

Evolution and Information in a Gift-Giving Game^{*}

By

Phillip Johnson, (johnsonp@yahoo.com)

Centro de Investigacion Economía, Instituto Tecnológico Autónomo de México,

David K. Levine, (dlevine@ucla.edu)

Department of Economics, UCLA,

and Wolfgang Pesendorfer, (pesendorfer@princeton.edu)

Department of Economics, Princeton University

Corresponding Author: David Levine

8th Floor Bunche Hall

Department of Economics

UCLA

Los Angeles, CA 90095

Fax: (310) 825-3810 or (310) 825-9528; Phone: (310) 825-3810

Email: dlevine@ucla.edu

^{*} We are grateful to financial support from National Science Foundation Grant SBR-93-20695 and the UCLA Academic Senate. We are especially grateful to two anonymous referees and an associate editor, whose comments led to a major rethinking of the model.

Abstract: We examine the stochastic stability of a process of learning and evolution in a gift-giving game. Overlapping generations of players are randomly matched to play the game. They may consult information systems to learn about the past behavior of their opponents. If the value of the gift is smaller than twice the cost, then gifts are not given. If the value of the gift is more than four times the cost, then gifts are exchanged. Moreover, in the stochastically stable equilibrium, a unique information system is selected to support cooperation.

Keywords: *evolution, information, learning, stochastic stability, repeated games.*

List of Symbols

\forall	Σ	\cup	\equiv	\in	$=$	\neq	\leq	\geq
$>$	$<$	$\{$	$\}$	$:=$	Φ	Θ		
α	β	ε	θ	ϕ	λ	μ	σ	ω
π	η							
B	C	F	G	I	K	L	M	R
S	U	W	a	f	g	h	I	k
l	r	s						

1. Introduction

This paper is about the emergence of cooperative behavior as the unique long-run result of learning in a repeated setting. There is a long-standing tension between the theory of repeated games and empirical evidence. The folk theorem asserts that when players are patient any payoff in a large and diverse set is possible in equilibrium, yet, common experience (supported by experimental research) suggests that cooperative or altruistic behavior is typical. Work by Young [1993] and Kandori, Mailath and Rob [1993] suggests that evolutionary forces can lead to unique outcomes in the long run, even in a setting where there are multiple equilibria. The goal of this paper is to apply that theory in the context of a repeated gift-giving game.

Evolution and learning are most easily studied in a setting of repeated interaction within a large population. We follow Kandori [1992] and consider a framework where players are randomly matched to play a one-shot game. Each player can get some information about the past behavior of his opponent. In the terminology of Kandori, this information is distributed by “information systems.”

To make the model tractable, we make a number of specialized assumptions. Agents live for two periods. An equal number of agents in each generation, referred to as “young” and “old,” are alive every period. Each young agent is matched with an old agent to play a simple gift giving game. The gift is worth 1 to the young agent, and $\alpha > 1$ to the recipient.

	Young	
	Give	Keep
Old	$\alpha, -1$	0,0

The players' objective is to maximize the sum of the payoffs over their lifetime. Taking into consideration only his immediate payoff, the young player's dominant strategy is to keep the gift for himself. The overlapping generations environment allows for a connection between actions when young and payoffs when old. A young player who behaves opportunistically may be punished by the young agent he will face in the next period. To implement this punishment, the young agent must have some information about his opponent's history. We consider a class of information systems sufficiently rich to allow altruistic outcomes as equilibria in the game. Each information system assigns players a *red flag* or a *green flag*. The class consists of every information system that assigns flags that depend on the player's action and on his opponent's flag.

We assume that each young player may consult one such information system. Hence a strategy for a young player consists of the choice of an information system together with a rule that specifies for which of his opponent's flag colors he will give the gift. For example, consider following information system:

- if the player gives to a player with a green flag he is assigned a green flag;
- if he does not give to a player with a red flag, he is assigned a green flag;
- in all other cases, the player is assigned a red flag.

This information system together with the rule that gives the gift if (and only if) the opponent has a green flag is referred to as the *green-team-strategy*. Members of the team, that is, players with a green flag, are given the gift whereas non-members are not. To become part of the team (that is, to receive a green flag) a player must either give to a member or not give to (punish) a non-member. Clearly, there is an analogous *red-team-strategy* in which the same information system is used but players give the gift only to opponents with a red flag. We refer to these two strategies as the *team strategies*.

Players are assumed to choose an optimal strategy given their beliefs about other players' behavior. Our model of learning specifies how these beliefs are formed. We assume that beliefs are derived from public observations of the strategies used in past play. Moreover, players assume that their opponents will not change their strategies – at least in the relevant future. Since the pool of observations is the same for every player, all players have the same beliefs.¹ To that model, we add a stochastic error: individual observations in the pool may randomly “mutate.”

Applying the methods of Ellison [1995] we characterize the stochastically stable distribution of outcomes, that is, we examine the long-run behavior of the process as the mutation rate converges to zero. We show the following results:

- Suppose the value of the gift to the old player is less than twice the value of the gift to the young player. In that case, there are altruistic equilibria in the model without mutations, but selfish behavior is the unique stochastically stable outcome.
- When the old players value the gift at least four times more than the young players do, gift-giving is the unique stochastically stable outcome. Gift-giving is supported by the *team strategies*. In the long run, the system is either in a steady state where all

players use the *red team-strategy* or in a steady state where all players use the *green team-strategy*.

- Our conclusion that altruism emerges in the long-run stochastically stable distribution does not mean that the first best is obtained. Because there is noise in the process, punishment takes place with positive probability. Consequently the long-run stochastically stable distribution, while it involves behaving altruistically “most of the time,” is nevertheless Pareto dominated by the non-equilibrium outcome of giving no matter what.

The team strategy is successful because it punishes any deviation. A player who fails to give to a deserving opponent is punished, as is player who gives to an undeserving opponent. For comparison consider the following strategy. The information system assigns the player a green flag if and only if he gives the gift. A player who consults this information system and gives the gift if (and only if) he encounters a green flag plays *tit-for-tat*. This strategy punishes a player for keeping the gift, regardless of the reason. In particular, a player who carries out a deserved punishment is punished. In our environment, tit-for-tat players believe they must carry out punishments a positive fraction of the time. Against other tit-for-tat players this is never optimal and hence tit-for-tat is not a best response to itself.

