The Evolution of Cooperation Through Imitation[1]

David K. Levine and Wolfgang Pesendorfer[2]

First version: September 29, 1999

This version: March 3, 2001

Abstract: We study when cooperation and conflict emerge in games such as the Prisoner's Dilemma. We make use of two simple ideas: existing strategies are more likely to be imitated than new strategies are to be introduced, and prior to interaction it is possible to identify how an opponent will behave. We use methods introduced by Kandori, Mailath and Rob [1993] and Young [1993] to examine the long-run evolutionary limit of this system. This limit contains only pure strategies. A sufficient condition for a unique limit is that a strategy beat all others in pairwise contests. When players can perfectly determine the strategy used by their opponent, full cooperation is achieved. When identification is imperfect, we characterize the degree of cooperation and conflict.

[1] We are grateful to support from National Science Foundation Grant SES-99-86170 and the UCLA Academic Senate. We would like to thank seminar participants at the Stanford and ITAM Theory Workshops and the Southern California Theory Conference. We owe a particular debts of gratitude to Philip Johnson who helped get us started on this line of research and to Pedro Dal Bo for pointing out numerous mistakes and typos, and correcting an erroneous proof.

[2] Departments of Economics UCLA and Princeton.

1. Introduction

Prisoner's dilemma type games have found widespread application in the study of problems ranging from public goods, to the conflict between nations and to the study of collusion in oligopoly. It is well known that both cooperation and conflict are possible in settings ranging from that of repeated games to games with overlapping generations of players [Kandori, 1992a], and populations of players with information about past play conveyed by information systems [Kandori, 1992b]. While versions of the folk theorem [Fudenberg and Maskin, 1986] assure us that both cooperation and conflict are possibilities, there is little systematic theory to tell us when one outcome emerges rather than the other. This paper attempts to provide such a theory.

Our starting point is a model of evolutionary learning. Players in this model 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 randomly matched into pairs to play a Prisoner Dilemma-type game. Unlike standard pairwise matching games, we allow a player's strategy to depend on the opponent's type. The type of a player is in turn determined by his behavior. We interpret this model as describing a situation where strategies are rules of thumb that players use in a wide variety of circumstances. Players are committed to a particular rule because it is too costly to change behavior in any particular match. Players can identify the opponent's rule either because they observe other interactions of the opponent or because such rules affect observable features of the opponent such as his language or his gestures.

For example, consider the strategy that takes a cooperative action if the opponent is of the same type, that is, has chosen the same strategy, and a punishment action if the opponent is of a different type. 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 rule that defects independently of the opponent's type. Again, this rule is an optimal choice when the opponent is expected to use it. Thus, with a sufficiently rich set of types, these games typically have multiple equilibria that resemble the equilibria in 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.

Our work 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.

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. This modified propagation mechanism 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 pairwise contests. As we illustrate through examples, this is implied by, but weaker than, the criterion of ½-dominance proposed by Morris, Rob and Shin [1993]. Ellison [2000] shows that if a strategy is ½-dominant, it is the unique long-run outcome.

In addition to the work mentioned above, there are several other papers that have a connection to our results. Bergin and Lipman [1994] show that the relative probabilities of different types of noise can make an enormous difference in long-run equilibrium; here we explore one 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 pairwise contests is sufficient for a strategy to be the long-run outcome. However, in their model winning all pairwise contests implies 1/2-dominance, which is not true in our applications.

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 symmetric normal form game played by a single population of players. There are finitely many pure strategies, denoted by $s \in S$. Mixed strategies are vectors of probabilities denoted by $\sigma \in \Sigma$. The support of a mixed strategy σ is denoted by $\mathrm{supp}(\sigma) := \{\mathrm{s} | \sigma(\mathrm{s}) > 0\}$. A mixed strategy is called *pure* if it puts unit weight on a single pure strategy; abusing notation, we denote by s the mixed strategy that puts probability one on s. The utility of a player depends on his own pure strategy and the mixed strategy played by the population. It is written $u(s,\sigma)$ and continuous in σ. A prototypical example is a game in which players from different populations are randomly matched to play particular player roles.

There are m players in the population, each of whom plays a pure strategy. The distribution of strategies at time t is denoted by $\sigma_t \in \Sigma$. Starting with an initial distribution σ_0, the distribution σ_t is determined from σ_{t-1} according to the following "imitative" process.

