When Naughty is Nice: Cooperation and Noisy Recognition of Ingroups*

Carmen Astorne-Figari †

This version: August 8, 2019

Abstract

This paper studies a prisoners' dilemma played between two individuals who exhibit altruistic preferences towards ingroups, but ingroup recognition is noisy. Players know the probability of playing with an ingroup, but, upon meeting, observe a noisy private signal of group membership, choosing to cooperate only with those identified as ingroups. When the prisoners' dilemma is played for two periods instead of one, noisy recognition of ingroups brings strategic considerations into the game even if the odds of facing another ingroup are small. There are Perfect Bayesian Equilibria in which players cooperate in the first round even after getting an outgroup signal. Since a player can make inferences about her opponent's signal based on first period actions, outgroups mimic ingroups to induce cooperation in the second period.

Keywords: cooperation; social identity; diversity; altruism; prisoners' dilemma

JEL Classification Numbers: C72, C73, D64, D91, Z13

*The author has benefited from insightful comments from John Nachbar, David Levine and Yuri Khoroshilov. This work stems from an earlier presented working paper entitled Kin Targeted Altruism with Noise.

†Department of Economics, University of Memphis, 427 Fogelman Admin Building, Memphis, Tennessee 38152; e-mail address: cmstrnfg@memphis.edu.

1 Introduction

An individual's behavior in strategic situations can be affected by her social identity, as established in the seminal work of Akerlof and Kranton [2000]. In particular, individuals perceive relevant social groups or categories and identify with some of them. Experimental evidence shows that, in many situations, individuals tend to favor other individuals they perceive as belonging to the same categories they do, called ingroups (Goette et al. [2006], Chen and Li [2009], Klor and Shayo [2010], Kranton and Sanders [2017]). However, other studies show that an individual's definition of self may be dependent on social context or priming (Charness et al. [2007], Eckel and Grossman [2005], Chen et al. [2014], Cadsby et al. [2013]). This suggests that, in some cases, recognition of ingroup members may be noisy.

This paper studies a prisoners' dilemma played by individuals with altruistic preferences towards other ingroup members who only cooperate with those perceived as ingroups in a one-shot game. In this context, I study whether repetition can increase cooperation when players are altruistic towards ingroups and ingroup recognition is noisy.

I find that, when the prisoners' dilemma is played for two periods (after observing the signals only once), there are cases where players always cooperate in the first period. In particular, players who observe the outgroup signal cooperate in the first period for strategic reasons. First-period cooperation conceals a player's signal from her opponent, who may have observed the ingroup signal. This allows her to defect in the second period against a potentially cooperating opponent. Here, when signals are sufficiently noisy, repetition can generate an incentive to cooperate even if the probability of facing an ingroup is not very high.

1.1 Related Literature

This paper makes two contributions to the literature on the economics of identity. The first is to introduce noise in the recognition of other in-group members, which plays a critical role in my results. The second is to examine the interaction between altruistic preferences towards ingroup members and strategic considerations brought by repeated interactions among players.

As in Akerlof and Kranton [2000, 2002, 2005, 2008] and Shayo [2007], there are social categories or groups, which are subsets of the overall population. Although I don't model these groups explicitly, we can think of individuals being ingroups whenever there exists at least one social group that they both identify with (i.e. they are both in-groups to some group), and outgroups otherwise. I am agnostic about how or why individuals identify with certain categories. That is, categories and player identification with them are exogenous in my model, but the probability that two individuals identify with the same category is common knowledge.

Noisy recognition of ingroups can be motivated by the fact that individuals may disagree about the saliency of characteristics of a group or category with which they identify. This contrasts with Shayo [2007], where all characteristics of social category prototypes are equally salient across all individuals.

Finally, my results resemble Healy [2007], which studies equilibria of a sequential prisoners' dilemma framed in a labor market context, where selfish workers choose to exert high effort to avoid hurting worker reputation and obtain high wages. As in this work, the presence of altruistic individuals generates an incentive for non-altruistic players to choose cooperative actions for strategic reasons.