As a second comparison, consider the information system that assigns a red flag when the player does not give to a player with a green flag; in all other cases it assigns a green flag. A player who consults this information system and gives to players with a green flag uses a *weak-team-strategy*. This strategy is similar to the team strategy in that

team members are altruistic to other team members and punish non-members. Players who punish non-members receive a green flag and are therefore rewarded. However, failure to punish non-members results also in a green flag and hence is not punished. To see why the team strategy is more successful than the weak-team strategy consider a situation where some fraction of the population is playing tit-for-tat. In this case, punishment of non-members is costly since it triggers punishment by players who use tit-for-tat. The weak-team strategy gives its members only a weak incentive to punish non-members whereas the team strategy gives its members a strong incentive to do so. Therefore, the team strategy is more robust to an invasion of players using tit-for-tat than the weak team strategy.

2. The Model

In this section we describe a simple altruism game played between overlapping generations of players. Players live for two periods; in the first period of life players are referred to as young, in the second period as old. Each generation has a constant and large population n . Each young player is randomly matched with an old player, and must choose between a gift of 0 or 1; the old player has no decision to make. If the young player chooses 0, both players receive a payoff of zero; if he chooses 1, he receives a payoff of -1 , while the old player receives a payoff of $\alpha > 1$. We denote the young player's action by $x \in \{0,1\}$. Lifetime utility is the sum of utility over the two periods of life.

The behavior of young players is reported on by information systems. After the player has chosen an action, each information system assigns the player a “red flag” or a “green flag.” Let $\{r, g\}$ be the set of flags. The assigned flag depends on the player's

action and on the flag of the old player he is matched with. Formally, an information system is a map $i : \{0,1\} \times \{r,g\} \rightarrow \{r,g\}$. It is easily checked that there are 16 information systems. We denote the set of information systems by I .

Let $F := \{r,g\}^{16}$ denote the set of flag vectors with $f \in F$ as a generic element. We write f_i for the flag corresponding to information system i . Every old player is characterized by a flag vector $f \in F$. With probability $1-\eta$ this flag vector is equal to the vector assigned to him by the information systems in the previous period. With probability $\eta > 0$ the old player is assigned a random flag vector according to a uniform distribution on F . We assume that the chance a player is assigned a random flag vector is small:

Assumption: $\beta := \alpha(1-\eta) > 1$.

A strategy $s = (a,i)$ consists of the choice of *one* information system $i \in I$ and a map $a : \{r,g\} \rightarrow \{0,1\}$ that assigns an action to a flag. Let S denote the set of strategies and observe that only pure strategies are allowed.

Young players in every period share common beliefs about the current flags of old players and about the distribution of strategies of next period's young players. These beliefs are assumed to be the frequencies of two samples, a sample of strategies and a sample of flag vectors. The sample of strategies is of fixed size k and is denoted by σ ; the sample of flag vectors is of fixed size l and denoted by ϕ . By π_σ and π_ϕ we denote the corresponding frequency distribution. Hence, $\pi_\sigma(s)$ denotes the fraction of strategies in the sample that are equal to s . Below, we describe in detail how these samples are generated.

A young player's *optimal behavior* is determined by maximizing the sum of payoffs of the two periods of his life, given the beliefs π_σ, π_ϕ . In case of a tie he uses a fixed tie-breaking rule that depends only on his beliefs. We denote the optimal behavior by $s(\phi, \sigma)$.

One class of strategies are those that keep the gift regardless of the information. As these strategies perform identically, we refer to them jointly as *the selfish strategy*. For convenience we make the following assumption about the tie-breaking rule.

Tie Breaking Assumption: If the selfish strategy is optimal, it is always used.

Although our theorems remain true without this assumption, as we discuss further below, it simplifies the proofs.

Let Φ denote the collection of flag vectors of old agents. The flag sample ϕ is derived by taking a random sample of l_1 flag vectors from Φ and adding to it a sample of size l_2 in which each flag vector is represented with equal weight. For simplicity, we assume that l_1 and l_2 are multiples of the number of possible flag vectors in F , denoted by $\#F$. Define $l = l_1 + l_2$.

A state is a triple $\theta = (\sigma, \phi, \Phi)$. By Θ we denote the set of states that have the property that ϕ is derived from Φ as described above. Play and beliefs over time are determined from a Markov process M on the state space Θ , which we now describe in detail.

- Given $\theta = (\sigma, \phi, \Phi)$ the next periods strategy sample σ' is generated as follows: first, h randomly chosen observations in the sample σ are replaced with h observations of the optimal behavior $s(\phi, \sigma)$. This yields a preliminary sample σ^* .

Second, each observation in the sample σ^* remains unchanged with probability $1 - \varepsilon$ and has an independent $\varepsilon[(\#S) - 1]^{-1}$ chance of mutating to each other strategy. This yields next period's sample σ' .

- Given $\theta = (\sigma, \phi, \Phi)$ the next period flag vectors of old players, Φ' is generated. For each observation f in Φ , a corresponding observation f' in Φ' is determined as follows:
 - a. With probability $\eta > 0$ the flag vector f' is chosen randomly according to a uniform distribution on F .
 - b. With probability $1 - \eta$ the flag vector f' is determined by the information systems and by the young player's optimal behavior. That is, if $s(\phi, \sigma) = (a^*, i^*)$ then $f'_i = i(f_i, a^*(f_i))$ for all $i \in I$.

By $\Pr(\theta' | \theta)$ we denote probability of transition to state θ' given θ . This Markov process is designed give a concrete representation of the type of process we have in mind. The following Lemma derives the key properties of the flag beliefs that will be used in the subsequent analysis.

Lemma 1: (i) There is $\omega > 0$ such that for every $\theta = (\sigma, \phi, \Phi) \in \Theta$, $\pi_\phi(f) \geq \omega$ for all $f \in F$; (ii) Independent of ε there is $c > 0$ such that for $\theta' = (\bar{\sigma}, \phi', \Phi') \in \Theta$ and $\theta'' = (\bar{\sigma}, \phi'', \Phi'') \in \Theta$, $\Pr(\theta'' | \theta) \geq c \Pr(\theta' | \theta)$.

Proof: Let $\omega = l_2 / (l \cdot \#F)$. Part (i) follows since for every $f \in F$, ϕ contains at least $l_2 / \#F$ flag vectors of type f . For part (ii) observe that any collection of flag vectors of old players, Φ'' , occurs with probability at least $(\eta / \#F)^n$ in every period. The

probability of selecting a sample of l_1 flag vectors that yields ϕ depends on Φ, ϕ only. For each $\theta = (\bar{\sigma}, \phi, \Phi) \in \Theta$ this probability is strictly positive. Since Θ is finite, part (ii) follows.