1) One player i is chosen at random. Only this player changes his strategy.
2) With probability $C\varepsilon$ player i chooses from S randomly using the probabilities σ_{t-1}. This is called *imitation*: strategies are chosen in proportion to how frequently they were played in the population in the previous period.
3) With probability ε^n player i chooses each strategy from S with equal probability. This is called *innovation*: strategies are picked regardless of how widely used they are, or how successful they are.
4) With probability $1 - C\varepsilon - \varepsilon^n$ player i randomizes with equal probability among the strategies that solve

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

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

This process gives rise to a Markov process M on the state space $\Sigma^m \subset \Sigma$ consisting of all mixed strategies consistent with the grid induced by each player playing a pure strategy. The process M is positively recurrent, and so has a unique invariant distribution μ^ε. Our goal is to characterize $\mu \equiv \lim_{\varepsilon \to 0} \mu^\varepsilon$.

Our main assumption is that imitation is much more likely than innovation.

Unlikely Innovation: $n > m$.

Unlikely Innovation implies that, as $\varepsilon \to 0$, the probability that *every* player changes strategy by imitation is much greater than the probability that a single player innovates. This assumption is maintained throughout the paper.

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.[3]

[3] 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. As long as σ_{t-1} is interior the relative best response is of course a true best response.

Like much of the existing literature, we consider a single population.[4] In our model the single-population assumption implies that all players are *a priori* identical, and that behavior is imitated throughout the population.

We first establish a basic result about the limit distribution $\mu = \lim_{\varepsilon \to 0} \mu^{\varepsilon}$. Unlikely innovation implies that mixed strategies are less stable than pure strategies. A mixed strategy can evolve into a pure strategy purely through imitation, while a pure strategy cannot evolve at all without at least one innovation. Theorem 1 confirms this intuition by showing that the limit invariant distribution μ places weight only on pure profiles in Σ^m.

Theorem 1: $\mu = \lim \mu^{\varepsilon}$ *exists and* $\mu(\sigma) > 0$ *implies that* σ *is a pure strategy.*[5]

Our main characterization result (Theorem 2) shows that if a pure strategy beats all other strategies in pairwise contests, then it is the unique stochastically stable state. We begin by explaining what it means to win pairwise contests. For $0 \leq \alpha \leq 1$, the mixed strategy that plays s with probability α and $\tilde{s}$ with probability $1-\alpha$ is denoted by $\alpha s + (1-\alpha)\tilde{s}$.

Definition 1: *The strategy* s *beats* $\tilde{s}$ *iff*

$$u(s, \alpha s + (1-\alpha)\tilde{s}) - u(\tilde{s}, \alpha s + (1-\alpha)\tilde{s}) > 0$$

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

Thus, a strategy s beats $\tilde{s}$ if s yields higher utility against any combination of s and $\tilde{s}$ that puts more weight on s than on $\tilde{s}$. In Definition 2, we weaken this concept to allow for ties.

Definition 2: *The strategy* s *weakly beats* $\tilde{s}$ *iff*

$$u(s, \alpha s + (1-\alpha)\tilde{s}) - u(\tilde{s}, \alpha s + (1-\alpha)\tilde{s}) > 0$$

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

[4] 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.

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

$$u(s, \alpha s + (1-\alpha)\tilde{s}) - u(\tilde{s}, \alpha s + (1-\alpha)\tilde{s}) = 0$$

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

A strategy "beats the field" if it beats every other pure strategy.

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

In Theorem 2, we show that if strategy s beats the field then s is the unique long run outcome. If strategy s weakly beats the field then the long run distribution must place strictly positive probability on s. In addition, s and $\tilde{s}$ must get the same payoff against a population that plays s and $\tilde{s}$ with equal probability. In the applications below, we will use this property to show that $\tilde{s}$ must be similar to s.

Theorem 2: *If* m *is sufficiently large and* s *beats the field then* $\mu(s) = 1$. *If* m *is sufficiently large and* s *weakly beats the field then* $\mu(s) > 0$. *Moreover, if* $\mu(\tilde{s}) > 0$ *then* $u(s, \tfrac{1}{2}s + \tfrac{1}{2}\tilde{s}) - u(\tilde{s}, \tfrac{1}{2}s + \tfrac{1}{2}\tilde{s}) = 0$.

