows that the long run outcome of the evolutionary dynamics will put positive probability on the strategy s_0 . Moreover, every other strategy s that is used with positive probability leads to a payoff similar to the payoff of s_0 : when s meets s both players receive the payoff $U(\bar{a}, \bar{a})$; when s meets s_0 both players receive the same payoff.

Theorem 3: $\mu(s_0) > 0$. *If $\mu(s) > 0$ then $u(s, s) = u(s_0, s_0) = U(\bar{a}, \bar{a})$ and $u(s, s_0) = u(s_0, s)$.*

Theorem 3 implies that if \bar{a} is the unique symmetric Pareto optimal outcome and if $U(\underline{a}, a) - U(a, \underline{a}) > 0$ for all $a \neq \underline{a}$ then the optimal punishment strategy is the unique outcome in the long-run limit. This will be the case in a generic game that satisfies

Assumptions 2 and 3. Otherwise, there may be two or more strategies that are long-run stable. Note that if we observe the system for a long time we will typically observe each player using the same long-run stable strategy and all players receiving the payoff $U(\bar{a}, \bar{a})$. However, occasionally there will be a transition from one long-run stable strategy to another. During this brief period of transition, players using different strategies will punish one another.

The proof of Theorem 3 in the Appendix shows that strategy s_0 weakly beats the field, that is; it weakly beats every other strategy. Let us briefly explain the role of the Assumptions in proving this. Assumption 1 guarantees that players can detect when opponents use a strategy that differs from their own. Assumptions 2 and 3 allow us to construct an optimal punishment strategy that punishes all opponents who use a different strategy with the same action. Assumption 3 implies that the action \underline{a} maximizes the payoff difference between a player and his opponent irrespective of the action chosen by the opponent. Assumption 2 implies that deviating to an asymmetric profile cannot increase everybody's payoff and therefore s_0 is a best response to itself. Without Assumption 3 the optimal punishment could depend on the particular strategy s used by other players. In that case, a long-run stable outcome will depend on the details of the information environment (Y, π) . If Assumption 2 were violated then there might be an action $a \in A$ such that $U(\bar{a}, \bar{a}) < U(a, \underline{a})$. In that case, s_0 would reward a deviation to a strategy that always plays a and – obviously – the result would fail. Assumption 2 seems natural given that our model considers single population dynamics. For games in which all Pareto optimal action profiles are asymmetric it seems more appropriate to consider an evolutionary model with two distinct populations of players, one corresponding to row players and the other corresponding to column players.

Note that the strategy s_0 need not be $\frac{1}{2}$ dominant. Suppose that the underlying game is a Prisoner's dilemma and let \tilde{s} be a constant strategy that always plays "defect". Suppose, moreover, that there are signals that enable a strategy \underline{s} to play "defect" against s_0 and "cooperate" against \tilde{s} . As defined above, s_0 plays "cooperate" against s_0 and the "defect" otherwise. Against $\frac{1}{2}s_0 + \frac{1}{2}\underline{s}$ the strategy \tilde{s} does better than s_0 and therefore s_0 is not $\frac{1}{2}$ dominant. Note that the strategy \underline{s} seems to serve no useful purpose except to make \tilde{s} look good against s_0 . Our theory of infrequent innovation provides a rigorous account of why we should not expect such strategies to play a role in

determining the long-run equilibrium: because they do not themselves do well against s_0 they will not remain around long enough for players to discover that they should play \tilde{s} .

4.2 Gift Exchange and Imperfect Identification

The long-run stable strategy identified in the previous section has a simple binary form. It uses a single cooperative action when it meets an opponent who uses the same strategy and it uses a single punishment action for opponents who use a different strategy. To obtain this simple form, it is necessary that agents be able to identify deviators (opponents who use a different strategy) without error.

Such a simple strategy is no longer optimal when deviations cannot be detected without error. In that case, the punishment or reward must take into account the likelihood of facing a deviator. In this section, we analyze the case of imperfect identification for the special case of an additively separable payoff structure.

We assume that each action a has a cost $c(a)$ and yields a benefit $b(a)$ for the opposing player. The payoff of a player who takes action a and whose opponent chooses action a' is

$$U(a, a') = b(a') - c(a) \quad (1)$$

We assume that $c(a) \geq 0$, with $c(\hat{a}) = 0$ for some action $\hat{a} \in A$. Note that the utility function (1) satisfies Assumptions 2 and 3. The payoff structure defined in (1) implies that the effect of a player's action on the opponent's payoff is independent of the action taken by the opponent. This feature of the payoffs will allow us to determine the long-run stable outcomes under our evolutionary dynamics; without it, optimal punishments will generally depend on opponents' strategies, and we do not get a clear result

We can interpret this game as describing a situation where two players meet and have an opportunity to exchange goods. The function c denotes the cost of the good to the giving player and b describes the benefit of the good for the receiving player. Games with this utility function resemble a prisoner's dilemma in that the cost minimizing action is dominant.

Assumption 4 describes the information structure analyzed in this section.

Assumption 4:

$$\pi(y | s, s') = \begin{cases} p(y) & \text{if } s = s' \\ q(y) & \text{if } s \neq s' \end{cases}$$

Thus, $p(y)$ describes the probability that a player receives the signal y if he and his opponent use the same strategy whereas $q(y)$ describes the probability when the two players use different strategies. Suppose the prior probability that a player uses s is α and that he receives the signal y . Then, the posterior probability that the opponent will play according to s is

$$\frac{\alpha p(y)}{\alpha p(y) + (1 - \alpha)q(y)}.$$

This posterior is greater than α when $p(y) > q(y)$ and less than α when $q(y) > p(y)$.