☑

The key elements of the beliefs about flags are: (1) they are non-doctrinaire, that is, players always believe that there is a minimal probability of meeting an opponent with any flag vector and (2) they are “volatile”, that is, any non-doctrinaire belief may occur. Non-doctrinaire beliefs are guaranteed by assuming that each sample ϕ contains a minimum number of flag vectors of each type. The volatility of beliefs is guaranteed by assuming that flags are recorded with error. This error is independent of the mutation probability ε and is therefore much larger than the mutation probability when we consider the limit, as $\varepsilon \rightarrow 0$. In our formulation, players do not use their beliefs about the distribution of strategies to formulate beliefs about the flag distribution. Rather, players rely only on the sample of current flag vectors. We could allow the strategy beliefs to influence the flag beliefs if at the same time we assumed that the flag sample is sufficiently large so that the distribution of strategies has only a small impact on the beliefs about the flag distribution.

Other key elements of the Markov process are: (1) beliefs about strategies adjust slowly to the actual distribution of strategies; and (2) strategies of young players are observable, not merely their actions. Assumption (1) is relatively standard and rules out the possibility of cobweb-type cycles. The observability of strategies is a significant assumption, and we discuss its merits in the conclusion.

The Markov process M on the state space Θ has been completely specified. To analyze the long-run dynamics of M on Θ , note that M is strictly positive since each flag sample and strategy sample has positive probability. It follows that the process L is ergodic with a unique stationary distribution μ^ε . Moreover, the transition probabilities are polynomials in ε . Consequently we can apply Theorem 4 from Young [1993] and conclude that $\lim_{\varepsilon \rightarrow 0} \mu^\varepsilon$ exists. We denote this limit as μ and refer to it as the *stochastically stable distribution*. From Young's theorem, this distribution places weight only on states that have positive weight in stationary distributions of the transition matrix for $\varepsilon = 0$. Our goal is to characterize the stochastically stable distribution for several special cases using methods developed by Ellison [1995].

3. Pure Steady States

According to Young's Theorem, the long-run behavior of the system for small ε is a subset of the long-run behavior of the system for $\varepsilon = 0$. Our first objective is to study the system in the latter case. We will not attempt to find all possible stationary distributions of M for $\varepsilon = 0$, but rather give some specific examples. Our goal is to illustrate that this game, like repeated public goods games more generally, has many equilibria, some cooperative and some not.

We will focus on a particular class of stationary distributions, which we refer to as steady states. In a *steady state*, beliefs about strategies do not change over time. Notice also that such a steady state is a Nash equilibrium in the ordinary meaning of the term: each player's strategy is optimal given the actual strategies played by opposing players.

The most obvious steady state is the one in which all players are believed to use the selfish strategy. Given that regardless of current play opponents are believed to be selfish next period, it is optimal for each player to be selfish now, and avoid the cost of the gift. As a result, next period beliefs about strategies will be unchanged.

There are, however, steady states in which players give the gift. Recall the red and green team strategies from the introduction. In these strategies the color of the flag can be interpreted as a “team” that the player is on. Formally, the team-strategies use the following information system:

$$i(g,1) = i(r,0) = g; i(r,1) = i(g,0) = r.$$

The green team strategy takes the actions $a(g) = 1, a(r) = 0$, and the *red-team strategy* takes the actions $a(r) = 1, a(g) = 0$. Thus, players give to team members but do not give to non-members. To be part of the team a player must follow this strategy. Any deviation leads to a flag that identifies the player as not part of the team. Therefore, any deviation from a team strategy yields at most 1 in the current period, but loses β next period against a player who uses the same team strategy. Since $\beta > 1$ each team strategy is therefore a best response to itself, regardless of the flag beliefs.

The team strategies are not the only pure steady states with gift-giving. For example, the team strategy rewards players for selfishness against a non-team member. This does not appear necessary as selfish behavior already provides an immediate benefit to a player. There are strategies we refer to as *weak team strategies* in which young players who meet a non-member always become part of the team regardless of their behavior and hence receive the gift from a team member when they are old. Formally, the green-weak-team-strategy uses the information system

$$i(g,0) = r; i(r,0) = g; i(r,1) = i(g,1) = g$$

and takes the actions $a(r) = 0, a(g) = 1$. Like the team strategies this strategy it is a best response to itself.

In summary, the picture here is very much like that of the usual folk theorem for repeated games. There are many possibilities ranging from complete selfishness to altruism. The theory for $\varepsilon = 0$ allows the possibility of altruism but provides no theory of whether or why it might occur.

4. Long-Run Behavior

We now ask whether evolutionary arguments can pick out a particular steady state. Our main results give sufficient conditions for the unique stochastically stable distribution to be either all players selfish, or one of the two team strategies. First we give conditions under which players are selfish in the stochastically stable distribution. Let $\Theta^S \subset \Theta$ denote the set of states where the strategy sample consists only of selfish strategies.² Proposition 1 says that if the gains to altruism are small then the stochastically stable distribution places all its weight on states in Θ^S .

Proposition 1: *If $\beta < 2$ then $\mu(\Theta^S) = 1$.*

Proof: Let $\Theta^{S/2}$ denote the set of states where the strategy sample has at least half the players playing a selfish strategy. From Ellison [1995] it is sufficient to show that at $\Theta^{S/2}$ the optimal behavior is to be selfish. At $\Theta^{S/2}$ the choice 0 earns an immediate payoff of 0. If the player chooses 1, his immediate payoff is -1 and since there is a $1/2$ chance of meeting an always-selfish opponent, the maximal gain when old is $\beta/2$. Since $\beta/2 < 1$ the optimal behavior is to choose 0.

☑

Our second result describes a range of parameters for which altruism will occur in the stochastically stable distribution. Let Θ^R denote the set of states where the strategy sample consists only of the red team strategy, and Θ^G be the states where the strategy sample consists only of the green team strategy. Let $\lambda = l_2 / l$ be the fraction of the flag sample that is the uniform portion.

Proposition 2: *If $1 - 2\lambda - 4(h + 4)/k > 0$ and*

$$\beta \geq \frac{4}{1 - 2\lambda - 4(h + 4)/k}$$

then $\mu(\Theta^R) = \mu(\Theta^G) = 1/2$.

We will prove this below. Notice that the bound given implies $\beta > 4$. Moreover, for any $\beta > 4$, we can find k is large enough, and h/k and λ are close enough to zero that the bound holds. Then Proposition 2 implies that for $\beta > 4$ cooperation will emerge and all players will use the same team strategy in the stochastically stable distribution.

