Extended Abstract: Cooperation without Association

Christopher Ahern Robin Clark Steve Kimbrough

January 31, 2014

Contents

1 Introduction

The work reported here is part of a research stream motivated fundamentally by two problems. The first is that of understanding and explaining cooperation.

How is it that cooperation can arise and be sustained among self-interested agents? Evidently it can, but the primary models/explanations, based on either kinship or association over time, are generally acceded to be insufficient.[1] In this paper, we present and develop a novel model, in which a society of agents will sustain cooperation in repeated play, even though the counter parties change randomly with each round of play. That is, our agents establish and maintain cooperation without (very much) association.

The second problem fundamentally motivating our research is the matter of understanding and modeling strategic behavior as it occurs embedded in social systems. Standard game-theoretic modeling abstracts away social context, with payoffs presented in utilities without regard to their sources. Society and its norms are left undistinguished from individual rewards; the two are conflated. How might we do modeling that teases apart these two sources of value and influence? Our aforementioned model affords this distinction and offers useful insights in that regard.

Models of cooperation predict that agents should cooperate either with their kin (Dugatkin, 2006; Harman, 2010) or with an agent with which they have a long-running association (Axelrod, 1984, 1996; Axelrod and Hamilton, 1981). The former source, although widely accepted, is presently contested in favor of group selection (Wilson, 2012; Sober and Wilson, 1998). The latter source is accepted as sufficient, when it applies, e.g.,

> The key to the evolution of cooperation, collective action, and social structure is correlation. Correlation of interactions allows the evolution of cooperative social structure that would otherwise be impossible. Social institutions enable, and to a large part consist of, correlated interactions. (Skyrms, 2004, pp. xii–xiii)

The question of cooperation, however, remains very much open.[2] While association or correlation *may* produce cooperation, the question remains when and where it actually does and has. There is the further question of whether other principles, other than kin selection, association, and group selection can be relevant. Our results in this extended abstract address specifically the latter issue and suggests that social factors may be relevant to understanding cooperation.

We focus mainly on presenting our model. In the concluding section we discuss its significance *vis à vis* sociality and directions for future research.

[1] Documenting this claim is not within the scope of this extended abstract, but see, e.g., (Tomasello, 2009, page 51) for one of many sources that identify understanding altruism and cooperation as an unsolved problem of paramount importance. See (Kimbrough, 2011, 2012) for an overview of the subject.

[2] And the subject of very much research, which we cannot review in this short space. See generally work on evolutionary games and especially coevolutionary games, the latter reviewed in (Perc and Szolnoki, 2010).

2 Defining the Game

On the standard view from game theory, strangers should be met with suspicion and, certainly, not with cooperation. Cooperation in a one-shot Prisoner's Dilemma should quickly be extinguished. Cooperation between strangers remains a mystery. In this paper, we propose to look directly at cooperation between strangers—one-shot games where the players cannot anticipate repeated play—and investigate whether some level of cooperation can be maintained.

Games are normally construed in terms of the actions of the players and the payoffs associated with the outcomes. A reasonable question is whether this information exhausts our strategic behavior. An agent deciding a policy of play might use information beyond utilities to inform her behavior. We propose to look at the "attractiveness" of the outcomes as another source of information that agents use in establishing their behavior (Camerer, 2003).

We use a standard one-shot Prisoner's Dilemma (PD) game with a form of sampling dynamics; unlike the usual replicator dynamics, our model is not strategy-centric but individual-centric. Moreover, unlike the replicator dynamics, our agents do not replicate; instead they may change the moves they make (e.g., cooperate, defect) in response to their experiences. That is, our agents change their behavior in response to the social world; they are not bare strategies that are replaced by other strategies.

The usual result in a PD game is that it is rational to play the dominant strategy of defect. Suppose that players have an additional piece of information beyond their utilities for their direct payoffs from the game; in particular, players are given information about whether their realized individual payoff is below mean or whether they received a payoff that is at the mean or above. That is, players are given a piece of social information about the payoffs of some subset of players. This opens the way for a strategy based on social information, for example Win, Stay; Lose, Shift where players keep the same stage-game strategy if their own payoff is at the mean or better and shift strategies if their own payoff is below.

3 Core Model

The evolution of cooperation and reciprocity has largely been studied from the perspective of repeated Prisoner's Dilemma games where the two players have an association; that is, the two players are selected and play between them is continued with some fixed probability. (The situation then is said to be *Iterated PD*.) This makes some sense, since cooperation is a dominated strategy in a one-shot PD. In the latter case, even in an environment where many players initially adopt the cooperate strategy, we would expect cooperation to be quickly extinguished; defection always pays better than cooperation when the players have no reason to suppose that they will ever play each other again.

From a broader social perspective, of course, the behavior exhibited in the one-shot PD makes little sense. In a population where there is widespread cooperation, the defectors make a quick profit. But soon enough, the cooperators learn their lesson and become defectors; cooperation is extinguished and the returns from defection plummet, an example of the tragedy of the commons (Hardin, 1968) where the defectors exhaust their best resource, cooperation. If only the population could independently maintain a sufficiently high level of cooperation, then both cooperators and defectors could maintain an expectation of reasonable payoff. This observation suggests a different approach to the one-shot PD. Suppose that we have repeated random play of a PD in a constant pool. That is, the play is repeated but not iterated, since the counter parties are randomly drawn in each instance of play. The agents learn only whether the strategy they played resulted in a reward at the mean or better:

> If the agent's payoff after the current play is less than the population mean, then the agent changes its stage-game strategy.