2 Setup

Two players, 1 and 2, meet to play a prisoner's dilemma, with payoffs shown in panel (a) of Table 1. The parameter $c > 0$ denotes the price of cooperating with one's opponent: if a player decides to cooperate with her opponent instead of defecting, she benefits the opponent in 1 dollar at a cost of c dollars that she could have kept for herself.[1]

Table 1: Prisoners' Dilemma

		Player 2	
		C	D
Player 1	C	$1, 1$	$-c, 1+c$
	D	$1+c, -c$	$0, 0$

(a) Players are outgroups

		Player 2	
		C	D
Player 1	C	$2, 2$	$1, 1$
	D	$1, 1$	$0, 0$

(b) Players are ingroups

Players can be ingroups (I) or outgroups (O). Players are ingroups with probability $p \in (0, 1)$, which is common knowledge. That is, p denotes the probability of positive assortative matching. When matching is assortative, cooperation can be sustained because the cost of cooperating may be repaid via higher probabilities of playing against a cooperating opponent (See Bergstrom [1995, 2003] for more references). We are not interested in the cases where matching is assortative, and will focus on cases where p is low. When players are ingroups, a player's payoff increases in that of her opponent.[2] Payoffs are shown on panel (b) of Table 1. Note that players always prefer to cooperate with another ingroup and defect against an outgroup.

[1] Similar to that in Andreoni and Vesterlund [2001].
[2] Similar to using Hamilton's Rule, as in Bergstrom [2003].

Upon meeting, each player observes a signal in $\{\iota, \omega\}$, where ι is the signal associated to ingroups, and ω is the signal associated to outgroups, each with different levels of noise. The distribution of the realized pair of signals conditional on the state of the world (I or O) is common knowledge.

A player observes the outgroup signal conditional on players being ingroups with probability $m < 1/2$. Similarly, a player observes the ingroup signal conditional on players being outgroups with probability $n < 1/2$. The parameters m and n capture the noise of the ingroup and the outgroup signal respectively. Thus, the probability that both players observe the right signal is $(1-m)^2$ conditional on being ingroups, and $(1-n)^2$ conditional on being outgroups.

Players meet, observe the private signal, and then play a one-shot prisoners' dilemma. Table 2 shows a player's expected payoffs conditional on observing signal $s \in \{\iota, \omega\}$.

Table 2: Expected Payoffs to Player 1 Conditional on Observing Signal s

		Player 2: C	Player 2: D
Player 1	C	$1 + \Pr[I\|s]$	$\Pr[I\|s] - c\Pr[O\|s]$
	D	$1 + c\Pr[O\|s]$	0

When a player observes the ingroup signal, $p \geq \frac{cn}{cn+(1-m)}$ guarantees that the expected gains to cooperation are higher than the cost of mistakenly favoring an outgroup, so the player will cooperate. Similarly, when a player observes the outgroup signal, $p \leq \frac{c(1-n)}{c(1-n)+m}$ guarantees that cooperating is too costly and the player will defect[3]. Thus, in the unique Nash equilibrium, players cooperate upon

[3]Note that $m < 1/2$ and $n, 1/2$ imply that $\frac{cn}{cn+(1-m)} < \frac{c(1-n)}{c(1-n)+m}$ always.

observing ι and defefct otherwise. If positive assortative matching were too likely, players would blindly cooperate with any opponent, regardless of the observed signal. On the other hand, if the probability of positive assortative matching were too low, players would prefer to defect all the time.

Before analyzing the two-period game, we explore what happens when players know that the realized signal pair is (ι, ω) or (ω, ι). Since both cases are computationally equivalent, assume the signal realization is (ω, ι). I call this game "artificial" because it might never be played, but has important implications on the two-times repeated game. Following a similar intuition to that of the one period game, defection is the unique Nash equilibrium of the artificial game if and only if $p \leq \frac{cn(1-n)}{cn(1-n)+m(1-m)}$. Otherwise, cooperation is the unique Nash equilibrium.