5. Proof of Proposition 2

We begin with a discussion of different types of stage game strategies. With probability η an old player's flag vector is randomly drawn and independent of his behavior in the previous period. With probability $1 - \eta$ his flag vector is determined by his action and his opponents flag in the previous period. In that case, the *reaction* of the strategy $s = (a, i)$ to a player who chooses x when his opponent's flag was f is

$a(i(x, f_i))$. We define a *strong* strategy to be a strategy for which the reaction always depends on the opponent's previous action. The following definition makes this precise.

Definition. A strategy $s = (a, i)$ is strong if $a(i(1, j)) \neq a(i(0, j))$ for all $j \in \{r, g\}$.

Strategies that are not strong are called weak: S^S denotes the set of strong strategies and S^W denotes the weak strategies.

In total there are 8 strong strategies. Every strong strategy consists of one of the following four information systems.

Team: i_t	$i_t(1, g) = g$	$i_t(0, g) = r$	$i_t(1, r) = r$	$i_t(0, r) = g$
Tit-for-tat: i_g	$i_g(1, g) = g$	$i_g(1, r) = g$	$i_g(0, g) = r$	$i_g(0, r) = r$
Tit-for-tat: i_r	$i_r(1, g) = r$	$i_r(1, r) = r$	$i_r(0, g) = g$	$i_r(0, r) = g$
Anti-team: i_a	$i_a(1, g) = r$	$i_a(1, r) = g$	$i_a(0, g) = g$	$i_a(0, r) = r$

Let a_r be defined as $a_r(g) = 0, a_r(r) = 1$ and let $a_g = 1 - a_r$. The strong strategies can be obtained by combining any of the four information systems above with a_r or a_g . Thus, (a_r, i_t) and (a_g, i_t) are the red and green team strategies; (a_g, i_g) and (a_r, i_r) are red and green tit-for-tat; (a_g, i_r) and (a_r, i_g) are red and green tat-for-tit, that is, strategies that reward opportunistic behavior but punish gift giving; finally, (a_r, i_a) and (a_g, i_a) are red and green anti-team, that is, strategies that punish members of the team but reward non-team members.

If a player expects to face a strong strategy $s = (a, i)$ in the following period then, for each flag f_i , there is a unique choice of current action $a^*(f_i)$ that yields an expected payoff of β next period. We define $B(s)$ to be the strategies that take that action for every flag vector with i th component f_i . If the unique action is independent of f_i , $a^*(f_i) = a$, $B(s)$ consists of all strategies that choose a independent of the flag vector. Note that all of these strategies have the same reaction. If the unique action depends on f_i then $B(s)$ contains the single strategy (a^*, i) . We somewhat loosely refer to $B(s)$ as the best response to s .

For the state $\theta = (\sigma, \phi, \Phi)$ we can define the *gain function* G of using $s = (a, i)$ instead of $\hat{s} = (\hat{a}, \hat{i})$ as follows:

$$G(s, \hat{s}, \theta) \equiv \sum_f \left(\hat{a}(f_i) - a(f_i) + \beta \sum_{(a', i') \in S} \left(a'(i'(a(f_i), f_{i'})) - a'(i'(\hat{a}(f_i), f_{i'})) \right) \pi_\sigma(a', i') \right) \pi_\phi(f)$$

To understand this expression, note that the term $\hat{a}(f_i) - a(f_i)$ is the expected difference in the first period payoff if the opponents flag is f . Recall that $\beta = \alpha(1 - \eta)$ is the value of the gift times the probability that the flag is assigned by the information systems (rather than randomly). To obtain the expected payoff difference in the second period, we multiply β by the difference between using a and \hat{a} in the probability of receiving the gift. When flags are assigned at random, the second period flag vector is independent of the action taken in the first period. In that case, the player's payoff is independent of the

action and hence these terms drop out. We conclude that the optimal behavior at θ is s if $G(s, \hat{s}, \theta) \geq 0$ for all $\hat{s} \neq s$.

Lemma 2: Suppose that \bar{s} is a strong strategy. If $s \in B(\bar{s})$ and $\hat{s} \notin B(\bar{s})$ then

$$G(s, \hat{s}, \theta) \geq \omega(-1 + (2\pi_\sigma(\bar{s}) - 1)\beta),$$

provided the RHS is non-negative.

Proof: Let $s = (a, i), \hat{s} = (\hat{a}, \hat{i})$. Let f be such that $\hat{a}(f_i) \neq a(f_i)$. Then, since \bar{s} is strong,

$$\begin{aligned} & \hat{a}(f_i) - a(f_i) + \beta \sum_{s'=(a',i') \in \mathcal{S}} \left(a'(i'(a(f_i), f_{i'})) - a'(i'(\hat{a}(f_i), f_{i'})) \right) \pi_\sigma(s') \\ & \geq -1 + \beta\pi_\sigma(\bar{s}) - \beta(1 - \pi_\sigma(\bar{s})) \\ & = -1 + \beta(2\pi_\sigma(\bar{s}) - 1) \end{aligned}$$

Since $s \in B(\bar{s}), \hat{s} \notin B(\bar{s})$, there is a flag vector f such that $a(f_i) \neq \hat{a}(f_i)$ and, by Lemma 1, $\pi_\sigma(f) \geq \omega$. Since $-1 + \beta(2\pi_\sigma(\bar{s}) - 1) \geq 0$ the lower bound follows.

☑

Team strategies are the only strong strategies that are best responses to themselves.

Lemma 3: If s is a strong strategy and $s \in B(s)$ then s is either the red or green team strategy.

Proof: Let $s = (a, i)$ and suppose that $a = a_g$. Suppose a young player using s meets f with $f_i = g$. Then, since s is a best response to itself and chooses 1, it must be in the expectation of receiving a green flag, and in the expectation that 0 will result in a red flag. Since s is strong, if the player meets a red flag, he must expect to get a green flag

for 0 and a red flag for 1. This uniquely defines the information system i_t and hence $s = (a_g, i_t)$, the green team-strategy. An analogous argument for $a = a_r$ yields (a_r, i_t) .

☑

From the Lemma 2 and 3, if s is one of the team strategies, and

$$\pi_\sigma(s) > \frac{1}{2} + \frac{1}{2\beta} \equiv 1 - \frac{R}{k},$$

then the optimal behavior of all players is to play s . This defines R – the *radius* of s – as defined in Ellison [1995]. Note that

$$\frac{R}{k} = \frac{1}{2} \left(\frac{\beta - 1}{\beta} \right).$$

The interpretation of the radius is that if the state is in $\Theta^R \cup \Theta^G$ fewer than R mutations followed by a sufficiently long period with no mutations will return the system to $\Theta^R \cup \Theta^G$.