In other words, the agents play Win, Stay; Lose, Shift given the population mean as their metric. That is, their strategy in the super game is Win, Stay; Lose, Shift, which drives the particular stage-game strategies they use as play unfolds. We will show, first, that this update rule maintains a 50% cooperation rate in the population and, second, that individual agents do not need information about the entire population. If they only have information about 1 agent, we replicate the standard result that cooperation extinguishes quickly. As agents are given information about the mean payoff of more individuals, cooperation increases and rapidly approaches the 50:50 limit.

In order to illustrate the intuition, we give an example of a PD in Figure 1; the idea is that one agent can give another agent a benefit, b, with a value of 10 units, but doing so exacts a cost, c, the agent 5 units. If two agents cooperate, they each get a payoff of $b - c$, or 5 units. If one agent defects and the other agent cooperates, the defector gets pure benefit, 10 units, and the cooperator suffers pure cost, -5 units. Finally, two defectors exchange nothing, for a payoff of 0.

	Cooperate	Defect
Cooperate	$5, 5$	$-5, 10$
Defect	$10, -5$	$0, 0$

Figure 1: A Prisoner's Dilemma Game

	Cooperate	Defect
Cooperate	Stay, Stay	Shift, Stay
Defect	Stay, Shift	Shift, Shift

Figure 2: Strategy Change Bimatrix

Suppose we have a population of agents, split evenly between cooperators and defectors. If agents are randomly paired, the population average is 2.5. A pair of cooperators and a defector playing a cooperator will do better than the population mean with payoffs of 5 and 10, respectively; these agents will keep their strategies. A pair of defectors and a cooperator playing a defector will do worse than the population mean with payoffs of 0 and -5, respectively; these agents will shift their strategies. We've summarized this in Figure 2.

3.1 Knowledge of the Population

For the sake of generalizing the above comments into a formal model, suppose that p is the proportion of cooperators in the population, so that $(1-p)$ is the proportion of defectors. Assume $0 < p < 1$ and random, independent pairing of agents in the population, each agent involved in exactly one pairing. Given an initial proportion of cooperators in the population, we can determine the proportion at the next point in time—next stage in the stage game—under these circumstances.

Using notation established by Axelrod (1984) and followed in the literature, we designate the reward for joint cooperation as R ($= (b-c)$ in our development below of the PD), the penalty for mutual defection as P ($= 0$ in our development), the temptation to defect as T ($= b$), and the sucker's payoff as S $(= -c)$. In a Prisoner's Dilemma game it is required that $T > R > P > S$ and that $2R > T + S$. Thus, by inspection it is evident that an agent obtaining a reward of T or R gets an above mean reward; similarly, P and S are below the mean.

The expected utility of cooperation and defection are the following:

$$\begin{aligned} EU(C) &= p(b-c) + (1-p)(-c) = pb - c \\ EU(D) &= p(b) + (1-p)(0) = pb \end{aligned} \tag{1}$$

The average payoff in the population as a whole is then:

$$EU = pEU(C) + (1-p)EU(D) = p(b-c) \tag{2}$$

The different possible interactions and decision according to the rule can be seen in Table 1. *Self* gives the strategy of the focal individual, *Other* gives the strategy of the opponent. The payoff to the focal individual is listed. Given the simple rule of staying when doing better than average, or moving when doing worse, WIN, STAY; LOSE, SHIFT, *PlayC* shows whether the individual will cooperate at the next point in time after a particular interaction. The probability of the interaction is given by Pr.

Self	Other	Self-Payoff	Other-Payoff	PlayC	Pr
C	C	$b-c$	$b-c$	1	p^2
C	D	$-c$	b	0	
D	C	b	$-c$	0	
D	D	0	0	1	$(1-p)^2$

Table 1: Possible interactions and outcomes

Cooperators who met with other cooperators will have an above mean return and so continue to cooperate. Defectors who meet other defectors will have a below mean return and so switch to cooperation. All others will defect. Thus, the proportion of cooperators in the population at the next point in time can be given as the following recursion.

$$\begin{aligned} p' &= p^2 + (1-p)(1-p) \\ &= 1 - 2p + 2p^2 \end{aligned} \tag{3}$$

It is clear that $p' \neq p$, unless $p = 1$ or $p = 0$ (a possibility we exclude by assumption) or $p = \frac{1}{2}$, in which case $p' = p$. This can be seen by considering the change in the proportion of cooperators from one point in time to the next, $\Delta p = p' - p$. There is no change in the proportion of cooperators when $p = \frac{1}{2}$.

$$\Delta p = 1 - 3p + 2p^2 \tag{4}$$

Note further that

$$\frac{dp'}{dp} = -2 + 4p \tag{5}$$

which equals 0 when (and only when) $p = \frac{1}{2}$, and thus is stable.