We begin by analyzing the pure strategy Nash equilibria of the static two-player game. As in the previous section, each player simultaneously chooses a strategy $s \in S$. Suppose the strategy s satisfies

$$\sum_Y p(s(y))(b(s(y)) - c(a)) \geq \sum_Y q(s(y))b(s(y)) \quad (\text{IC})$$

Inequality (IC) says that a player who plays s receives a higher payoff than a player who always plays the cost minimizing action \hat{a} . (Recall that $c(\hat{a}) = 0$). Since the information system sends the same signals for every deviation from s , a player who wishes to deviate will always do best by deviating to the cost minimizing strategy. Consequently, the inequality (IC) is necessary and sufficient for $s \in S$ to constitute a pure strategy Nash equilibrium.

Benchmark Theorem: *The strategy $s \in S$ is a pure strategy Nash equilibrium if and only if it satisfies (IC).*

The strategy that chooses the cost minimizing action for every signal is always a Nash equilibrium. If the signal is uninformative ($p = q$) then this will be the only equilibrium. If the signal is informative and the benefits of exchange are sufficiently great then there will be Nash equilibria in which gift exchange takes place.

Next, we characterize the long-run stable outcome implied by our evolutionary model. The strategy s_0 is defined as follows. For every signal y the action $s_0(y)$ solves

$$\max_{a \in A} [p(y) - q(y)]b(a) - [p(y) + q(y)]c(a) \quad (*)$$

We assume that the maximization problem (*) has a unique solution for every y . The strategy s_0 rewards opponent types when $p(y) > q(y)$ and punishes opponent types when $q(y) > p(y)$. In the limiting case where the type allows no inference about his play ($p(y) = q(y)$) the strategy s_0 minimizes the cost c .

Theorem 4, proven in the Appendix, shows that s_0 is the unique long-run outcome.

Theorem 4: $\mu(s_0) = 1$.

To prove Theorem 4 we show that s_0 weakly beats the field and that $u(s_0, \frac{1}{2}s_0 + \frac{1}{2}s) = u(s, \frac{1}{2}s_0 + \frac{1}{2}s)$ implies $s = s_0$. Therefore, we can apply Theorem 2 to get the desired conclusion. Theorem 4 uses the special structure of the gift-exchange game to find an optimal punishment strategy that “works” irrespective of the behavior of the opponent. In particular, the simple characterization of the long-run outcome depends on the fact that the impact of a player’s action on his opponent’s payoff is independent of the opponent’s action. The separability of payoffs assumed in this section plays a role similar to the role played by Assumption 3 in the previous section. Without the separable payoffs, the optimal punishments and rewards could depend on the strategy of the opponent. In that case, there may be no strategy that weakly beats the field and hence our characterization theorem could not be applied.

Under our interpretation of the environment as gift exchange game, the gift is decreasing as the signal indicates an opponent who is more likely to play a different strategy. It is worth emphasizing that in a long-run stable outcome typically all players are choosing the same strategy. Hence, the probability that a player is using s_0 is one, irrespective of the realization of the signal. Nevertheless, players will punish each other for appearing to be different. This implies that the equilibrium is inefficient. In contrast to the case of Theorem 3, where a socially efficient payoff was realized in the long-run stable outcome, here inefficiencies persist in the long-run because players respond to signals *as if* one-half of the population were using a different strategy and hence needed to be punished.

4.3 Informational Dominance

In Theorem 4 every strategy generates the same information. We now relax that assumption and consider strategies that may differ in their ability to identify the behavior

of opponents. For example, a strategy may have an advantage in determining whether an opponent uses the same strategy. Alternatively, a strategy may be good at masquerading and hence be hard to distinguish from other strategies.

Specifically, consider the optimal punishment strategy s_0 defined above. This strategy uses a symmetric information structure defined by (p, q) and therefore generates the same information for every opponent who does not use s_0 . We continue to suppose that s_0 uses this information structure, that is

$$\pi(y | s_0, s) = \begin{cases} p(y) & \text{if } s_0 = s \\ q(y) & \text{if } s_0 \neq s \end{cases}$$

However, for strategies other than s_0 we now drop Assumption 4 and allow general signal distributions $\pi(\cdot | s, \cdot)$. Nevertheless, we show that if s_0 is informationally dominant, it emerges as a long run outcome.

Consider a situation where only strategies s_0 and s are played. Strategy s_0 is *informationally superior* to strategy s if the signal generated by s_0 provides better information about the opponent's strategy than the signal generated by s . The signal generated by s_0 provides better information (in the sense of Blackwell (1954)) than the signal generated by s if there is a non-negative matrix

$$\left(\lambda_{yz} \right)_{y \in Y, z \in Y}$$

such that

$$\begin{aligned} \sum_{y \in Y} \lambda_{yz} &= 1, \forall z \\ \pi(y | s, s) &= \sum_{z \in Y} \lambda_{yz} p(z), \\ \pi(y | s, s_0) &= \sum_{z \in Y} \lambda_{yz} q(z); \end{aligned}$$

In other words, the signals generated by $\pi(\cdot | s, \cdot)$ are a *garbling* of the signals generated by s_0 .