Ellison [1995] shows that a state is in the stochastically stable distribution if the radius is larger than the *co-radius*. The co-radius is calculated by finding for each ergodic class with $\varepsilon = 0$ the minimum number of mutations needed to get back to $\Theta^R \cup \Theta^G$, then maximizing over ergodic classes. We complete the proof of Proposition 2 by demonstrating that, under the hypothesis of Proposition 2, the co-radius of $\Theta^R \cup \Theta^G$ is smaller than the radius, R .

The proof will utilize the fact that if $\bar{\mu}$ is an ergodic distribution for $\varepsilon = 0$ with $\bar{\mu}(\theta) > 0$, $\theta = (\sigma, \phi, \Phi)$, then, by Lemma 1(ii), $(\sigma, \bar{\phi}, \bar{\Phi}) \in \Theta$ is in the same ergodic class as θ . Furthermore, if $\bar{\theta}$ is such that $\Pr(\bar{\theta} | (\sigma, \bar{\phi}, \bar{\Phi})) > 0$ then also $\bar{\theta}$ is in the

same ergodic class as θ . So in computing the co-radius, we can limit attention to states with “favorable” flag distributions.

Favorable flag distributions make some information system independent in the sense that the believed probability of the realization of flags from that information system is independent of the realization of flags from other information systems. More precisely, consider the random variables $f_i \in \{r, g\}$, where $\Pr(f_i = r) = \pi_\phi(\{f : f_i = r\})$. The information system i is *independent at ϕ* if the random variable f_i is independent of $f_j, \forall j \neq i$. For example, in the sample that has an equal number of flag vectors of each type, every information system is independent.

Lemma 4: Suppose $\bar{\mu}$ is a stationary distribution of M for $\varepsilon = 0$. For any state $\theta = (\sigma, \phi, \Phi)$ with $\bar{\mu}(\theta) > 0$ there exists a state $\bar{\theta}$ with $\bar{\mu}(\bar{\theta}) > 0$ such that (i) if $s = (a, i)$ is a tit-for-tat or a tat-for-tit strategy then $\pi_{\bar{\sigma}}(s) = 0$; (ii) if \hat{s} is an anti-team strategy then

$$\pi_{\bar{\sigma}}(\hat{s}) \leq \frac{1}{2} + \frac{1}{2\beta} + h/k$$

Proof: Let $\bar{\Phi}$ be the sample with an equal number of flag vectors of each type, and $\bar{\phi}$ the corresponding beliefs. Let $\bar{\theta} = (\sigma, \bar{\phi}, \bar{\Phi})$.

We first show that at $\bar{\theta}$ the strategy s is not chosen by any player if it is tit-for-tat or tat-for-tit. Let $s = (a, i)$ be green tit-for-tat strategy. Notice that all information systems, and i in particular, are independent at $\bar{\phi}$.

Suppose that s is optimal. Consider the expected utility conditional on the opponent flag being green, u_g , and conditional on it being red, u_r . The tit-for-tat information system does not depend on flag of the opposing player and, by independence of $\bar{\phi}$, the probability of opponent flags f_j for $j \neq i$ do not depend on the opponent flag f_i . Consequently, given the action that is taken, second period expected utility does not depend on the first period opposing player value of f_i . Of course, first period utility depends only on the action taken, so expected utility given the action taken is independent of the flag. It follows that $u_g = u_r$, since if $u_g > u_r$ always altruistic would be better than green tit-for-tat, and conversely if $u_r > u_g$ always-selfish would be better. However, since $u_r = u_g$, the always-selfish strategy does exactly as well as green tit-for-tat, and therefore, the tie breaking assumption implies the selfish strategy is chosen.

A similar argument holds if s is red tit-for-tat or either tat-for-tit. Now observe that there is a strictly positive probability that $\bar{\phi}$ recurs for k/h consecutive periods. Therefore, with strictly positive probability all observations of s are removed from the strategy sample σ and replaced with always-selfish, resulting in a $\bar{\theta}$ satisfying (i).

To prove (ii), note that, by Lemma 3, the anti-team strategy is not a best response to itself. Hence, by Lemma 2, if \hat{s} is an anti-team strategy and

$$\pi_{\sigma}(\hat{s}) > \frac{1}{2} + \frac{1}{2\beta}$$

then all players choose a best response to \hat{s} . Since h of the k observations are replaced each period, part (ii) follows. ☑

Remark: Lemma 4 is the only part of the proof that uses the tie breaking assumption. We use it to break ties between tit-for-tat and tat-for-tit and the selfish strategy. However, we could also break the tie by introducing a single mutation that favored the selfish strategy (that is, one player switches to the selfish strategy). This would increase the calculation of the co-radius by 1 in what follows, with an insubstantial change in the final result.

From Lemma 4, we may restrict attention to “favorable” states satisfying (i) and (ii). Starting at such a “favorable” state, we now define even “more favorable” states that places a lower bound on the gain to switching to a team strategy.

Lemma 5: Suppose $\beta > 4$. Let s, \tilde{s} be the team strategies, $\hat{s}, \hat{\tilde{s}}$ be the anti-team strategies. Suppose σ has no observations of tit-for-tat or tat-for-tit and $\pi_\sigma(\hat{s}) \geq \pi_\sigma(\hat{\tilde{s}})$.

There exist $\bar{\phi}, \bar{\Phi}$ such that $\bar{\theta} = (\sigma, \bar{\phi}, \bar{\Phi}) \in \Theta$ and

$$G(s, \hat{s}, \bar{\theta}) \geq \frac{1}{2} \left(-1 + \beta (\pi_\sigma(s) - \pi_\sigma(\tilde{s}) - \lambda \pi_\sigma(S^w)) \right) + \left(1 - \frac{\lambda + \beta}{2} \right) (\pi_\sigma(\hat{s}) - \pi_\sigma(\hat{\tilde{s}})).$$

Proof: Notice that the information system of a weak strategy has the property that for one of the two flag realizations, the action of the young player does not affect the assigned flag. Let \check{f} be such that for every weak information system i , \check{f}_i has the property that the action does not influence the flag assigned to the player. Suppose, in

addition, that for the anti-team information system $i = i_a, \check{f}_i$ is such that gift giving is a best response to \hat{s} , the more frequent anti-team strategy. Let \check{F} denote the set of flag vectors \check{f} for which $\check{f}_i = \check{f}_i, i \neq i_t, i_r, i_g$. Let $\bar{\phi}_1$ be a sample of size l_1 , which is uniform over \check{F} . Let $\bar{\phi}$ be the sample of size $l_1 + l_2$ obtained by adding to $\bar{\phi}_1$ a sample of size l_2 , which is uniform over F . The sample $\bar{\phi}$ has the following properties.

$$(1) \pi_{\bar{\phi}}(\{f': f_i' = g\}) = 1/2 ;$$

$$(2) \pi_{\bar{\phi}}(\{f': f_i' \neq \check{f}_i\}) = \frac{l_2}{2l} = \frac{\lambda}{2} \text{ for } i \notin \{i_t, i_r, i_g\};$$

(3) The information systems i_t, i_r, i_g are independent.