To visualize, consider the plot of p versus p' in Figure 3. In what follows, open circles indicate unstable equilibria and closed circles indicate stable equilibria. For example, note that the homogeneous populations where all agents cooperate or all agents defect are not stable. In contrast, the internal equilibrium where the half the population cooperates at any point in time is a stable equilibrium. What the graph tells us is that if p is less than $\frac{1}{2}$, the next value—p'—will be $> \frac{1}{2}$ and if $p > \frac{1}{2}$ the next value will be asymptotically closer to, but greater than, $\frac{1}{2}$. In short, $p = \frac{1}{2}$ is the single stable equilibrium point, which in general will be approached from above.

This version of the model assumes that every agent at every round of play receives information about its position in the entire population. In the next sections, we consider cases in which this requirement is relaxed.

3.2 Knowledge of a Single Opponent

Let the proportion of cooperators in the population be p and the proportion of defectors be $(1-p)$. We can calculate the likelihood of certain interactions and the impact of those interactions on the composition of the population over time. The different interactions are shown in Table 2.

Self	Other	Self-Payoff	Other-Payoff	PlayC	Pr
C	C	$b-c$	$b-c$	1	p^2
C	D	$-c$	b	0	
D	C	b	$-c$	0	
D	D	0	0	0	

Table 2: Possible interactions and outcomes

If a cooperator meets another cooperator, or a defector meets another defector, they will both receive the same payoff. Thus, these individuals will do just as well as what they take to be the "average". This means that

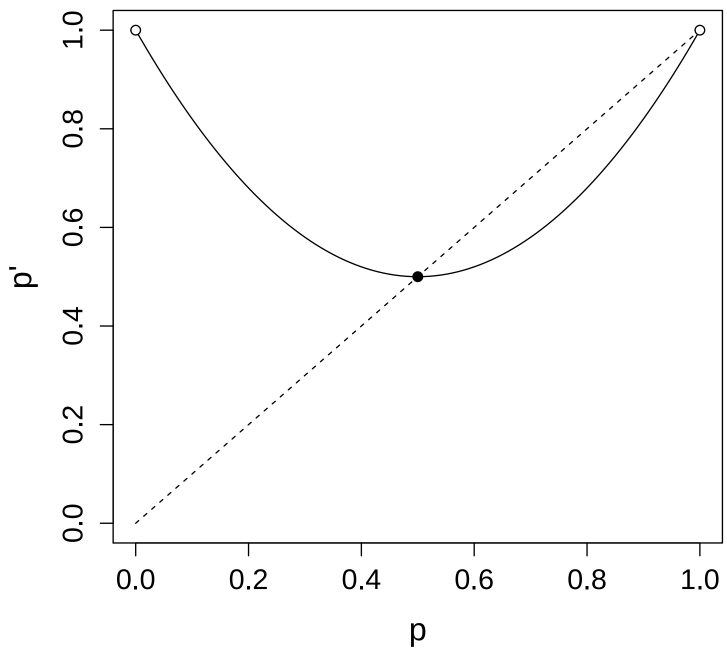

Figure 3: Change of proportion of cooperators with knowledge of population average over time. p' as a function of p: $p' = 1 - 2p + 2p^2$.

cooperators who meet cooperators will continue to cooperate and defectors who meet defectors will continue to defect. In contrast, a cooperator who meets a defector does worse than what they take to be the "average" and will switch to defecting. From the opposite perspective, a defector who meets a cooperator will continue to defect because it is better than average.

Give an initial proportion of cooperators in the population, we can determine the proportion at the next point in time under these circumstances. Only cooperators who met with other cooperators will continue to cooperate. All others will defect. Thus, the proportion of cooperators in the population at the next point in time can be given as the following recursion.

$$p' = p^2 \tag{6}$$

It is clear that unless $p = 1$ the proportion of cooperators at the next point in time will be less than the current proportion. This follows from the fact that p is a proportion. Over time p will grow smaller and smaller and cooperation will be driven out of the population.

The difference equation can be given as follows.

$$\Delta p = p^2 - p \tag{7}$$

The equilibria in the population are the compositions where the proportion of cooperators does not change from one point in time to the next. This is true just when $p = 0$ or $p = 1$. We can determine the stability of these equilibria by evaluating the derivative of the recursion.

$$\frac{dp'}{dp} = 2p \tag{8}$$

Note that only the equilibrium where $p = 0$ is stable. This follows from the fact that the derivative is greater than -1 and less than 1. The equilibrium where $p = 1$ is thus unstable. This simply acts as a check on the intuition we noted above. Namely, if p is a proportion, and we continue to square it, then it will vanish in the limit.

We can recover this intuition by visual inspection of the recursion in Figure 4. For any value of p, p' will necessarily be smaller. This process continues and cooperation disappears from the population.