3 Effects of repetition

I restrict attention to the case where there is defection in the artificial game.[4] Here, observing the outgroup signal is sufficient for a player to defect. On the contrary, a player who observes the ingroup signal prefers to cooperate in the first period, but learning that her opponent observed the outgroup signal would lead her to defect in the second period.

The following proposition summarizes the main result. Intuition for the proof is provided below, with all derivations in the Appendix.

[4]When there is cooperation in the artificial game, there is always more cooperation in the twice repeated game. Observing the ingroup signal is sufficeint for a player to cooperate, and a player who observes the outgroup signal can be persuaded to cooperate in the second period if she learns her opponent observed the ingroup signal.

Proposition 1. *For $n \geq n^*$ and $p \in \left[\max\left\{p^*, \frac{cn}{cn+(1-m)}\right\}, \frac{cn(1-n)}{cn(1-n)+m(1-m)}\right]$, the unique Perfect Bayesian equilibrium of the twice-repeated prisoners' dilemma is:*

- *Regardless of whether ι or ω is observed, cooperate in the first period.*
- *If (a) ι is observed, and (b) both players cooperate in the first period, cooperate in the second period. Otherwise, defect in the second period.*
- *If a player observes defection in the first period, she believes the opponent observed the outgroup signal ω.*

Where $n^* = \frac{c(1-m)}{(1-m)+c(3-2m)} \in \left(0, \frac{1}{2}\right)$ and $p^* = \max\left\{0, \frac{(1-n)[c-(c+1)n]}{(2-m)m+(1-n)(c-(c+1)n)}\right\}$.

Suppose that player 1 (male) observes the outgroup signal and player 2 (female) observes the ingroup signal. Given his signal, player 1's expected first period payoff is always higher when he defects. However, defecting in the first period would reveal to player 2 that he observed the outgroup signal, and she would switch to defecting in the second period. Since, from player 1's perspective, there is a positive probability that player 2 observed the ingroup signal, he may have an incentive to cooperate in the first period.

When a player observes the outgroup signal, he benefits from cooperating in the first period whenever the opponent is an ingroup (regardless of the opponent's first period action), and whenever the opponent sees the ingroup signal (because the opponent cooperates in the first period). First period cooperation is costly whenever the opponent was correctly identified as an outgroup (because the opponent will not cooperate in the second period). When $p > p^*$, the expected benefit of cooperating in the first period (and waiting until the second period to defect) is greater than the expected benefit of defecting in the first period and being punished in the second period.[5] Since there is defection in the artificial game, it must

[5]When $p < p^* < \frac{cn(1-n)}{cn(1-n)+m(1-m)}$, the effect of repetition is the opposite and repetition

be the case that $p^* < \frac{cn(1-n)}{cn(1-n)+m(1-m)}$ for players to cooperate in the first period. This occurs whenever the outgroup signal is sufficiently noisy, namely $n \in \left[n^*, \frac{1}{2}\right)$, with $n^* = \frac{c(1-m)}{(1-m)+c(3-2m)} > 0$ for all values of m and c.

When the outgroup signal is more noisy i.e. as n increases, there is a higher probability that an outgroup observes the ingroup signal and players have an incentive to use first period cooperation to conceal their signal in the twice-repeated game. As the noise of the ingroup signal m increases, a less noisy outgroup signal can still enable more cooperation, as seen in Figure 1. The area between the dashed line and the solid line represents the parameter region with defection in the artificial game. The shaded area shows cases where all players cooperate in the first period.

increases defection. Players who observe the outgroup signal have no incentive to conceal their signal from their opponents because the rewards of waiting one period to defect are too low.