(4) There is $\bar{\Phi}$ such that $(\sigma, \bar{\phi}, \bar{\Phi}) \in \Theta$.

Let $s = (a, i_t), \hat{s} = (\hat{a}, \hat{i})$. We proceed to calculate a lower bound for $G(s, \hat{s}, \theta)$. Let

$$U((a, i), \theta) = \sum_f \left(-a(f_i) + \beta \sum_{(a', i') \in S} a'(i'(a(f_i), f_i')) \pi_{\sigma}(a', i') \right) \pi_{\bar{\phi}}(f).$$

Then,

$$G(s, \hat{s}, \theta) = U((a, i_t), \theta) - U((\hat{a}, \hat{i}), \theta)$$

First, we find a lower bound for $U((a, i_t), \theta)$. By (1) above, half of the i_t flags are green and half are red. That means that s gives the gift half the time, and therefore the payoff in the first period is

$$\sum_f (-a(f_i)) \pi_\phi(f) = -1/2.$$

If the opponent also chooses s , that is, $(a', i') = s$, then the second period payoff is

$$\sum_f a' (i' (a(f_i), f_{i'})) \pi_\phi(f) = 1.$$

If (a', i') is anti-team, since i_i is independent and s gives a gift in $1/2$ of the matches, s must receive a gift from either anti-team strategy in at least $1/2$ the matches, and the second period payoff is

$$\sum_f a' (i' (a(f_i), f_{i'})) \pi_\phi(f) \geq 1/2.$$

By assumption σ has no tit-for-tat or tat-for-tit observations, so that leaves only the weak strategies, and the second period payoff is

$$\sum_f \left(\sum_{(a', i') \in S^W} (a' (i' (a(f_i), f_{i'}))) \pi_\sigma(a', i') \right) \pi_\phi(f).$$

Adding terms we get

$$U((a, i), \theta) \geq -1/2 + \beta \left(\pi_\sigma(s) + \frac{\pi_\sigma(\hat{s}) + \pi_\sigma(\hat{\hat{s}})}{2} + \sum_f \left(\sum_{(a', i') \in S^W} (a' (i' (a(f_i), f_{i'}))) \pi_\sigma(a', i') \right) \right) \pi_\phi(f).$$

From this we subtract an upper bound for $U((\hat{a}, \hat{i}), \theta)$. We may assume that the player is told whether or not he will face an anti-team strategy in the second period of life before choosing his own strategy: clearly this cannot lower payoffs, and it turns out to simplify computations. Observe that $\pi_\sigma(\hat{s}) \geq \pi_\sigma(\hat{\hat{s}})$, and recall that both \hat{s} and $\hat{\hat{s}}$ both use the same information system i_a . Consequently, knowing that an anti-team opponent

will be faced next period, it is always optimal to use $i = i_a$. If the flag f_i is such that \hat{s} will not give a gift next period, it is clearly optimal not to give a gift. This means that the only two possible optimal strategies are either to play some $s^* \in B(\hat{s})$ or to play the always-selfish strategy. We will calculate an upper bound on the payoff to \hat{s} for each case.

If in the first period always-selfish is employed against anti-team, then

$$\sum_f (-\hat{a}(f_i)) \pi_{\bar{\phi}}(f) = 0.$$

If s^* is employed against anti-team, we chose \check{f}_i so that a gift is given by s^* . The probability of \check{f}_i is equal to $1 - \lambda/2$. This follows since with probability $1 - \lambda$ the flag is from $\bar{\phi}_1$ (in which case \check{f}_i is chosen with probability 1) and with probability λ the flag is from a uniform distribution over all flags (in which case \check{f}_i is chosen with probability $1/2$). Consequently, if a best-response to \hat{s} is used against \hat{s} ,

$$\sum_f (-\hat{a}(f_i)) \pi_{\bar{\phi}}(f) \leq -(1 - \lambda/2) (\pi_{\sigma}(\hat{s}) + \pi_{\sigma}(\hat{\hat{s}})).$$

Now consider the second period, and suppose that a', i' is anti-team. The always-selfish strategy gives

$$\begin{aligned} \sum_f a'(i'(\hat{a}(f_i), f_i)) \pi_{\bar{\phi}}(f) &\leq (1 - \lambda) \pi_{\sigma}(\hat{\hat{s}}) + \lambda (\pi_{\sigma}(\hat{s}) + \pi_{\sigma}(\hat{\hat{s}})) / 2 \\ &= (1 - \lambda/2) \pi_{\sigma}(\hat{\hat{s}}) + (\lambda/2) \pi_{\sigma}(\hat{s}) \end{aligned}$$

If $s^* \in B(\hat{s})$ then

$$\sum_f a'(i'(\hat{a}(f_i), f_i)) \pi_{\bar{\phi}}(f) \leq \pi_{\sigma}(\hat{s})$$

Now suppose that (a', i') is a team strategy. Since \hat{s} is not itself a team strategy, the information system i_t is not used, and since by (3) i_t is independent, and by (2) half the i_t flags are red and half green, \hat{s} receives a gift at most half the time from the team strategies, so

$$\sum_f a' \left(i' \left(\hat{a}(f_i), f_{i'} \right) \right) \pi_\phi(f) \leq 1/2.$$

Finally, suppose that a', i' is weak. For the flag $\check{f}_{i'}$, i' does not depend on the player's action, so \hat{a} receives exactly the same payoff as a . The flag $f_{i'} \neq \check{f}_{i'}$ occurs with probability $\lambda/2$ (if the flag is drawn from the uniform sample, $f_{i'} \neq \check{f}_{i'}$ has probability $1/2$). Consequently, we have the upper bound

$$\begin{aligned} & \sum_f \left(\sum_{(a', i') \in S^W} \left(a' \left(i' \left(\hat{a}(f_i), f_{i'} \right) \right) \right) \pi_\sigma(a', i') \right) \pi_\phi(f) \\ & \leq \lambda/2 + \sum_f \left(\sum_{(a', i') \in S^W} \left(a' \left(i' \left(a(f_{i_t}), f_{i'} \right) \right) \right) \pi_\sigma(a', i') \right) \pi_\phi(f) \end{aligned}$$