3.3 Knowledge of Opponent and One More

Thus far we have staked out the extreme cases where knowledge is rather limited, or rather robust. However, we might consider those intermediate

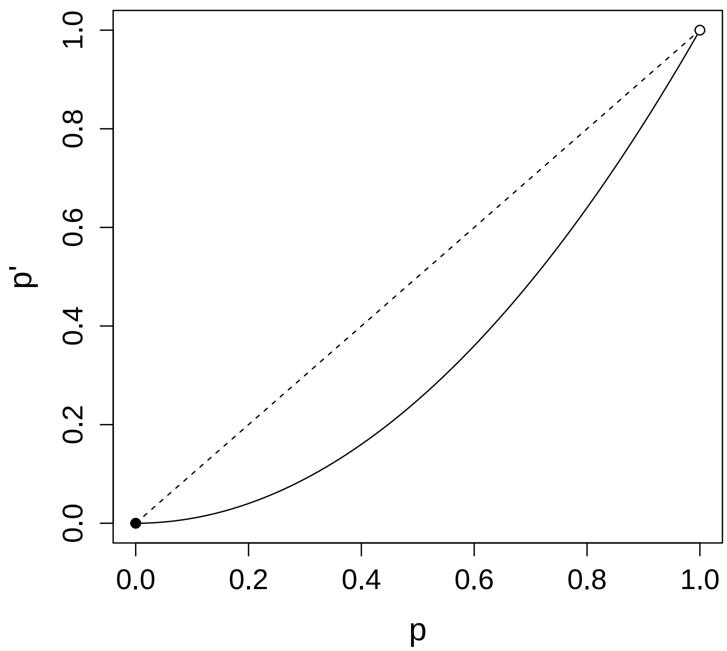

Figure 4: Change of proportion of cooperators with knowledge of single opponent over time. p' as a function of p: $p' = p^2$.

Self	Other	Self-Payoff	Other-Payoff	Obs	Obs-Opp	Obs-Payoff	Avg-Payoff	PlayC	Pr
C	C	$b-c$	$b-c$	C	C	$b-c$	$b-c$	1	p^4
C	C	$b-c$	$b-c$	C	D	$-c$	$\frac{1}{2}b-c$	1	$p^3(1-p)$
C	C	$b-c$	$b-c$	D	C	b	$b-\frac{1}{2}c$	0	
C	C	$b-c$	$b-c$	D	D	0	$\frac{1}{2}(b-c)$	1	$p^2(1-p)^2$
D	D	0	0	C	C	$b-c$	$\frac{1}{2}(b-c)$	1	$p^2(1-p)^2$
D	D	0	0	C	D	$-c$	$-\frac{1}{2}c$	0	
D	D	0	0	D	C	b	$\frac{1}{2}b$	1	$p(1-p)^3$
D	D	0	0	D	D	0	0	0	

Table 3: Knowledge of Opponent and One More

cases to see if there is some crucial amount of knowledge that is sufficient to sustain cooperation in a population. We do so by considering first the addition of a single observation to the case of limited knowledge. We show that knowing your opponent's payoff along with another individual's is not sufficient to maintain cooperation in the population. However, we show that knowing your opponent's payoff along with the payoff of two other individuals is sufficient to maintain cooperation in the population.

The addition of another individual's payoff to the decision making process can be calculated from the details in Table 3.[3] In this case, *Obs* indicates the action taken by the additional agent observed by the focal individual. *Obs-Opp* indicates the action taken by the observed agent's opponent. *Obs-Payoff* indicates the payoff to the observed agent given these actions. The average payoff of the focal individual's opponent and the observed agent are listed in *Avg-Payoff*.

We can give the recursion for this as the following:

$$\begin{aligned} p' &= p^4 + p^3(1-p) + 2p^2(1-p)^2 + p(1-p)^3 \\ &= 2p^3 - 2p^2 + p \end{aligned} \tag{9}$$

The difference equation is then the following, with equilibria where $p = 0$ and $p = 1$.

$$\Delta p = 2p^3 - 2p^2 \tag{10}$$

Examining the derivative shows that only the equilibrium where $p = 0$ meets the requirements of stability.

[3]Note that a cooperator who meets a defector will never continue cooperating, no matter what she observes. Similarly, a defector who meets a cooperator will never observe anything that makes her switch to cooperating. We leave these possibilities for the focal individual out of the table.

$$\frac{dp'}{dp} = 6p^2 - 4p + 1 \tag{11}$$

Given that p is a proportion, then $2p^3 < 2p^2$, thus p' will always be less than p. Thus, as in the case of information about a single opponent, adding knowledge about another individual does not stop the downward spiral to defection. Again, we can see these results in a plot of the recursion in These results can be seen in Figure 5.