The strategy s_0 is *informationally dominant*, if it is informationally superior to every other strategy s . Note that informational dominance only requires that strategy s_0 generates better information in situations where s_0 and *one* other competing strategy are

played. Thus, s_0 may be an informationally dominant strategy even though strategy s does better at identifying a third strategy \bar{s} .

A trivial example of an informationally dominant strategy is a strategy that cannot be distinguished from any other strategy. In that case, $\pi(y | s, s_0) = \pi(y | s, s)$ for all s and hence strategy s_0 is informationally dominant even if strategy s_0 does not generate any information, that is, $p(y) = q(y)$ for all y . This is a case where strategy s_0 is informationally dominant because it successfully masquerades as other strategies.

Theorem 5 shows that when strategy s_0 is informationally dominant, it emerges as an outcome of the long-run stable distribution. Moreover, every strategy that is a long-run stable outcome is similar to strategy s_0 in payoff. In particular, if $\mu(s) > 0$ then the payoff when s meets s is the same as the payoff when s_0 meets s_0 .

Theorem 5: *If s_0 is informationally dominant then $\mu(s_0) > 0$. Moreover, for every strategy s with $\mu(s) > 0$ we have that $u(s, s) = u(s_0, s_0)$ and that $u(s, s_0) = u(s_0, s)$.*

In this section, we have restricted the informationally dominant strategy to generate symmetric information, that is, to generate the same information for every opponent. This allowed us to identify a behavior (a map from signals to actions) that is successful against every opponent. The symmetry assumption in this section is therefore more than a convenience. It implies that strategy s_0 is informationally superior to every other strategy with a uniform interpretation of the signals. If we were to forego this symmetry assumption we would need to replace it with a requirement that would preserve this uniformity. For example, we could assume that there is a reference strategy s such that any signal realization generated by s_0 against an arbitrary opponent is at least as informative as it is against strategy s . Informational dominance would then require that the signal generated against s be informationally superior to the signal generated by any opponent.

4.4 Examples

We conclude this section by illustrating Theorems 4 and 5 in the following four examples.

Example 1: First, consider the case where every strategy uses the same symmetric information structure (p, q) and hence Theorem 4 applies. Moreover, there is a signal y_0

with the property that $p(y_0) = 1, q(y_0) = 0$, that is, players can perfectly identify if their opponents use the same strategy. In such a case Theorem 4 is of course a special case of Theorem 3. If a player uses strategy s_0 and meets a player who also uses s_0 , then both players are assigned the type y_0 . Since $p(y_0) = 1, q(y_0) = 0$, the action taken by both players solves

$$\max_{a \in A} b(a) - c(a)$$

Note that $b(a) - c(a)$ is the social benefit of action a and hence we have an efficient outcome in this case. If a player uses a strategy other than s_0 then his opponent receives the signal $y \neq y_0$. Hence $p(y) = 0$, and so s_0 punishes the player by choosing the action that solves

$$\max_{a \in A} -b(a) - c(a) = -\min_{a \in A} b(a) + c(a)$$

Note that this action maximizes the payoff difference between the two players, as required by Theorem 3. Since the punishment action minimizes the sum of a player's cost and his opponent's benefit, that player is willing to incur such a cost since it leads to a negative payoff for his non- s_0 playing opponent.

Example 2. To the environment of Example 1, we add the strategy \bar{s} , which can successfully masquerade as any other strategy. Thus, a player using strategy s cannot distinguish between opponents who use \bar{s} or s , so that $\pi(y | s, \bar{s}) = \pi(y | s, s)$ for all signals $y \in Y$. In addition, players who use \bar{s} do not receive informative signals about their opponents. Hence, we can describe their information by a symmetric information structure (\bar{p}, \bar{q}) with $\bar{p}(y) = \bar{q}(y)$. The strategy \bar{s} is informationally dominant and hence we can apply Theorem 5. Since signals are not informative it follows that \bar{s} is a long-run stable outcome if it takes the least cost action \hat{a} for every signal realization. In that case, Theorem 5 implies that every strategy that is a long-run stable outcome must play the least cost action. Hence, the introduction of a strategy that successfully masquerades as other strategies eliminates cooperation between players.

Example 3. This example serves to emphasize that s_0 need not be $\frac{1}{2}$ dominant in the ordinary sense. Consider the environment of Theorem 5 and assume that s_0 , the

informationally dominant long-run strategy, is not constant. Let \tilde{s} be a constant strategy that always plays \hat{a} . Suppose also that there are signals that enable a strategy \underline{s} to identify \tilde{s} with certainty and to choose an action that maximizes b . Otherwise, \underline{s} chooses \hat{a} . For an appropriate choice of b , the strategy \tilde{s} does better than s_0 against $\frac{1}{2}s_0 + \frac{1}{2}\underline{s}$ and therefore s_0 is not $\frac{1}{2}$ dominant.