To gain an intuition for Theorem 2, consider a pair of pure strategies, s, s'. Let $P(s \to s')$ denote the probability that the Markov process to move from the state where all players choose the strategy s to a state where all players choose s'. We first compute an upper bound for this probability. Note that one innovation is required to leave any state where all agents play the same strategy. Moreover, since players choose a relative best response with probability close to one, we need a number of imitations that ensures that s' is a relative best response. Otherwise, the Markov process simply returns to the state where all players choose s. Assume that s beats s'. Since s yields a strictly higher payoff than s' to $\alpha s + (1-\alpha)s', \alpha \geq 1/2$, at least $m/2$ imitations are required for the transition from s to s' when m is even. For small ε, we therefore conclude that

$$P(s \to s') \leq \varepsilon^n (C\varepsilon)^{m/2}$$

A similar argument allows us to find a lower bound on the transition from the state where all players choose s' to the state where all players choose s. The transition requires again one innovation and, since s beats s', at most $m/2 - 1$ imitations. A lower bound for the transition probability is therefore

$$P(s' \to s) \le \varepsilon^n \, (C\varepsilon)^{m/2-1}$$

Thus, the ratio of the two probabilities is

$$\frac{P(s' \to s)}{P(s \to s')} \approx \frac{\varepsilon^n \, (C\varepsilon)^{m/2-1}}{\varepsilon^n \, (C\varepsilon)^{m/2}} = \frac{1}{C\varepsilon}$$

and hence we conclude that it is far more likely that the Markov process makes a transition from s' to s than it is to make a transition from s' to s. By Theorem 1 only pure strategies are candidates for states to occur with significant probability in the long run when ε is small. Moreover, if s beats the field, the argument above applies to *every* strategy $\tilde{s} \neq s$. Hence, it follows that in the long run only the state where all players choose s can occur with significant probability.

The hypothesis, that a strategy beats the field is connected to the idea of ½-dominance introduced by Morris, Rob and Shin [1993]. The concept of ½-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 pure profiles, 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, ½-dominance clearly implies winning all pairwise contests, so if there is a ½-dominant strategy, from Morris, Rob and Shin [1993] it is stochastically stable with respect to the usual evolutionary dynamic, and it is also stochastically stable when innovation is unlikely.

3. Matching Games with Behavioral Types

In this section, we apply the characterization results to a pairwise matching game. Every period players are matched into pairs to play a symmetric normal form game. Prior 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

$$U(a,a')$$

Players simultaneously choose strategies from an abstract strategy space S. Strategies serve two roles. Depending on the chosen strategies, players receive information about the opponent's strategy choice. Based on that information, each strategy takes an action $a \in A$. Thus, strategies serve a dual role: first, they influence the information that is generated about the player and his opponent and second, they govern the behavior as a function of the generated information.

Formally, each player receives a signal $y \in Y$, a finite set. The probability of receiving a signal depends on the strategies used by both players and is given by $\pi(y \mid s,s')$ if the player uses s and the opponent s'. These signals are private information. In our interpretation, a signal reflects what the opponent can learn about the player's behavior prior to the interaction.

Each strategy also determines how the player will respond to different signals. That is, each strategy s gives rise to a map (which we denote by the same symbol) $s : Y \to A$. Notice that several strategies may induce the same map, yet differ in the probability with which they send signals. We will assume that for every map from signals to actions there is some strategy that induces that map.

Assumption 0: *If* $f : Y \to A$ *there is a strategy* $s \in S$ *with* $s(y) = f(y)$.

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 signals.

To motivate this construction, assume that strategies govern the behavior of agents in many games. Players are committed to a particular strategy because it is too costly to change behavior in any particular match. If a player can observe previous interactions of the opponent he is able to predict the opponent's behavior. Alternatively, it may be the case that strategies are directly observable. An individual who rarely lies may blush whenever he is dishonest. Seeing the opponent blush indicates that he is unlikely to be dishonest in future interactions. (This example is due to Frank (1987)). The "signal" encapsulates such information in our model.

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 the opponent uses the same strategy 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.