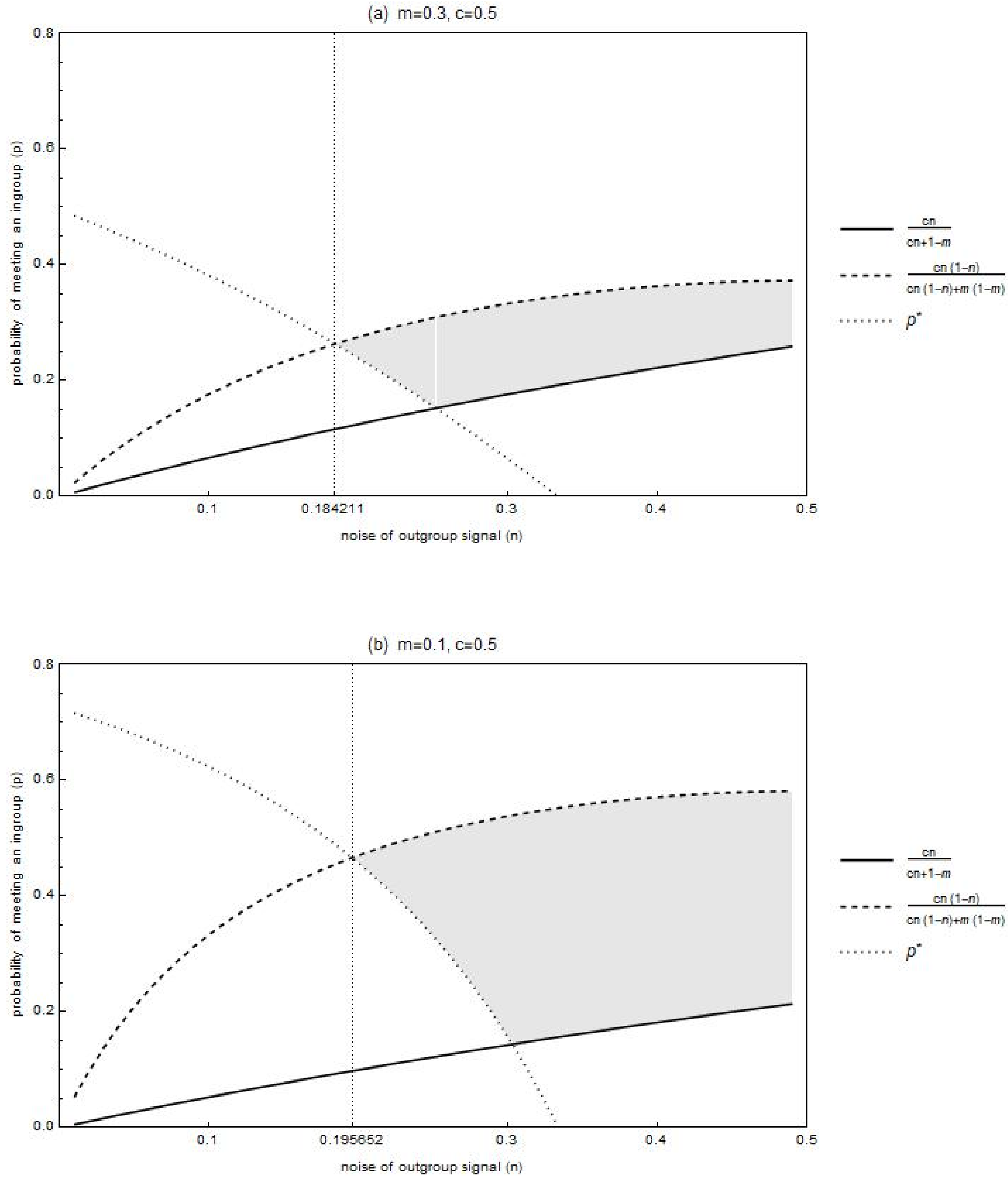

Figure 1: Parameter Region with More Cooperation

4 Discussion

These results are particularly suggestive in light of experimental evidence suggesting that individuals' identities are sensitive to the context in which interactions take place. If individuals in our society feel primed by social categories such as race, ethnicity and gender—where recognition of other ingroup or outgroup members is not likely to be noisy—we may observe less cooperative behavior. As an example, Alesina et al. [1999] find that, in U.S. cities that are more ethnically diverse, the amount spent on productive public goods such as education, sewer or roads is lower. If those individuals' identities consist primarily of ethnicity, the results are consistent with my framework. One wonders if results like Alesina et al. [1999] would be less likely to occur in a society where individual identities are based on characteristics that are harder to observe, such as education, work ethic, or family values.