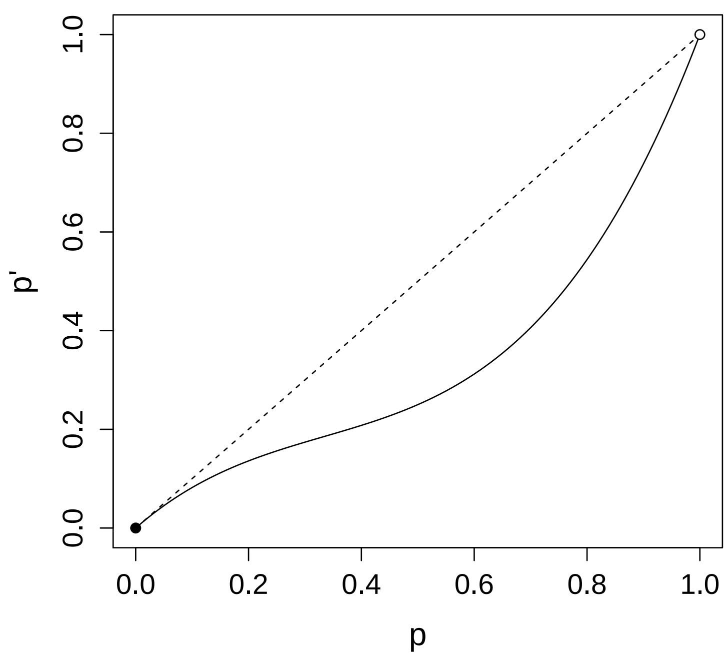

Figure 5: Change of proportion of cooperators with knowledge of single opponent and one other over time. p' as a function of p: $p' = 2p^3 - 2p^2 + p$.

3.4 Knowledge of Opponent and Two More

The addition of two individuals' payoffs to the decision making process can be calculated from the details in Table 4. *Obs1* indicates the first and *Obs2*

Self	Other	Self-Payoff	Other-Payoff	Obs1	Obs1-Opp	Obs2	Obs2-Opp	Avg-Payoff	PlayC	Pr
C	C	$b-c$	$b-c$	C	C	C	C	$b-c$	1	p^6
C	C	$b-c$	$b-c$	C	C	C	D	$\frac{2}{3}b-c$	1	$p^5(1-p)$
C	C	$b-c$	$b-c$	C	C	D	C	$b-\frac{2}{3}c$	0	
C	C	$b-c$	$b-c$	C	C	D	D	$\frac{2}{3}(b-c)$	1	$p^4(1-p)^2$
C	C	$b-c$	$b-c$	C	D	C	C	$\frac{2}{3}b-c$	0	
C	C	$b-c$	$b-c$	C	D	C	D	$\frac{1}{3}b-c$	1	$p^4(1-p)^2$
C	C	$b-c$	$b-c$	C	D	D	C	$\frac{2}{3}(b-c)$	1	$p^4(1-p)^2$
C	C	$b-c$	$b-c$	C	D	D	D	$\frac{1}{3}(b-2c)$	1	$p^3(1-p)^3$
C	C	$b-c$	$b-c$	D	C	C	C	$\frac{2}{3}(b-c)$	1	$p^5(1-p)$
C	C	$b-c$	$b-c$	D	C	C	D	$\frac{2}{3}(b-c)$	1	$p^5(1-p)$
C	C	$b-c$	$b-c$	D	C	D	C	$b-\frac{1}{3}c$	0	
C	C	$b-c$	$b-c$	D	C	D	D	$\frac{1}{3}(2b-c)$	?	$p^3(1-p)^3$
C	C	$b-c$	$b-c$	D	D	C	C	$\frac{2}{3}(b-c)$	1	$p^4(1-p)^2$
C	C	$b-c$	$b-c$	D	D	C	D	$\frac{1}{3}(b-2c)$	1	$p^3(1-p)^3$
C	C	$b-c$	$b-c$	D	D	D	C	$\frac{1}{3}(2b-c)$	?	$p^3(1-p)^3$
C	C	$b-c$	$b-c$	D	D	D	D	$\frac{1}{3}(b-c)$	1	$p^2(1-p)^4$
D	D	0	0	C	C	C	C	$\frac{2}{3}(b-c)$	1	$p^4(1-p)^2$
D	D	0	0	C	C	D	C	$\frac{1}{3}(2b-c)$	1	$p^3(1-p)^3$
D	D	0	0	C	C	C	D	$\frac{1}{3}(b-2c)$	?	$p^3(1-p)^3$
D	D	0	0	C	C	D	D	$\frac{1}{3}(b-c)$	1	$p^2(1-p)^4$
D	D	0	0	C	D	C	C	$\frac{1}{3}(b-2c)$	?	$p^3(1-p)^3$
D	D	0	0	C	D	C	D	$-\frac{2}{3}c$	0	
D	D	0	0	C	D	D	C	$\frac{1}{3}(b-c)$	1	$p^2(1-p)^4$
D	D	0	0	C	D	D	D	$-\frac{1}{3}c$	0	
D	D	0	0	D	C	C	C	$\frac{1}{3}(2b-c)$	1	$p^3(1-p)^3$
D	D	0	0	D	C	C	D	$\frac{1}{3}(b-c)$	1	$p^2(1-p)^4$
D	D	0	0	D	C	D	C	$\frac{2}{3}b$	1	$p^2(1-p)^4$
D	D	0	0	D	C	D	D	$\frac{1}{3}b$	1	$p(1-p)^5$
D	D	0	0	D	D	C	C	$\frac{1}{3}(b-c)$	1	$p^2(1-p)^4$
D	D	0	0	D	D	C	D	$-\frac{1}{3}c$	0	
D	D	0	0	D	D	D	C	$\frac{1}{3}b$	1	$p(1-p)^5$
D	D	0	0	D	D	D	D	0	0	

Table 4: Knowledge of Opponent and Two More

the second agent observed. Note that those rows with ? in *PlayC* are 1 if $\frac{b}{c} > 2$ and 0 otherwise. We consider these two cases in turn.