Note that signals affect payoffs only indirectly by affecting behavior. If player i uses strategy s and his opponent uses strategy s' then the expected payoff of player i is given by

$$\sum_{y' \in Y} \sum_{y \in Y} U(s(y), s'(y'))\pi(y \mid s, s')\pi(y' \mid s', s)\sigma(s')$$

Therefore, the function $u(s,\sigma)$ (as defined in the previous section) is given by

$$u(s,\sigma) = \sum_{s' \in S} \sum_{y' \in Y} \sum_{y \in Y} U(s(y), s'(y'))\pi(y \mid s, s')\pi(y' \mid s', s)\sigma(s')$$

We consider three scenarios. In the first – analyzed in section 3.1 – a player is able to determine with certainty whether the opponent uses the same strategy as he does. For this case, we show for a general class of utility functions that full cooperation will emerge as the long run outcome. In the second scenario – analyzed in section 3.2 – we consider the case of noisy identification of the opponent's behavior for a restricted class of utility functions. For that environment, we relate the degree of cooperation to the ability of agents to identify types who use similar rules. The model of section 3.2 supposes a great deal of symmetry in the signaling mechanism. The third scenario – analyzed also in section 3.2 – replaces the symmetry assumption with an informational dominance condition.

3.1 Perfect Identification

The first scenario assumes that each player can identify with certainty whether the opponent is using the same strategy. For notational simplicity, we assume that there is one signal, denoted by y_0, that perfectly identifies that the opponent is using the same strategy.

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

We make two assumptions on U. Assumption 2 requires that there is no asymmetric profile that makes both players strictly better off than any symmetric profile.

Assumption 2: $U(a, a') > \max_a U(a, a) \Rightarrow U(a', a) \leq \max_a U(a, a)$

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 $\overline{a}$ be a Pareto optimal symmetric outcome

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

By assumption 0, there is a strategy s_0 for which.

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

The strategy s_0 takes the Pareto efficient action when the opponent uses the same strategy and punishes the opponent by taking action $\underline{a}$ when the opponent chooses a different strategy. Note that the punishment action maximizes the minimum difference between the player's payoff and his opponent's payoff.

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 used with positive probability is similar to s_0: when s meets s both players receive the payoff $U(\overline{a},\overline{a})$; when s meets s_0 both players receive the same payoff.

Theorem 3: Under A0-A3, $\mu(s_0) > 0$. *If* $\mu(s) > 0$ *then* $u(s,s) = u(s_0,s_0) = U(\overline{a},\overline{a})$ *and* $u(s,s_0) = u(s_0,s)$.

Theorem 3 implies that if $\overline{a}$ 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 s_0 is the unique outcome in the long-run limit. On the other hand, suppose there is some redundancy in the description of the game because, for example, there may be other strategies that induce the same map as s_0 or because there is an action $\hat{a}$ with $U(\hat{a},a) = U(\overline{a},a)$, for all $a \in A$. In either case, there are two or more strategies that satisfy the requirement of s_0. These strategies differ in extraneous detail only but will not recognize each other as the "same strategy". The long-run distribution places positive weight on every such strategy and if we observe the system for a long time we will typically observe each player using the same version of s_0. However, occasionally there will be a transition from one version of s_0 to another. During this brief period of transition, players using different versions will punish one another by choosing $\underline{a}$, the action that maximizes the difference between the two player's payoff.

The proof of Theorem 3 in the Appendix shows that the strategy s_0 weakly beats the field, that is; weakly beats every other strategy in a pairwise contest. However, the strategy s_0 need not be ½ dominant in the ordinary sense. 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 "defect" otherwise. Against $\frac{1}{2}s + \frac{1}{2}\underline{s}$ the strategy $\tilde{s}$ does better than s and therefore s is not ½ dominant. Such a strategy seems to serve no useful purpose except to make $\overline{s}$

look good against s . 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 they will not remain around long enough for players to discover that they should play $\overline{s}$.

3.2 Gift Exchange and Imperfect Identification

In this section, we specialize to a simple additively separable structure. 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

$$U(a,a') = b(a') - c(a) \tag{1}$$

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 and b describes the benefit of the good for the opposing player. Games with this utility function resemble a prisoner's dilemma in that the cost minimizing action is dominant.

We assume that $c(a) \geq 0$, with $c(a) = 0$ for some action $\hat{a} \in A$. Note that the utility function (1) satisfies Assumptions 2 and 3.