5 Appendix

Proof of Proposition 1

Second Period

On equilibrium path, checking for perfection of the second period actions is immediate from the one-shot game, since no information is revealed at the end of the first period. Off the equilibrium path, a player who observes defection in the first period believes the opponent observed the outgroup signal. Given that there is defection in the artificial game, players who believe their opponent observed the outgroup signal would always choose to defect in the second period regardless of the signal they observed.

First Period

As noted above, a player who observes the ingroup signal believes that only players who observe the outgroup signal would defect in the first period. Given these beliefs, and given that there is defection in the artificial game, a player who observes the ingroup signal would only cooperate in the second period if her opponent cooperated in the first period, and defect otherwise.

Since defecting in the first period would induce defection in the second period, a player who observes the outgroup signal may wish to cooperate in the first period to conceal her signal from her opponent. Note that a player who observes the outgroup signal will always defect in the second period.

Cooperating in the first period and defecting in the second period yields

$$\underbrace{1+\underbrace{\frac{pm}{pm+(1-p)(1-n)}}_{\Pr[I|\omega]}}_{\text{1st period payoff}}+\underbrace{\underbrace{\frac{(1-p)n(1-n)+pm(1-m)}{pm+(1-p)(1-n)}}_{\Pr[(\omega,\iota)|\omega]}+c\underbrace{\frac{n(1-n)(1-p)}{pm+(1-p)(1-n)}}_{\Pr[(\omega,\iota),O|\omega]}}_{\text{2nd period payoff}} \tag{1}$$

Defecting in the first period and defecting in the second period yields

$$\underbrace{1+c\underbrace{\frac{(1-p)(1-n)}{pm+(1-p)(1-n)}}_{\Pr[O|\omega]}}_{\text{1st period payoff}}+\underbrace{0}_{\text{2nd period payoff}} \tag{2}$$

A player who observes ω prefers to cooperate in the first period and wait until the second period to defect whenever Expression (1) is greater than Expression (2).

Setting Expresion (1) greater than Expression (2) and rearranging yields the following inequality.

$$pm(2-m) \geq (1-p)(1-n)[c-n(c+1)] \tag{3}$$

Note that whenever $n > c/(c+1)$, Inequality (3) always holds. Otherwise, for $n < c/(c+1)$, Inequality (3) holds whenever

$$p \geq p^* = \max\left\{0, \frac{(1-n)\,[c-(c+1)n]}{(2-m)m+(1-n)\,(c-(c+1)n)}\right\} \tag{4}$$

As we can see in Figure 1, the upper bound for p, $\frac{cn}{cn+(1-m)}$, is increasing in n, whereas p^* is decreasing in n. Setting the upper bound for p equal to p^* yields $n = \frac{c(1-m)}{(1-m)+c(3-2m)} = n^* < \min\left\{\frac{c}{c+1}, 0.5\right\}$. Thus, for $n > n^*$, there exist $p \in \left[\max\left\{p^*, \frac{cn}{(1-m)+cn}\right\}, \frac{cn(1-n)}{cn(1-n)+m(1-m)}\right]$ that a player who observes ω always cooperates in the first period.

A player who observes the ingroup signal chooses between cooperating in both periods, or defecting in the first period and cooperating in the second period.