3.4.1 $\frac{b}{c} \leq 2$

The recursion, given the information in Table 4, is the following:

$$\begin{aligned} p' &= p^6 + 3p^5(1-p) + 5p^4(1-p)^2 + 4p^3(1-p)^3 + 6p^2(1-p)^4 + 2p(1-p)^5 \\ &= 3p^6 - 9p^5 + 9p^4 - 4p^2 + 2p \end{aligned} \tag{12}$$

Thus the difference equation is the following, with real roots at $p = 0$, $p = 1$, $p = .29$.

$$\Delta p = 3p^6 - 9p^5 + 9p^4 - 4p^2 + p \tag{13}$$

Only the last of these is a stable equilibrium, as can be seen from the derivative.

$$\frac{dp'}{dp} = 18p^5 - 45p^4 + 36p^3 - 8p + 2 \tag{14}$$

We can see this visually in Figure 6 where $p = .29$ is the only stable equilibrium.

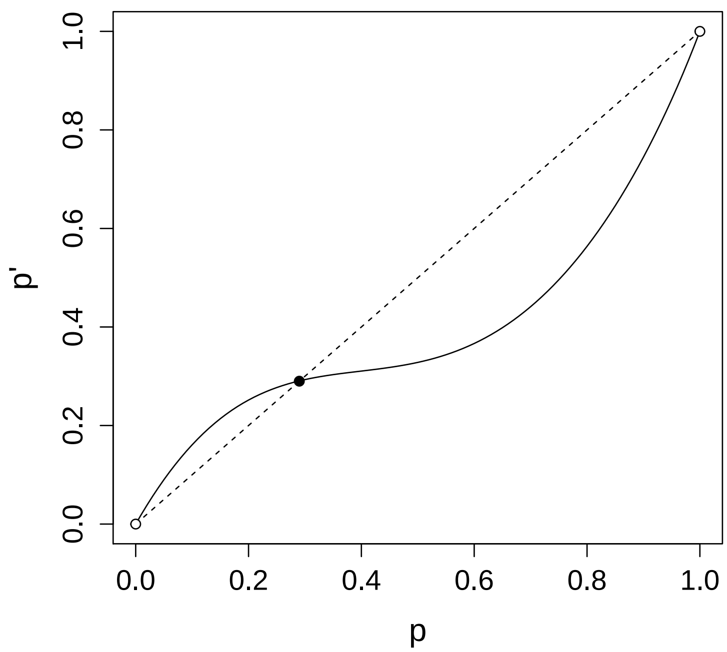

Figure 6: Change of proportion of cooperators with knowledge of single opponent and two others over time, when $\frac{b}{c} \leq 2$. p' as a function of p: $p' = 3p^6 - 9p^5 + 9p^4 - 4p^2 + 2p$

3.4.2 $\frac{b}{c} > 2$

As we noted above, if b is more than twice as large as c then more interactions will yield cooperation. The recursion, if this condition is met, is increased by $4p^3(1-p)^3$.

$$\begin{aligned} p' &= p^6 + 3p^5(1-p) + 5p^4(1-p)^2 + 8p^3(1-p)^3 + 6p^2(1-p)^4 + 2p(1-p)^5 \\ &= -p^6 + 3p^5 - 3p^4 + 4p^3 - 4p^2 + 2p \end{aligned} \tag{15}$$

Thus the difference equation is the following, with real roots at $p = 0$, $p = 1$, $p = .35$

$$\Delta p = -p^6 + 3p^5 - 3p^4 + 4p^3 - 4p^2 + p \tag{16}$$

Only the last of these is a stable equilibrium, as can be seen from the derivative.

$$\frac{dp'}{dp} = -6p^5 + 15p^4 - 12p^3 + 12p^2 - 8p + 2 \tag{17}$$

Again, we can see this visually in Figure 6 where $p = .35$ is the only stable equilibrium. The difference between the two conditions expressed in Figures 6 and 7 is visually subtle, but directly reflects the impact of the cost to benefit ratio. However, as the number of individuals observed increases towards the entire number of the population, the impact of this ratio vanishes.

4 Discussion

In our model, agents receive *social* information—how well they did compared to other members of the society—and the agents follow a social strategy, WIN, STAY; LOSE, SHIFT, with respect to their socially-conditioned rewards. Remarkably, this leads to a substantial level of cooperation in repeated Prisoner's Dilemma. Thus, we have produced a model of cooperation without kin selection, group selection, or association. Of course, presence together in a society might be counted as association, but if so, it is highly attenuated compared to the usual story. Our agents arrive at cooperation without iterated play, that is, without repeated play with fixed counter parties. The "shadow of society" (Kimbrough, 2012) rather than the "shadow of the future" (Axelrod, 1984) indeed looms large. Moreover, as we have