Example 4: Consider a symmetric two-signal scenario, $Y = \{0,1\}$ and $p(0) = q(1) = p, p \geq 1/2$. If the signal is $y = 0$ then this is an indication that the two players are using the same strategy whereas if the signal is $y = 1$ it is an indication that the strategies are different. Suppose there are three actions $a \in \{-1,0,1\}$, $b(a) = \beta a, c(a) = 2a^2 + a$. This is a trading game with a cooperative action ($a = 1$), a no-trade action ($a = 0$), and a hostile action ($a = -1$). Both the hostile and the cooperative action are costly for players, whereas the no-trade action is costless. In this example, we can apply Theorem 4 and distinguish the following cases. When

$$\frac{1}{2p-1} > \beta,$$

then in the unique long-run outcome all players take the no-trade action. When

$$\beta > \frac{3}{2p-1},$$

then in the unique long-run outcome players choose the cooperative action when the signal is 0 and the hostile action when the signal is 1. And finally when

$$\frac{3}{2p-1} > \beta > \frac{1}{2p-1},$$

then in the unique long-run outcome players take the no trade action when the signal is 0 and the hostile action when the signal is 1. In this case, the long run outcome is therefore worse than the unique equilibrium of the normal form game. Players choose no trade and hostility and do not realize any of the gains from trade.

This example points out a significant non-monotonicity. Suppose that $9 > \beta > 1$. Then when $p = 1$ we are in the second, cooperative, case. When p approaches $\frac{1}{2}$ we are in the first no trade case. But for intermediate values of p , for example $p = 5/6$, we are

in the final case where efficiency is actually less than the autarkic solution that occurs when p approaches $\frac{1}{2}$.

5. The Masquerade

A standard and fundamental question about systems in which strategies are directly observable is why evolution does not favor free-riding strategies that pretend to be something they are not. In particular, our results imply that a free-riding strategy that can perfectly mimic other strategies must be the long-term winner. If free-riding is not to emerge in the very long-run then it must in some respect be costly.

A good general discussion of the issue of free-riding and selfishness can be found in Bergstrom [2002]. To ask why we do not see only free-riding masqueraders we return to the standard environment of perfect identification and focus on the special case of gift exchange games. In this environment, we have shown that the unique stochastically stable outcome is the optimal punishment strategy that is altruistic towards itself and spiteful towards other strategies. Implicitly, we have assumed that such a masquerade is infinitely costly to implement. However, if there were a strategy that sent exactly the same signal as the optimal punishment strategy, but behaved selfishly in each match, then it would be the strategy to emerge as the long-term winner rather than the optimal punishment strategy.

The most interesting case to consider is the intermediate one in which a masquerade is neither costless nor infinitely costly. Recall the assumption of perfect identification: a signal y_0 such that $\pi(y_0 | s, s) = 1$ and $\pi(y_0 | s, s') = 0$ for $s \neq s'$. The strategy s_0 is called the optimal punishment strategy. Let the action $a^m = \arg \max_a U(a, \bar{a})$ be the best response to altruism. Let us add to this environment a single *free-riding* strategy s^m with the property that $s^m(y) = a^m$ and $\pi(y_0 | s, s^m) = 1$ for all $s \in S$. The free-riding strategy perfectly mimics every other strategy and exploits the optimal punishment strategy to the maximum extent possible. Utility to the strategy s^m is given by $U(a, a^m) - \xi$, where $\xi > 0$ is the cost of masquerading. Then the optimal punishment strategy will continue to beat the field provided it can beat the free-riding strategy.

Theorem 6: *If $\max\{U(a^m, \bar{a}) - U(\bar{a}, \bar{a}), U(a^m, a^m) - U(\bar{a}, a^m)\} \leq \xi$ then s_0 beats the field.*

Proof: The optimal punishment strategy earns

$$u(s_0, \alpha s_0 + (1 - \alpha)s^m) = \alpha U(\bar{a}, \bar{a}) + (1 - \alpha)U(\bar{a}, a^m)$$

while the free-riding strategy gets

$$u(s^m, \alpha s_0 + (1 - \alpha)s^m) = \alpha U(a^m, \bar{a}) + (1 - \alpha)U(a^m, a^m) - \xi,$$

from which the results follows directly. □

The implication of this seems clear enough: we would not expect strategies to evolve that tell “small lies,” that is costly lies with little benefit. However, we might expect that highly evolved strategies will tell low cost lies that have a large benefit. No one expects a defendant in a murder trial to admit that he committed the murder. Witnesses with less at stake are viewed as more likely to tell the truth.

A particularly significant aspect of this theory is what it implies for games where cooperation is enforced by punishments off the equilibrium path. Consider for example, a simple noisy prisoner’s dilemma game followed by a “punishment round.” In this game, there is a gift that costs one to give. The gift is worth G and if it is given it is received with probability $1 - p$. In the second round, each player may optionally impose a penalty of P on the other player at a cost of C . Assume that $P > 1$ and $(1 - p)G \geq p(C + P) + 1$. Then, the optimal punishment strategy is to give the gift and punish if a gift is not received in turn. The best-response to the optimal punishment strategy is not to punish, and

$$U(a^m, \bar{a}) - U(\bar{a}, \bar{a}) = U(a^m, a^m) - U(\bar{a}, a^m) = pC.$$

If the signal is not very noisy, then $pC < \xi$ and it would not be worth employing the free riding strategy.

Appendix

Below we interpret μ as the measure describing the limit of the invariant distributions of the perturbed ($\varepsilon > 0$) Markov process.

Let μ^0 be an irreducible invariant measure of the Markov process in which $\varepsilon = 0$. Let ω be the set of profiles in the state space $(S^0)^m$ that this invariant distribution gives positive weight to. We call such an ω an *ergodic set*. Let Ω be the set

of all such ω . Note that this is a set of sets. Let $\text{supp}(\sigma)$ denote the set of pure strategies in the support of σ . First we establish some basic facts about Ω .