Adding terms, we get

$$\begin{aligned} U((\hat{a}, \hat{i}), \theta) & \leq (1/2)(\pi_\sigma(s) + \pi_\sigma(\tilde{s}))\beta \\ & + \max\{(1 - \lambda/2)\pi_\sigma(\hat{s})\beta + (\lambda/2)\pi_\sigma(\check{s})\beta, -(1 - \lambda/2)[\pi_\sigma(\hat{s}) + \pi_\sigma(\check{s})] + \pi_\sigma(\hat{s})\beta\} \\ & + (\lambda/2)\pi_\sigma(S^W)\beta + \sum_f \left(\sum_{(a', i') \in S^W} \left(a' \left(i' \left(a(f_{i_t}), f_{i'} \right) \right) \right) \pi_\sigma(a', i') \right) \pi_\phi(f)\beta \end{aligned}$$

Subtracting this from the lower bound on the payoff from s , we get

$$\begin{aligned}
G(s, \hat{s}, \theta) &\geq -\frac{1}{2} + \left(\pi_\sigma(s) + \frac{\pi_\sigma(\hat{s}) + \pi_\sigma(\tilde{s})}{2} \right) \beta - \frac{1}{2} (\pi_\sigma(s) + \pi_\sigma(\tilde{s})) \beta \\
&\quad - \max \left\{ \left(1 - \frac{\lambda}{2} \right) \pi_\sigma(\hat{s}) \beta + \frac{\lambda}{2} \pi_\sigma(\tilde{s}) \beta, - \left(1 - \frac{\lambda}{2} \right) [\pi_\sigma(\hat{s}) + \pi_\sigma(\tilde{s})] + \pi_\sigma(\hat{s}) \beta \right\} \\
&\quad - \frac{\lambda}{2} \pi_\sigma(S^W) \beta \\
&\geq -\frac{1}{2} + \frac{1}{2} \pi_\sigma(s) \beta - \frac{1}{2} \pi_\sigma(\tilde{s}) \beta - \frac{\lambda}{2} \pi_\sigma(S^W) \beta \\
&\quad - \max \left\{ 0, - \left(1 - \frac{\lambda}{2} \right) [\pi_\sigma(\hat{s}) + \pi_\sigma(\tilde{s})] + \frac{\pi_\sigma(\hat{s}) - \pi_\sigma(\tilde{s})}{2} \beta \right\} \\
&\geq -\frac{1}{2} + \frac{1}{2} \pi_\sigma(s) \beta - \frac{1}{2} \pi_\sigma(\tilde{s}) \beta - \frac{\lambda}{2} \pi_\sigma(S^W) \beta \\
&\quad + \left(1 - \frac{\lambda}{2} - \frac{\beta}{2} \right) [\pi_\sigma(\hat{s}) - \pi_\sigma(\tilde{s})]
\end{aligned}$$

□

Proof of Proposition 2: Recall that the radius is the number of mutations required to escape the basin of one of the team strategies s , and we argued above that it is given by

$$\frac{R}{k} = \frac{1}{2} \left\lfloor \frac{\beta - 1}{\beta} \right\rfloor.$$

From Ellison [1995], to prove Proposition 2, we need to show under the hypothesis of the Proposition that the co-radius is less than the radius. The co-radius is calculated by finding for each ergodic class with $\varepsilon = 0$ the minimum number of mutations needed to get back to $\Theta^R \cup \Theta^G$, then maximizing over ergodic classes. Define C by

$$\frac{C}{k} = \frac{1}{2} \left\lfloor \frac{\beta - 1}{\beta} \right\rfloor - \frac{2}{k}.$$

This accounts for the fact that R and C may not be integers, so that the least integer greater than or equal C is smaller than the greatest integer less than or equal R . To

demonstrate the Proposition, for each ergodic class with $\varepsilon = 0$, we must find a member of that ergodic class for which C mutations returns us to the basin of $\Theta^R \cup \Theta^G$. From Lemma 4, we may assume that we begin at a state $\theta = (\sigma, \phi, \Phi)$ such that σ has no observations of tit-for-tat, and if \hat{s} is an anti-team strategy then

$$\pi_\sigma(\hat{s}) \leq \frac{1}{2} + \frac{1}{2\beta} + h/k \equiv \hat{\pi}.$$

Without loss of generality, we let s, \tilde{s} be the team strategies, and $\hat{s}, \hat{\tilde{s}}$ the anti-team strategies with $\sigma(s) \geq \sigma(\tilde{s}), \sigma(\hat{s}) \geq \sigma(\hat{\tilde{s}})$. From Lemma 5, we may assume that we begin at a flag distribution $\bar{\phi}, \bar{\Phi}$ such that the bound in that Lemma holds.

We now suppose that there are C mutations each resulting in s being played, and such that the strategy that mutates is \hat{s} as until there are equal numbers of $\hat{s}, \hat{\tilde{s}}$, and following which the number of \hat{s} remaining is never less than the number of $\hat{\tilde{s}}$. In addition, we suppose that the flag distribution remains $\bar{\phi}, \bar{\Phi}$. Let $\bar{\sigma}$ denote the strategy distribution following the mutations. This means that the bound in Lemma 5 applies. We will use this bound show that $G(s, \hat{s}, (\bar{\sigma}, \bar{\phi}, \bar{\Phi})) \geq 0$ for all \hat{s} . Following this, the $\varepsilon = 0$ dynamic assigns positive probability to the flag distribution remaining unchanged, all strategies being changed to optimal play in a finite amount of time, and the strategies \hat{s} changing to optimal play before any other strategies change. Since the optimal play with $G \geq 0$ is to play s , examination of the bound in Lemma 5 shows that G remains positive as each strategy is changed to optimal play, so this implies that there is positive probability of reaching $\Theta^R \cup \Theta^G$ which shows that C is indeed the coradius.