Cooperating in the first period and cooperating in the second period yields

$$\underbrace{1+\underbrace{\frac{(1-m)p}{p(1-m)+(1-p)n}}_{\Pr[I|\iota]}}_{\text{1st period payoff}} + \underbrace{\underbrace{\frac{(1-m)^2p}{p(1-m)+(1-p)n}}_{\Pr[(\iota,\iota)|\iota]} + \underbrace{\frac{(1-m)p}{p(1-m)+(1-p)n}}_{\Pr[I|\iota]} - c\Pr[(\iota,\omega),O|\iota]}_{\text{2nd period payoff}} \tag{5}$$

Defecting in the first period and cooperating in the second period yields

$$\underbrace{1+c\underbrace{\frac{n(1-p)}{p(1-m)+(1-p)n}}_{\Pr[O|\iota]}}_{\text{1st period payoff}} + \underbrace{\underbrace{\frac{(1-m)p}{p(1-m)+(1-p)n}}_{\Pr[I|\iota]} - c\underbrace{\frac{n(1-p)}{p(1-m)+(1-p)n}}_{\Pr[O|\iota]}}_{\text{2nd period payoff}} \tag{6}$$

A player who observes ι prefers to cooperate in the first period as long as Expression (5) is greater than Expression (6), which yields.

$$p \geq \frac{cn(1-n)}{cn(1-n)+(2-m)(1-m)} \tag{7}$$

Since $m, n < 1/2$, the right hand side of Inequality (7) is always smaller than the lower bound for p, $\frac{cn}{cn+(1-m)}$, and, therefore, Inequality (7) always holds.

References

George A Akerlof and Rachel E Kranton. Economics and identity. *The Quarterly Journal of Economics*, 115(3):715–753, 2000.

George A Akerlof and Rachel E Kranton. Identity and schooling: Some lessons for the economics of education. *Journal of economic literature*, 40(4):1167–1201, 2002.

George A Akerlof and Rachel E Kranton. Identity and the economics of organizations. *Journal of Economic perspectives*, 19(1):9–32, 2005.

George A Akerlof and Rachel E Kranton. Identity, supervision, and work groups. *American Economic Review*, 98(2):212–17, 2008.

Alberto Alesina, Reza Baqir, and William Easterly. Public goods and ethnic divisions. *The Quarterly Journal of Economics*, 114(4):1243–1284, 1999.

James Andreoni and Lise Vesterlund. Which is the fair sex? gender differences in altruism. *The Quarterly Journal of Economics*, 116(1):293–312, 2001.

Theodore C Bergstrom. On the evolution of altruistic ethical rules for siblings. *The American Economic Review*, pages 58–81, 1995.

Theodore C Bergstrom. The algebra of assortative encounters and the evolution of cooperation. *International Game Theory Review*, 5(03):211–228, 2003.

C Bram Cadsby, Maroš Servátka, and Fei Song. How competitive are female professionals? a tale of identity conflict. *Journal of Economic Behavior & Organization*, 92:284–303, 2013.

Gary Charness, Luca Rigotti, and Aldo Rustichini. Individual behavior and group membership. *American Economic Review*, 97(4):1340–1352, 2007.

Yan Chen and Sherry Xin Li. Group identity and social preferences. *American Economic Review*, 99(1):431–57, 2009.

Yan Chen, Sherry Xin Li, Tracy Xiao Liu, and Margaret Shih. Which hat to wear? impact of natural identities on coordination and cooperation. *Games and Economic Behavior*, 84:58–86, 2014.

Catherine C Eckel and Philip J Grossman. Managing diversity by creating team identity. *Journal of Economic Behavior & Organization*, 58(3):371–392, 2005.

Lorenz Goette, David Huffman, and Stephan Meier. The impact of group membership on cooperation and norm enforcement: Evidence using random assignment to real social groups. *American Economic Review*, 96(2):212–216, 2006.

Paul J Healy. Group reputations, stereotypes, and cooperation in a repeated labor market. *American Economic Review*, 97(5):1751–1773, 2007.

Esteban F Klor and Moses Shayo. Social identity and preferences over redistribution. *Journal of Public Economics*, 94(3-4):269–278, 2010.

Rachel E Kranton and Seth G Sanders. Groupy versus non-groupy social preferences: personality, region, and political party. *American Economic Review*, 107 (5):65–69, 2017.

Moses Shayo. A theory of social identity with an application to redistribution. 2007.