In this section, we consider strategies that may not be able to identify with certainty when the opponent is using the same strategy. We first analyze how this noisy information about the opponent's behavior affects the strategies that emerge in the long run.

To keep things simple, we begin by assuming that all strategies use the same symmetric information structure described in Assumption 4.

Assumption 4:

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

Thus, $p(y)$ describes the probability that a player has type y if he and his opponent use the same strategy whereas $q(y)$ describes the same 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 also 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 $q(y) > p(y)$.

The strategy s_0 is defined as follows. For every signal y the action $s_0(y)$ solves

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

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 in the long run only strategies that behave like s_0 are played.

Theorem 4: $\mu(s) > 0$ if and only if $s(y) = s_0(y)$ for all $y \in Y$.

When the signal is y the strategy s_0 takes the action that maximizes

$$\frac{p(y)-q(y)}{p(y)+q(y)}b(a)-c(a)$$

Let s denote the opponent's strategy choice, and suppose that there a prior of $\frac{1}{2}$ that the opponent is using s_0. In that case,

$$\Pr(s = s_0 \mid y) = \frac{p(y)}{p(y)+q(y)}; \Pr(s \neq s_0 \mid y) = \frac{q(y)}{p(y)+q(y)}$$

and therefore

$$\frac{p(y)-q(y)}{p(y)+q(y)} = \Pr(s = s_0 \mid y) - \Pr(s \neq s_0 \mid y)$$

Hence, the objective function puts a larger weight on the opponent's benefit when he is more likely to use strategy s_0. In our gift exchange interpretation, the gift is decreasing as the signal indicates an opponent that is more likely to be different. It is worth emphasizing that in a long-run stable outcome typically all players are choosing the same strategy. Hence, the probability that the 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. Unlike the case of Theorem 3 where a socially efficient payoff was realized in the long-run stable outcome, here inefficiencies persist because players respond to signals *as if* one-half of the population were using a different strategy and hence needed to be punished.

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 at determining whether the 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 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 that does not use s_0. Suppose, however, that other strategies are described by general signal distributions $\pi(\cdot \mid s,\cdot)$. Below, we define the concept of informational dominance and 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 signals generated by s if there is a non-negative matrix

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

such that

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

In other words, the signals generated by $\pi(\cdot \mid s,\cdot)$ 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 informationally dominant strategy even though strategy s does better at identifying a third strategy $\overline{s}$.

A trivial example of an informationally dominant strategy is a strategy that cannot be distinguished from any other strategy. In that case, $\pi(y \mid s, s_0) = \pi(y \mid s, s)$ 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 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 *informationally dominant then,* $\mu(s_0) > 0$. *Moreover, for every strategies* s *with* $\mu(s) > 0$ *we have* $u(s,s) = u(s_0, s_0)$ *and* $u(s, s_0) = u(s_0, s)$.

In this section, we have restricted the informationally dominant strategy to generate symmetric information, that is, 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. Hence, the symmetry assumption in this section is 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 give up the symmetry assumption we must replace it with a requirement that preserves 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 is informationally superior to the signal generated by any opponent.

We conclude this section by illustrating Theorems 4 and 5 in the following 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 the opponent uses the same strategy. In that case, Theorem 4 is 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 the outcome is efficient 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 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 between the two players, as required by Theorem 3. Since the punishment action minimizes the sum of the player's cost and of the opponent's benefit, the player is willing to incur a cost if it leads to a negative payoff of an opponent who does not use s_0.

Example 2. To the environment of Example 1, we add the strategy $\bar{s}$, which can masquerade as any other strategy. Thus, a player using strategy s cannot determine whether the opponent uses $\bar{s}$ or s, hence $\pi(y \mid s,\bar{s}) = \pi(y \mid s,s)$ 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 ½ dominant in the ordinary sense. Consider the environment of Theorem 5 and assume that s_0, the informationally dominant long-run outcome is not constant. Let $\tilde{s}$ be a constant strategy that always plays $\hat{a}$. Suppose, that there are signals that enable a strategy $\underline{s}$ to identify

$\tilde{s}$ with certainty and 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 against $\frac{1}{2}s + \frac{1}{2}\underline{s}$ and therefore s is not ½ 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

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

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

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

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. When

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

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.