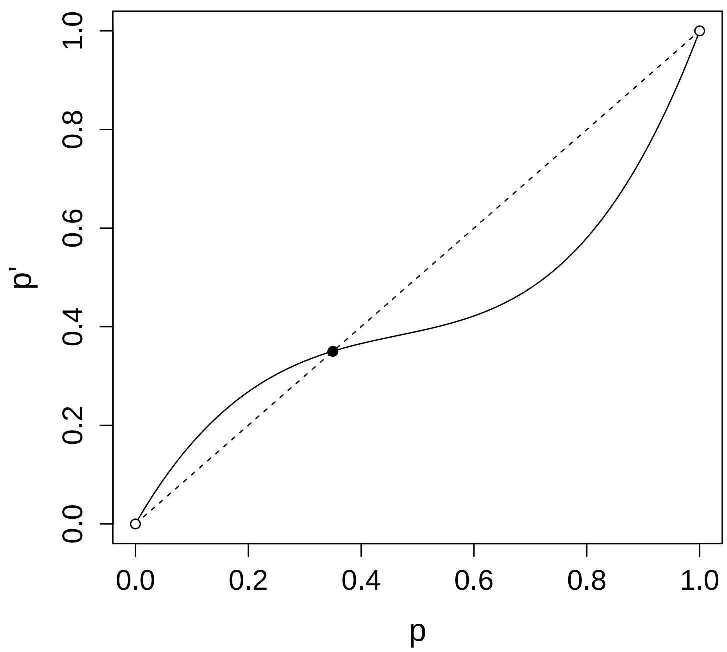

Figure 7: Change of proportion of cooperators with knowledge of single opponent and two others over time, when $\frac{b}{c} > 2$. p' as a function of p: $p' = -p^6 + 3p^5 - 3p^4 + 4p^3 - 4p^2 + 2p$.

shown, the epistemic requirements are modest; social information from a rather small subset of the society suffices to establish much cooperation. Further, the results are robust even in the presence of committed defectors. See the supplementary material, in appendix §A.

We think very much can be achieved by extending this approach. Social factors may be elaborated and investigated. For example, Win, Stay; Lose, Shift might be seen as a social norm, whether descriptive or prescriptive (Bicchieri, 2006) and so investigated, along with other contenders. Reputation (Mailath and Samuelson, 2006) and punishment (Boyd et al., 2003; Boyd and Richerson, 1992) are rife for additions to the model. Our thought is that with socially-conditioned behavior, punishment need not be altruistic, but prudent. This conjecture, of course, awaits work beyond the scope of this abstract. We note, further, that very much learning has to be seen as social (Bandura, 1976; Hirsch, 1976; Richerson and Boyd, 2005; Richerson et al., 2003) as must much of strategic interaction, particularly with regard to linguistic behavior (Clark, 2012).

In terms of specific game representations, a natural way of abstracting and generalizing our approach is to model these social factors as constituting something resembling a dynamic Bayesian game. In a simple version, *ex ante* each player is subject to a probability distribution from "nature," representing the player's propensities to pick various strategies. Nature moves and each of the players independently picks a strategy with which to play the series of stage games. In our core model, players deterministically pick Win, Stay; Lose, Shift. With the strategies selected, rounds of play ensue as described above and we observe the results. Notice that unlike as in standard Bayesian games, the types of our players are identified by their strategies, not their stage game payoffs. We submit that the strategies can be identified as, or at least linked with, social norms. Our basic finding, then, is that the social norm of Win, Stay; Lose, Shift results in a surprising degree of cooperation and is robust to defections by players choosing Always Defect. The larger point to hand here is that this conception answers well to our second motivation—representing and investigation games embedded in social contexts—and is naturally extended to models with larger consideration sets of strategies, with learning (agents switching strategies in response to experience), with gossip and punishment, and so on.

The ramifications of this simple model of sociality and strategy are indeed manifold.

References

Axelrod, R. (1984). *The Evolution of Cooperation.* Basic Books, Inc., New York, NY.

Axelrod, R. (1996). An evolutionary approach to norms. *American Political Science Review*, 80(4):1095–1111.

Axelrod, R. and Hamilton, W. (1981). The evolution of cooperation. *Science*, 211:1390–1396.

Bandura, A. (1976). *Social Learning Theory.* Prentice Hall/Pearson, New York, NY.

Bicchieri, C. (2006). *The Grammar of Society: The Nature and Dynamics of Social Norms.* Cambridge University Press, Cambridge, UK.

Boyd, R., Gintis, H., Bowles, S., and Richerson, P. J. (2003). The evolution of altruistic punishment. *PNAS*, 60(6):3531–3535.

Boyd, R. and Richerson, P. J. (1992). Punishment allows the evolution of cooperation (or anything else) in sizable groups. *Ethology and Sociobiology*, 13:171–195.

Camerer, C. F. (2003). *Behavioral Game Theory: Experiments in Strategic Interaction.* Russell Sage Foundation and Princeton University Press, New York, NY and Princeton, NJ.

Clark, R. (2012). *Meaningful Games.* The MIT Press, Cambridge, MA.

Dugatkin, L. A. (2006). *The Altruism Equation: Seven Scientists Search for the Origins of Goodness.* Princeton University Press, Princeton, NJ.

Hardin, G. (1968). The tragedy of the commons. *Science*, 162(3859):1243–1248.

Harman, O. (2010). *The Price of Altruism: George Price and the Search for the Origins of Kindness.* W.W. Norton, New York, NY.

Hirsch, F. (1976). *Social Limits to Growth.* Harvard University Press, Cambridge, MA.