Lemma A1: (1) $\omega, \omega' \in \Omega \Rightarrow \omega \cap \omega' = \emptyset$; (2) If $\sigma, \sigma' \in \omega \in \Omega$ then $\text{supp}(\sigma) = \text{supp}(\sigma')$; (3) $\{\sigma(s)\} \in \Omega$ for all $s \in S^0$.

Proof: When $\varepsilon = 0$ we have the *relative best-response dynamic* in which one player switches with equal probability to one of the relative best-responses to the current state. The sets ω are by definition minimal invariant sets under the relative best-response dynamic. That these sets are disjoint is immediate from the definition. Pure profiles are absorbing since no strategy can be used unless it is already in use. This means that every set ω consisting of a singleton pure profile is in Ω . To see that $\text{supp}(\sigma) = \text{supp}(\sigma')$, observe that the relative best-response dynamic cannot ever increase the set of strategies in use. If there is a point $s \in \text{supp}(\sigma), s \notin \text{supp}(\sigma')$ then the probability that the relative best-response dynamic goes from σ to σ' is zero, which is inconsistent with the two strategies lying in the same ergodic set. □

Lemma A1 (2) implies that for each $\omega \in \Omega$ we may assign a unique set of pure strategies $S(\omega)$ corresponding to $\text{supp}(\sigma), \sigma \in \omega$.

To prove our results, we will use the characterization of μ given by Young [1993].¹⁰ Let τ be a tree whose nodes are the set Ω . We denote by $\tau(\omega)$ the unique predecessor of ω . An ω -tree is a tree whose root is ω . For any two points $\omega, \tilde{\omega} \in \Omega$ we define the resistance $r(\omega, \tilde{\omega})$ as follows. First, a path from ω to $\tilde{\omega}$ is a sequence of points $(\sigma_0, \dots, \sigma_K)$ with $\sigma_0 \in \omega, \sigma_K \in \tilde{\omega}$ and σ_{k+1} reachable from σ_k by a single player changing strategy. If the change from σ_k to σ_{k+1} is a relative best-response, the resistance of σ_k is 0; if the change is an imitation the resistance is 1; if the change is an innovation the resistance is n . The resistance of a path is the sum of the resistance of each point in the sequence. The resistance $r(\omega, \tilde{\omega})$ is the least resistance of any path from ω to $\tilde{\omega}$. The resistance $r(\tau)$ of the ω -tree τ is the sum over non-root nodes of $r(\tilde{\omega}, \tau(\tilde{\omega}))$. The resistance of ω , $r(\omega)$ is the least resistance of any ω -tree. The following Theorem is proved in Young [1993].

¹⁰ Although the standard convention in game theory is that a tree begins at the root, Young [1993] followed the mathematical convention that it ends there. We have used the usual game-theoretic convention, so our trees go in the opposite direction of Young's.

Young's Theorem: $\mu = \lim \mu^\varepsilon$ exists and $\mu(\omega) > 0$ if and only if

$$r(\omega) = \min_{\tilde{\omega} \in \Omega} r(\tilde{\omega})$$

Remark: The set of ω for which $\mu(\omega) > 0$ is called the *stochastically stable set*.

The basic tool for analyzing μ is tree surgery, by which we transform one tree into another and compare the resistances of the two trees. Suppose that τ is an ω -tree. For any nodes $\tilde{\omega} \neq \omega$ we *cut* the $\tilde{\omega}$ -subtree separating the original tree into two trees; one of which is the $\tilde{\omega}$ -subtree and the other that which is left over. This reduces the resistance by $r(\tilde{\omega}, \tau(\tilde{\omega}))$. If $\hat{\omega}$ is a node in either of the two trees, and $\hat{\omega}$ is the root of the other tree, we may *paste* $\hat{\omega}$ to $\hat{\omega}$ by defining $\tau(\hat{\omega}) = \hat{\omega}$. This tree has the root of the tree containing $\hat{\omega}$. The paste operation increases the resistance by $r(\hat{\omega}, \hat{\omega})$, so the new tree has resistance $r(\tau) + r(\hat{\omega}, \hat{\omega}) - r(\tilde{\omega}, \tau(\tilde{\omega}))$. These operations can be used to characterize classes of least resistance trees, by showing that certain operations do not increase resistance. They can also be used (as below) in proofs by contradiction, showing that certain trees cannot be least-resistance because it is possible to cut and paste in such a way as to reduce resistance.

Theorem 1: $\mu = \lim \mu^\varepsilon$ exists and $\mu(\omega) > 0$ implies that $\omega = \{\sigma(s)\}$ for some $s \in S^0$.

Proof: Existence of μ follows from Young's theorem. Suppose that $\mu(\omega) > 0$ and that ω is not a singleton pure profile. Let τ be a least resistance ω -tree. Let $\tilde{\omega} = \{\sigma(s)\}$ be a singleton pure strategy that is in some $\sigma \in \omega$, that is, $s \in S(\omega)$. Cut $\tilde{\omega}$ and paste the root ω to it. Since $\tilde{\omega}$ is a singleton pure profile, it requires at least one innovation to go anywhere, so that cutting it reduces resistance by at least n . On the other hand, since $\sigma \in \omega$ and $s \in S(\omega)$, we can go from ω to $\tilde{\omega}$ by no more than m imitations, so that pasting the root to $\tilde{\omega}$ increases resistance by at most m . By the assumption of unlikely innovation (that is, $n > m$), this implies that the new tree has strictly less resistance than the old tree, thus contradicting Young's Theorem. □

In what follows we simplify notation by writing $\mu(\sigma)$ instead of $\mu(\{\sigma\})$.