4. Conclusion

In conclusion, we offer an alternative interpretation of our model as describing the evolution of interdependent preferences. The formulation is based on work by Gul and Pesendorfer (2001).[6]

Consider the separable example and assume that when an individual takes action a his monetary payoff is

$$b(a^{-i}) - c(a^i)$$

Hence, monetary payoffs depend on the opponent's action only through the lump sum $b(a^{-i})$. Suppose that the utility of a player depends both on his monetary payoff and on his opponent's monetary payoff and preferences. The signal $y \in Y$ encapsulates the information the player has about his opponent's preference type. The function $s \in S$ describes for a particular preference, the optimal choice from A for every signal y.

To re-interpret the evolutionary dynamic, assume that each period one player dies and is replaced by a new player. With probability $C\varepsilon$ the new player is the offspring of a player randomly selected from the existing population of players. With probability $1 - C\varepsilon - \varepsilon^n$ the new player is the offspring of the most successful player in the population. In either case, the new player inherits the preference type of his parent. We can view this process as a version of the replicator dynamic in which the most successful type reproduces with much larger probability than other types. Finally, with probability ε^n the new player is assigned an arbitrary preference type at random. This corresponds to a mutation.

Our results show that evolution can be expected to select preferences that treat opponents differently depending on their preferences. In particular, in the long-run we expect to find types that behave altruistically towards opponents with the same preferences and spitefully towards opponents with different preferences. Casual observation suggests that individuals indeed behave more altruistically when they can "identify with" the beneficiary of their altruism.

A critical assumption is that preferences are observable at least with some probability. As before, this assumption may be justified because preferences affect

---

[6] Gul and Pesendorfer (2001) develop a canonical model of interdependent preferences for which a signal space representation can be found.

observable behavior in a variety of contexts unrelated to the game in question. An example used by Frank (1987) is that of a person who blushes upon telling a lie and thereby reveals a disutility of dishonest behavior. As Frank points out, a blusher will have an advantage in situations that require trust. Thus, it can be expected that evolution endows honest types with behavioral traits that make their honesty observable.

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 mixed strategies in the state space Σ^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 $S(\sigma)$ denote the set of pure strategies used with positive probability in σ. First we establish some basic facts about Ω.

Lemma A1: *The sets* ω *are disjoint. Each set consisting of a singleton pure profile* $\{s\} \in \Omega$. *If* $\sigma, \sigma' \in \omega \in \Omega$ *then* $S(\sigma) = S(\sigma')$.

Proof: When $\varepsilon = 0$ we have the *relative best-response dynamic* in which in each period 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 single pure strategy is in Ω. To see that have $S(\sigma) = S(\sigma')$, observe that the relative best-response dynamic cannot ever increase the set of strategies in use. If there is a point $s \in S(\sigma), s \notin S(\sigma')$ then the probability that the best-response dynamic goes from σ to σ' is zero, which is inconsistent with the two strategies lying in the same ergodic set.

☑

The third part of the Lemma means that for each $\omega \in \Omega$ we may assign a unique set of pure strategies $S(\omega)$ corresponding to $S(\sigma), \sigma \in \omega$.

To prove our results, we will use the characterization of μ given by Young [1993].[7] 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 $(\sigma_0, \ldots, \sigma_K) \subset \Sigma^m$ with $\sigma_0 \in \omega$, $\sigma_K \in \tilde{\omega}$ and σ_{k+1} reachable from σ_k by a

[7] 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 the opposite direction of Young's.

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 proven in Young [1993].

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 the $\tilde{\omega}$-subtree and the other what is left over. This reduces the resistance by $r(\tilde{\omega},\tau(\tilde{\omega}))$. If $\widehat{\omega}$ 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 $\widehat{\omega}$ by defining $\tau(\hat{\omega}) = \widehat{\omega}$. This tree has the root of the tree containing $\widehat{\omega}$. The paste operation increases the resistance by $r(\hat{\omega},\widehat{\omega})$, so the new tree has resistance $r(\tau) + r(\hat{\omega},\widehat{\omega}) - r(\tilde{\omega},\tau(\tilde{\omega}))$. These operations can be used to characterize classes of least resistance trees, by showing certain operation do not increase the resistance. They can also be used as below in proof by contradiction, showing that certain trees cannot be least resistance because it is possible to cut and paste in such a way that the resistance is reduced.