It remains to show that that $G(s, \hat{s}, (\bar{\sigma}, \bar{\phi}, \bar{\Phi})) \geq 0$ for all \hat{s} . From Lemma 5

$$G(s, \hat{s}, \bar{\theta}) \geq -1/2 + (1/2)\pi_{\bar{\sigma}}(s)\beta - (1/2)\pi_{\bar{\sigma}}(\hat{s})\beta - (\lambda/2)\pi_{\bar{\sigma}}(S^w)\beta + (1 - \lambda/2 - \beta/2)\left[\pi_{\bar{\sigma}}(\hat{s}) - \pi_{\bar{\sigma}}(\widehat{\hat{s}})\right]$$

From the description of the mutations, we have $\pi_{\bar{\sigma}}(s) \geq \pi_{\bar{\sigma}}(\hat{s}) + (C/k)$, and since $\pi_{\bar{\sigma}}(\hat{s}) \leq \widehat{\pi}$, $(C/k) \leq 1/2$ we also have $\pi_{\bar{\sigma}}(\hat{s}) - \pi_{\bar{\sigma}}(\widehat{\hat{s}}) \leq \widehat{\pi} - (C/k)$. Consequently, we may find a lower bound on the RHS, to conclude that

$$G(s, \hat{s}, \bar{\theta}) \geq -1/2 + (1/2)(C/k)\beta - (\lambda/2)(1 - \widehat{\pi})\beta + (1 - \lambda/2 - \beta/2)\left[\widehat{\pi} - (C/k)\right].$$

Plugging in for C and $\widehat{\pi}$ we have that

$$\begin{aligned} G(s, \hat{s}, \bar{\theta}) &\geq -\frac{1}{2} + \left(\frac{\beta-1}{4\beta} - \frac{1}{k}\right)\beta - \frac{\lambda}{2}\left(1 - \left(\frac{1}{2} + \frac{1}{2\beta} + \frac{h}{k}\right)\right)\beta \\ &\quad + \left(1 - \frac{\lambda}{2} - \frac{\beta}{2}\right)\left[\left(\frac{1}{2} + \frac{1}{2\beta} + \frac{h}{k}\right) - \left(\frac{\beta-1}{2\beta} - \frac{2}{k}\right)\right] \\ &\geq \frac{\beta}{4} + \frac{1}{\beta} - \frac{5}{4} - \frac{\lambda}{2}\left(\frac{\beta}{2} - \frac{1}{2}\right) - \left(\frac{\beta}{2} - 1\right)\left(\frac{h+2}{k}\right) - \frac{\beta}{k} - \frac{\lambda}{2}\left(\frac{1}{\beta} + \frac{h+2}{k}\right) \\ &\geq \frac{(\beta-1)(\beta-4)}{4\beta} - (\beta-1)\left[\frac{\lambda}{2} + \frac{h+2}{k} + \frac{2}{k}\right] \end{aligned}$$

where the last inequality uses the fact that $\beta \geq 4$ and $(1/\beta + (h+2)/k) < 1$.

Collecting terms we get

$$G(s, \hat{s}, \bar{\theta}) \geq \frac{\beta-1}{4\beta}(\beta[1 - 2\lambda - 4(h+4)/k] - 4)$$

which is non-negative under the hypothesis of the Proposition, yielding the desired result.

6. Conclusion

We examine a gift-giving game with a particular learning process with noise in both the transmission of social and individual information about past play. In this setting, we show that for some parameter values cooperative play emerges as the unique long-run stochastically stable distribution. We conclude by examining the robustness of this result to variations in both the assumptions of the model and the details of the dynamic process.

Our analysis depends on the overlapping-generations assumption. The key simplifying element is that the consequences of current behavior do not extend beyond next period. An alternative approach, that would extend the results to games such as a recurrent Prisoners' Dilemma, would be to give players a discount factor. Appropriate restrictions on the parameters would ensure that, like the overlapping generations model, the effective consequences of current behavior do not extend beyond next period. So long as a player's next period flags depend only on his current action and his current opponent's flags this raises no major difficulties. Eliminating this restriction on the strategy space imposes some cost in increased complexity of the analysis. We could then rely on conditions on the range of parameters to ensure that next period cooperation is "important" and later cooperation is not.

A second essential element of the analysis is the assumption that strategies are observable. To see the importance of this assumption, consider an alternative scenario where strategies are unobserved. In order to have a well-posed optimization problem, players must have beliefs about the distribution of opponents' actions conditional on their own flag. They may form these beliefs based on observations of the action, flag and

information system used by past players. Because there is flag noise, typically there will be enough red and green flags in the sample to draw reliable inferences about the conditional distribution of actions. Occasionally, however, there may be only a few flags of a particular color, and this will make reliable inferences about the strategies being employed difficult. Moreover, the frequency with which the flag pool will have few or no flags of a particular color is much greater by assumption than the frequency of mutations. This means that the dynamic process will be driven by the inferential procedure, rather than the mutations, so our analysis does not apply.

Although this is an important limitation on our analysis, we think that it is less important than it seems for two reasons. First, it is possible to construct a mechanism by which strategies are directly observable; for example, players can write their strategy down and have an agent play on their behalf. At the end of the period, the paper is revealed. Indeed, we can allow the information system to depend on whether a written strategy of this type is played, or whether the player plays on his own behalf. A variation on the green team strategy which assigns a green flag only if the observable green team strategy is employed, together with its red counterpart will then be the unique long-run steady state. In other words, if some strategies are observable, and others not, the evolutionary process will itself choose the observable strategy, especially if punishment is given for failing to use an observable strategy.

Second, it is possible to consider sampling procedures that include and discard observations based on the color of the flag. For example, a rule can be employed that if for a particular information system there are fewer than K red flags, then observations with red flags are never discarded from the sample unless they are replaced with another

red flag observation. This means that inferences about the distribution of actions conditional on flags are always dominated by sample information rather than priors. Moreover, the employment of these sampling procedures makes sense, as the goal of players is to draw inferences based on data rather than priors.

References

- G. Ellison, "Basins of Attraction and Long-Run Equilibria," MIT, 1995.
- D. Fudenberg and D. K. Levine, "Learning in Games," Cambridge: MIT Press, Cambridge, 1998.
- M. Kandori, Social Norms and Community Enforcement, *Review of Economic Studies*, 59 (1992), 61-80.
- M. Kandori, G. Mailath and R. Rob, Learning, Mutation and Long Run Equilibria in Games, *Econometrica*, 61 (1993), 27-56.
- S. Morris, R. Rob and H. Shin, p-dominance and Belief Potential, *Econometrica*, 63 (1993): 145-158.
- P. Young, The Evolution of Conventions, *Econometrica*, 61 (1993), 57-83.

Footnotes

¹ The assumption that all players have access to a common pool of observations of past play is made only for tractability. If the agents had access to many private observations or so long as the public pool was sufficiently larger than the private pool, beliefs of the agents would be similar.

² Each information system has a corresponding always-selfish strategy; which one is used is irrelevant.