Theorem 2: *In the global matching or local interaction model, if s beats the field then $\mu(\sigma(s)) = 1$. If s weakly beats the field then $\mu(\sigma(s)) > 0$. Moreover, if $\mu(\sigma(s')) > 0$ then $u(s, \frac{1}{2}s + \frac{1}{2}s') = u(s', \frac{1}{2}s + \frac{1}{2}s')$.*

Proof: The proof is in two parts, concerning first the strict case and then the weak case.

1. Strict Case: Suppose that there is some ω with $\mu(\omega) > 0$. By Theorem 1, $\omega = \{\sigma(s')\}$ for some pure strategy s' . Let τ be the least resistance of the ω -tree. If $\{\sigma(s)\}$ is not the root, we may suppose that it is attached to some $\tilde{\omega}$, and consider cutting it and pasting the root to it. We now have two different arguments depending on whether matching is global or local.

Global Case: Considering first a transition from $\sigma(s)$ to $\sigma(s')$, it takes at least one innovation combined with (since s beats any point in $\tilde{\omega}$) $m/2$ imitations to arrive at $\tilde{\omega}$. So this reduces resistance by at least $n + m/2$. Considering next the reverse transition from $\sigma(s')$ to $\sigma(s)$, since s beats s' such a transition can be accomplished with one innovation and no more than $(m-1)/2$ imitations. So resistance is strictly reduced under this combined operation thus leading to a contradiction of Young's Theorem.

Local Case: As in the intuitive discussion in the text, when the configuration is such that all players $\ell = L, L+1, \dots, L+2k$ are all playing s we say that the state has *contiguous s* . We reiterate the observation that when s beats the field, a state with contiguous s is in the basin of $\sigma(s)$: the players in a contiguous set of s never find it a relative best response to switch from s , and players on the border of a contiguous set always find it a relative best response to switch to s . Starting from $\sigma(s)$ we must eliminate all contiguous sets of s in order to be able to exit the basin of $\sigma(s)$. This requires at least $m/(2k+1)$ players switching to a different strategy. So cutting $\{\sigma(s)\}$ from $\tilde{\omega}$ reduces resistance by at least $(m/(2k+1)) - 1$. On the other hand, starting from $\sigma(s')$, it is sufficient (following an innovation into s) for $2k$ players to switch to s in order to be able to establish a contiguous set and hence a basin of $\sigma(s)$. In this case resistance will have been increased by at most $2k$. Given the assumption made in the main body of the article that $m > (2k+1)^2$, it follows that resistance has been strictly reduced by this combined operation, thus leading to the contradiction of Young's Theorem.

2. Weak Case: In this case we can only conclude (for both global and local interaction) that resistance is not increased after our tree-cutting operation. This conclusion then

implies that $\{\sigma(s)\}$ must be at the root of a least-cost tree, giving us $\mu(\sigma(s)) > 0$ as required.

If we have $\mu(\sigma(s')) > 0$ and $u(s, \frac{1}{2}s + \frac{1}{2}s') - u(s', \frac{1}{2}s + \frac{1}{2}s') \neq 0$ (note the inequality here) then, since s weakly beats the field by assumption, it follows that s beats s' . From this we can conclude that $\{\sigma(s')\}$ cannot be the root of a least cost tree, and so it must be that $\mu(\sigma(s')) = 0$, thus proving the last statement in the theorem. \square

Theorem 3: (1) $\mu(\sigma(s_0)) > 0$. (2) If $\mu(\sigma(s)) > 0$ then $U(s(y_0), s(y_0)) = U(\bar{a}, \bar{a})$ and $U(s(y), \underline{a}) = U(\underline{a}, s(y))$ for all y such that $\pi(y | s_0, s) > 0$.

Proof: We first show that s_0 weakly beats the field. Suppose that $s \neq s_0$. Let $s_\alpha := \alpha s_0 + (1 - \alpha)s$, then

$$\begin{aligned} u(s_0, s_\alpha) - u(s, s_\alpha) &\geq \alpha U(\bar{a}, \bar{a}) - (1 - \alpha)U(s(y_0), s(y_0)) \\ &\quad + \sum_{y \in Y} ((1 - \alpha)U(\underline{a}, s(y)) - \alpha U(s(y), \underline{a}))\pi(y | s, s_0) \end{aligned}$$

Since this expression is linear in α , to show that s weakly beats the field, it suffices to show that it is non-negative both for $\alpha = \frac{1}{2}$ and $\alpha = 1$. When $\alpha = \frac{1}{2}$ we have

$$\begin{aligned} u(s_0, s_\alpha) - u(s, s_\alpha) &\geq \frac{1}{2}[U(\bar{a}, \bar{a}) - U(s'(y_0), s'(y_0))] \\ &\quad + \frac{1}{2} \min_{a' \in A} [U(\underline{a}, a') - U(a', \underline{a})] \end{aligned}$$

The first term is non-negative by the definition of \bar{a} ; the second term is non-negative by Assumption 3. For $\alpha = 1$, we have

$$u(s_0, s_\alpha) - u(s, s_\alpha) \geq U(\bar{a}, \bar{a}) - \max_{a' \in A} U(a', \underline{a})$$