Theorem 1: $\mu = \lim \mu^{\varepsilon}$ *exists and* $\mu(\sigma) > 0$ *implies that* σ *is a pure strategy.*

Proof of Theorem 1: 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 $\tilde{\omega} = \{s\}$ be a singleton pure strategy that is played with positive probability by some $\sigma \in \omega$, that is, $s \in S(\omega)$. Cutting $\tilde{\omega}$ and pasting the root ω to it. Since $\tilde{\omega}$ is a singleton pure profile, it requires at least one innovation to go anywhere, so cutting reduces the resistance by at least n. On the other hand, since $\sigma \in \omega$ and $\sigma(\tilde{\omega}) > 0$, we can go from ω to $\tilde{\omega}$ by no more than m imitations, pasting the root to $\tilde{\omega}$ increases the resistance by at most m. By the assumption of unlikely innovation, this implies that the new tree has strictly less resistance than the old contradicting Young's Theorem.

☑

Theorem 2: *If* m *is sufficiently large and* s *beats the field then* $\mu(s)=1$. *If If* m *is sufficiently large and* s *weakly beats the field then* $\mu(s)>0$. *Moreover, if* $\mu(\tilde{s})>0$ *then* $u(s, \frac{1}{2}s+\frac{1}{2}\tilde{s})-u(\tilde{s}, \frac{1}{2}s+\frac{1}{2}\tilde{s})=0$.

Proof: Consider first the strict case. Suppose that there is some other ω with $\mu(\omega)>0$. By Theorem 1, $\omega=\{\hat{s}\}$ for some pure strategy $\hat{s}$. Let τ be the least resistance ω-tree. Since it is not the root, we may suppose that $\{s\}$ is attached to some $\tilde{\omega}$, and consider cutting it and pasting the root to it. It took at least one innovation plus, since s beats any point in $\tilde{\omega}$, more than $m/2$ imitations to get to $\tilde{\omega}$, so the resistance is reduced by strictly more than $n+m/2$. However, since s beats $\hat{s}$ we can get from $\omega=\{\hat{s}\}$ to $\{s\}$ with one innovation and no more than $m/2$ imitations. So resistance is strictly reduced contradicting Young's Theorem.

In the weak case, we can only conclude that the resistance is not increased, however, this implies that $\{s\}$ is at the root of a least-cost tree, which gives $\mu(\{s\})>0$. If $\mu(\{\tilde{s}\})>0$ and $u(s, \frac{1}{2}s+\frac{1}{2}\tilde{s})-u(\tilde{s}, \frac{1}{2}s+\frac{1}{2}\tilde{s})\neq 0$ then $\tilde{s}$ is not tied with s, so s weakly beats $\tilde{s}$ by assumption. However, if s weakly beats $\tilde{s}$ and $u(s, \frac{1}{2}s+\frac{1}{2}\tilde{s})-u(\tilde{s}, \frac{1}{2}s+\frac{1}{2}\tilde{s})\neq 0$ then s beats $\tilde{s}$, which by the argument in the previous paragraph implies that $\{\tilde{s}\}$ is not the root of a least cost tree, so $\mu(\{\tilde{s}\})=0$.

☑

Theorem 3: $\mu(s_0)>0$. *If* $\mu(s)>0$ *then* $U(s(y_0),s(y_0))=U(\overline{a},\overline{a})$ *and* $U(s(y),\underline{a})=U(\underline{a},s(y))$ *for all* y *such that* $\pi(y \mid s_0, s)>0$.

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

$$\begin{aligned} u(s_0,\sigma(\alpha))-u(s,\sigma(\alpha)) \geq{} & \alpha U(\overline{a},\overline{a})-(1-\alpha)U(s(y_0),s(y_0)) \\ & +\sum\nolimits_{y\in Y}\{(1-\alpha)U(\underline{a},s(y))-\alpha U(s(y),\underline{a})\}\pi(y \mid s,s_0) \end{aligned}$$