Kimbrough, S. O. (2011). On the production and ramification of cooperation: The cooperation afforder with framing hypothesis. *Philosophy of the Social Sciences*, 41(1):111–136.

Kimbrough, S. O. (2012). *Agents, Games, and Evolution: Strategies at Work and Play.* CRC Press, Boca Raton, FL.

Mailath, G. J. and Samuelson, L. (2006). *Repeated Games and Reputations: Long-Run Relationships.* Oxford University Press, New York, NY.

Perc, M. and Szolnoki, A. (2010). Coevolutionary games—a mini review. *BioSystems*, 99:109–125.

Richerson, P. J. and Boyd, R. (2005). *Not by Genes Alone: How Culture Transformed Human Evolution.* The University of Chicago Press, Chicago, IL.

Richerson, P. J., Boyd, R. T., and Henrich, J. (2003). Cultural evolution of human cooperation. In Hammerstein, P., editor, *Genetic and Cultural Evolution of Cooperation*, pages 357–388. The MIT Press, Cambridge, MA.

Skyrms, B. (2004). *The Stag Hunt and the Evolution of Social Structure.* Cambridge University Press, Cambridge, UK.

Sober, E. and Wilson, D. S. (1998). *Unto Others: The Evolution and Psychology of Unselfish Behavior.* Harvard University Press, Cambridge, MA.

Tomasello, M. (2009). *Why We Cooperate.* The MIT Press, Cambridge, MA.

Wilson, E. O. (2012). *The Social Conquest of Earth.* Liveright Publishing Corporation, New York, NY.

A Supplementary Material: Social Learning with Committed Defectors

As a slight variation on our model, we consider the case where some agents are committed to defection while all others observe a WIN-STAY, LOSE-SHIFT rule according to the average payoff of the population. So now we have two super game strategies in play.

Let q be the probability that an agent will be a committed defector and $(1-q)$ the probability that she will adopt the contingent rule of changing her strategy, i.e., WIN-STAY, LOSE-SHIFT. Further, let p be the proportion of rule-following agents cooperating and $(1-p)$ be the proportion of rule-following agents defecting in a given round of play. We can calculate the expected utility of the different stage-game strategies in the population.

$$
\begin{aligned}
EU(C) &= (1-q)[p(b-c)+(1-p)(-c)]+q(-c) = (1-q)pb-c \\
EU(D) &= (1-q)[p(b)+(1-p)(0)]+q(0) = (1-q)pb
\end{aligned}
\tag{18}
$$

This, in turn, allows us to calculate the average payoff in the population as a whole:

$$
\begin{aligned}
EU &= (1-q)[pEU(C)+(1-p)EU(D)]+qEU(D) \\
&= (1-q)p(b-c)
\end{aligned}
\tag{19}
$$

If individuals only adopt a different strategy when they do worse than the population average, then we can determine which interactions will lead to cooperation or defection as shown in Table 5. We represent the committed defectors as E in the table. Here we assume that committed defectors, as the term suggests, are committed to defecting and will not change strategy regardless of any information they may receive. We also assume that agents playing according to the contingent rule cannot make a committment to defection.

Self	Other	Self-Payoff	Other-Payoff	PlayC	Pr
C	C	$b-c$	$b-c$	1	$(1-q)^2p^2$
C	D	$-c$	b	0	
C	E	$-c$	b	0	
D	C	b	$-c$	0	
D	D	0	0	1	$(1-q)^2(1-p)^2$
D	E	0	0	1	$(1-q)(1-p)q$
E	C	b	$-c$	0	
E	D	0	0	0	
E	D	0	0	0	

Table 5: Possible interactions and outcomes

When two cooperators meet the payoff to the focal individual will be $b-c$. This is greater than or equal to the average payoff for any non-homogeneous population where $0 < p, q < 1$. Thus cooperators who meet cooperators will always continue to cooperate. In contrast, a cooperator who meets a defector, either contingent or committed, will do worse than the average payoff. A defector who meets a cooperator will do much better than the average payoff. Both will begin or continue to defect. Finally, for all populations where there is a nonzero proportion of cooperators, a contingent defector meeting another defector will do worse than the population average. These contingent defectors will cooperate moving forward.

Given the proportion of cooperators in a population at one time, we can give the proportion at the next point in time:

$$\begin{aligned}(1-q)p' &= (1-q)^2p^2 + (1-q)^2(1-p)^2 + (1-q)(1-p)q \\ p' &= (1-q)2p^2 + (q-2)p + 1\end{aligned} \tag{20}$$

Note that, as we would expect, when $q = 0$ this reduces to the case discussed above. This allows us to calculate the proportion of cooperation amongst those agents following the rule. The change of the proportion of cooperation from one time period to the next is:

$$\Delta p = (1-q)2p^2 + (q-3)p + 1 \tag{21}$$

If there is a nonzero proportion of committed defectors then the rate of cooperation does not change when $p = \frac{1}{2}$. We can see that this internal fixed point is the only stable one:

$$\frac{\partial p'}{\partial p} = (1-q)4p + (q-2) \tag{22}$$

Agents who are committed to defection, in essence, squeeze out the space for cooperation, but the general result holds. All agents that abide by the rule and take winning to be doing better than the population average cooperate half of the time.