Assumption 3 implies $U(\underline{a}, a') \geq U(a', \underline{a})$ for all $a' \in A$ and Assumption 2 implies $U(\bar{a}, \bar{a}) \geq \min\{U(\underline{a}, a'), U(a', \underline{a})\}$. Hence, it follows that $U(\bar{a}, \bar{a}) - U(a', \underline{a}) \geq 0$ for all $a' \in A$ which shows that s_0 weakly beats the field. Theorem 2 therefore implies that $\mu(\sigma(s_0)) > 0$ and that if $\mu(\sigma(s)) > 0$ then

$$\begin{aligned} 0 &= u(s_0, s_{\frac{1}{2}}) - u(s, s_{\frac{1}{2}}) \\ &\geq \frac{1}{2}[U(\bar{a}, \bar{a}) - U(a, a)] + \frac{1}{2} \sum_{y \in Y} \{U(\underline{a}, s(y)) - U(s(y), \underline{a})\}\pi(y | s, s_0) \end{aligned}$$

From the definitions of \bar{a} and \underline{a} as well as from the expression above, we see that this is possible only if $U(a, a) = U(\bar{a}, \bar{a})$ and $U(s(y), \underline{a}) = U(\underline{a}, s(y))$ for all t with

$$\pi(y | s, s_0) > 0.$$

☑

Theorem 4: $\mu(s_0) = 1$.

Proof: Theorem 4 follows from Theorem 5 (proved below) since the action that defines $s_0(y)$ is unique. In particular, Theorem 5 implies that $\mu(\sigma(s)) > 0$ if and only if $s(y) = s_0(y)$ for all $y \in Y$ and therefore $\mu(\sigma(s)) > 0$ implies $s = s_0$

☑

Theorem 5: *If s_0 is informationally dominant then, $\mu(\sigma(s_0)) > 0$. Moreover, for every strategy $\sigma(s)$ with $\mu(\sigma(s)) > 0$ we have $u(s, s) = u(s_0, s_0)$ and $u(s, s_0) = u(s_0, s)$.*

Proof: We first show that s_0 weakly beats the field. Let $s_\alpha = \alpha s_0 + (1 - \alpha)s$ for some $s \in S$. We must show that $u(s_0, s_\alpha) - u(s, s_\alpha) \geq 0$ for $\alpha \in [\frac{1}{2}, 1)$. Since $u(s_0, s_\alpha) - u(s, s_\alpha)$ is linear in α it suffices to show that it is non-negative at both $\alpha = \frac{1}{2}$ and $\alpha = 1$.

$$\begin{aligned} & u(s_0, s_{\frac{1}{2}}) - u(s, s_{\frac{1}{2}}) \\ &= \frac{1}{2} \sum_{y \in Y} \{ [p(y) - q(y)]b(s_0(y)) - [p(y) + q(y)]c(s_0(y)) \\ &\quad - [\pi(y | s, s) - \pi(y | s, s_0)]b(s(y)) - [\pi(y | s, s) + \pi(y | s, s_0)]c(s(y)) \} \\ &= \frac{1}{2} \sum_{y \in Y} \{ [p(y) - q(y)]b(s_0(y)) - (p(y) + q(y))c(s_0(y)) \\ &\quad - \sum_{z \in Y} \lambda_{zy} ([p(y) - q(y)]b(s(z)) - (p(y) + q(y))c(s(z))) \} \\ &\geq 0 \end{aligned}$$

where the last inequality follows since $s_0(y)$ maximizes

$$[p(y) - q(y)]b(a) - [p(y) + q(y)]c(a)$$

and $\sum_{z \in Y} \lambda_{zy} = 1$.

Next we consider the case $\alpha = 1$. Let $q'(y) = \pi(y | s, s')$. We may write the utility difference as

$$\begin{aligned}
& u(s_0, s_1) - u(s, s_1) = \\
& \sum_{y \in Y} \{ (p(y)(b(s_0(y)) - c(s_0(y))) - q(y)(b_0(s(y)) - c(s(y)))) \} \geq \\
& \sum_{y \in Y} \{ (p(y) - q(y))b(s_0(y)) - p(y)c(s_0(y)) \}
\end{aligned}$$

where the inequality follows from $c(a) \geq 0$. Recall that there is an action \hat{a} with $c(\hat{a}) = 0$ and therefore, by the definition of s_0

$$\begin{aligned}
(p(y) - q(y))b(s_0(y)) - p(y)c(s_0(y)) & \geq (p(y) - q(y))b(s_0(y)) - (p(y) + q(y))c(s_0(y)) \\
& \geq (p(y) - q(y))b(\hat{a})
\end{aligned}$$

Hence,

$$u(s_0, s_1) - u(s, s_1) \geq \sum_{y \in Y} (p(y) - q(y))b(\hat{a}) = 0$$

where the last equality follows from $\sum p(y) = \sum q(y) = 1$. This shows that s_0 weakly beats the field.