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 $\alpha = 1/2$ and $\alpha = 1$. When $\alpha = 1/2$ we have

$$\begin{aligned} u(s_0, \sigma(\alpha)) - u(s, \sigma(\alpha)) \geq {}& 1/2\left[U(\overline{a}, \overline{a}) - U(s'(y_0), s'(y_0))\right] \\ & + 1/2 \min_{a' \in A} \left[U(\underline{a}, a') - U(a', \underline{a})\right] \end{aligned}$$

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

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

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

$$\begin{aligned} 0 &= u(s_0, \sigma(1/2)) - u(s, \sigma(1/2)) \\ &\geq 1/2\left[U(\overline{a}, \overline{a}) - U(a, a)\right] \\ &\quad + 1/2 \sum_{y \in Y} \{U(\underline{a}, s(y)) - U(s(y), \underline{a})\} \pi(y | s, s_0) \end{aligned}$$

From the definition of $\overline{a}, \underline{a}$ and the expression above, we see that this is possible only if $U(a, a) = U(\overline{a}, \overline{a})$ and $U(s(y), \underline{a}) = U(\underline{a}, s(y))$ for all y with $\pi(y | s, s_0) > 0$.

☑

Theorem 4: $\mu(s) > 0$ if and only if $s(y) = s_0(y)$ for all $y \in Y$.

Proof: Theorem 4 follows from Theorem 5 below since the action that defines $s_0(y)$ is unique.

☑

Theorem 5: *If* s_0 *informationally dominant then,* $\mu(s_0) > 0$. *Moreover, for every strategies* s *with* $\mu(s) > 0$ *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 $\sigma(\alpha) = \alpha s_0 + (1 - \alpha) s$ for some $s \in S$. We must show that $u(s_0, \sigma(\alpha)) - u(s, \sigma(\alpha)) \geq 0$ for $\alpha \in [1/2, 1)$. Since

$u(s_0,\sigma(\alpha)) - u(s,\sigma(\alpha))$ is linear in α it suffices to show that it is non-negative at both $\alpha = 1/2$ and $\alpha = 1$.

We may write the utility difference between s_0 and s at $\alpha = 1/2$ as

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

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 $\alpha = 1$. In the case of Theorem 6, let $q'(y) = \pi(y \mid s,s')$. We may write the utility difference as

$$
\begin{aligned}
&u(s_0,\sigma(1)) - u(s,\sigma(1)) = \\
&\sum_{y\in Y}\{(p(y)(b(s_0(y)) - c(s_0(y)))) - q(y)(b_0(s(y)) - c(s(y)))\} \geq \\
&\sum_{y\in Y}\{(p(y)-q(y))b(s_0(y)) - p(y)c(s_0(y))\}
\end{aligned}
$$

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,

$$u(s_0,\sigma(1)) - u(s,\sigma(1)) \geq \sum_{y\in Y}(p(y)-q(y))b(\hat{a}) = 0$$

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

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

$$u(s_0, \sigma(½)) - u(s, \sigma(½)) =$$
$$\frac{1}{2}\sum_{y\in Y}\{[\,p(y) - q(y)\,]b(s_0(y)) - (p(y) + q(y))c(s_0(y))$$
$$-\sum_{z\in Y}\lambda_{zy}\,([\,p(y) - q(y)\,]b(s(z)) - (p(y) + q(y))c(s(z)))\} = 0$$

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

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

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.

Blackwell (1954) (…)

Dawid, H. [1999]: *Adaptive Learning by Genetic Algorithms*, (Berlin: Springer).

Gul, F. and W. Pesendorfer [2001]: "Interdependent Preferences", work in progress.

Ellison, G. [2000]: "Basins of Attraction, Long-Run Stochastic Stability, and the Speed of Step-by-Step Evolution," Review of Economic Studies, 67:17-47.

Frank, Robert H. [1987], "If Homo Economicus Could Choose His Own Utility Function, Would He Want One with a Conscience? " *American Economic Review*, 77: 593-604.

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 [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.

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.

Morris, S., R. Rob and H. Shin [1993]: “p-dominance and Belief Potential,” *Econometrica*, 63: 145-158.

Levine, D. K. and W. Pesendorfer [2000]: "Evolution Through Imitation in a Single Population," *UCLA.*

Van Damme, E. and J. W. Weibull [1998]: “Evolution with Mutations Driven by Control Costs,” Tilburg.

Young, P. [1993]: “The Evolution of Conventions,” *Econometrica*, 61: 57-83.