When $\mu(\sigma(s)) > 0$ then it follows that

$$\begin{aligned}
& u(s_0, s_{1/2}) - u(s, s_{1/2}) = \\
& \frac{1}{2} \sum_{y \in Y} \{ [p(y) - q(y)]b(s_0(y)) - (p(y) + q(y))c(s_0(y)) \\
& - \sum_{z \in Y} \lambda_{zy} ([p(y) - q(y)]b(s(z)) - (p(y) + q(y))c(s(z))) \} = 0
\end{aligned}$$

Since $s_0(y)$ is the unique optimal action it follows that for $\lambda_{zy} > 0$, $s(z) = s_0(y)$ and hence

$$\begin{aligned}
u(s, s) & = \sum_{y \in Y} \sum_{z \in Y} \lambda_{zy} (p(y)b(s(z)) - p(y)c(s(z))) \\
& = \sum_{y \in Y} (p(y)b(s_0(y)) - p(y)c(s_0(y))) \\
& = u(s_0, s_0)
\end{aligned}$$

An analogous argument shows that $u(s_0, s) = u(s, s_0)$.

□

References

- Bergin, J. and B. L. Lipman [1994]: "Evolution with state-dependent Mutations," UCLA Theory Workshop, 4/14/94.
- Bergstrom, T. [2002]: "Evolution of Social Behavior: Individual and Group Selection Models," forthcoming, *Journal of Economic Perspectives*.
- Binmore, K. and L. Samuelson [1992]: "Evolutionary Stability in Repeated Games Played by Finite Automata," *Journal of Economic Theory* 57, 278-305.
- Dawid, H. [1999]: *Adaptive Learning by Genetic Algorithms*, (Berlin: Springer).
- Dutta, J. and K. Prasad, "Imitation and Long Run Outcomes," *The B.E. Journals in Theoretical Economics, Topics in Theoretical Economics*, 4, 1, Article 7.
- Ellison, G. [1995]: "Basins of Attraction, Long Run Equilibria, and the Speed of Step-by-Step Evolution," *Review of Economic Studies*, 67 (1), 17-45, 2000.
- Eshel, I., L. Samuelson, and A. Shaked [1998], "Altruists, Egoists, and Hooligans in a Local Interaction Model, *American Economic Review* 88 (1998), 157-179.
- Frank, R. H. [1987], "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience?," *American Economic Review*, 77 593-604.
- Fehr, E. and S. Gächter [2000]: "Cooperation and Punishment in Public Goods Experiments," *American Economic Review*, 90: 980-994.
- Friedlin, M. and A. Wentzell [1982]: *Random Perturbations of Dynamical Systems*, (New York: Springer Verlag).
- Friedman, D. [1998]: "On Economic Applications of Evolutionary Game Theory," *Journal of Evolutionary Economics*, 8: 15-43.
- Fudenberg, D. and D. Kreps [1993]: "Learning Mixed Equilibria," *Games and Economic Behavior*, 5: 320-367.
- Fudenberg, D. and D. K. Levine [1997]: "Measuring Subject's Losses in Experimental Games," *Quarterly Journal of Economics*, 112: 508-536.
- Fudenberg, D. and D. K. Levine [1998]: *The Theory of Learning in Games*, (Cambridge: MIT Press).
- Fudenberg, D. and E. Maskin [1986]: "The Folk Theorem for Repeated Games with Discounting and Incomplete Information," *Econometrica*, 54: 533-54.

- Kim, Y. and J. Sobel, [1995]: "An Evolutionary Approach to Pre-Play Communication", *Econometrica*, 63, 1181-1193.
- Kirman, A. [1993]: Ants, Rationality, and Recruitment, *The Quarterly Journal of Economics*, CVIII, 137-156.
- Hahn, S. [1995]: "The Long Run Equilibrium in an asymmetric Coordination Game," Harvard.
- Johnson, P., D. K. Levine and W. Pesendorfer [2000]: "Evolution and Information in a Gift Giving Game," forthcoming, *Journal of Economic Theory*.
- Kandori, M. [1992a]: "Repeated Games played by Overlapping Generations of Players," *Review of Economic Studies*, 59: 81-92
- Kandori, M. [1992b]: "Social Norms and Community Enforcement," *Review of Economic Studies*, 59: 61-80.
- Kandori, M. and R. Rob [1993]: "Bandwagon effects and long run technology choice," University of Tokyo Discussion Paper 93-f-2.
- Kandori, M., G. Mailath and R. Rob [1993]: "Learning, Mutation and Long Run Equilibria in Games," *Econometrica*, 61: 27-56.
- Levine, D. K. [1998]: "Modeling Altruism and Spitefulness in Experiments," *Review of Economic Dynamics*, 1: 593-622.
- Levine, D. K. and W. Pesendorfer [2000]: "Evolution Through Imitation in a Single Population," *UCLA*.
- Morris, S., R. Rob and H. Shin [1993]: "p-dominance and Belief Potential," *Econometrica*, 63: 145-158.
- Robson, A. [1990]: "Efficiency in Evolutionary Games: Darwin, Nash and the Secret Handshake," *Journal of Theoretical Biology*, 144, 379-396.
- Schlag, K. H. [1999], "Why Imitate, and If So, How?" A Bounded Rationality Approach to Multi-armed Bandits, *Journal of Economic Theory*, 78, 130-156.
- Van Damme, E. and J. W. Weibull [1998]: "Evolution with Mutations Driven by Control Costs," Tilburg.
- Vega-Redondo, F. [1997]: "The Evolution of Walrasian Behavior," *Econometrica* 65, 375-384.
- Weibull, J. [1995], *Evolutionary Game Theory*, MIT Press.
- Young, P. [1993]: "The Evolution of Conventions," *Econometrica*, 61: 57-83.