IKERLANAK

DISCRIMINATING BY TAGGING: ARTIFICIAL DISTINCTION, REAL DISCRIMINATION

by

Elena Iñarra and Annick Laruelle

2011

Working Paper Series: IL. 50/11

Departamento de Fundamentos del Análisis Económico I

Ekonomi Analisiaren Oinarriak I Saila

University of the Basque Country

Discriminating by Tagging: Artificial Distinction, Real Discrimination*

By Elena Inarra[†] and Annick Laruelle[‡§]

February 8, 2011

Abstract

We introduce a new variation of the hawk-dove game suggested by an experiment that studies the behavior of a group of domestic fowls when a subgroup has been marked. Specifically we consider a population formed by two types of individual that fail to recognize their own type but do recognize the other type. In this game we find two evolutionarily stable strategies. In each of them, individuals from one type are always attacked more, whatever proportion of the population they represent. Our theoretical results are consistent with the conclusions drawn from experimental work, where marked fowls received more pecks than their unmarked counterparts.(JEL C72)

*This research is supported by the Spanish Ministerio de Ciencia e Innovación under projects SEJ2006-05455 and ECO2009-11213, co-funded by ERDF, and by Basque Government funding for Grupo Consolidado GIC07/146-IT-377-07. We thank Inma Estevez for drawing our attention to this case study. We also thank Andre Barreira da Silva Rocha, Ken Binmore, Federico Grafe, Alan Grafen and Federico Valenciano for their comments. We also thank Marta San Martin for her help in the drawings. Elena Iñarra also gratefully acknowledges the hospitality of Oxford University and financial support from the Basque Goverment.

[†] *BRiDGE,* Departamento de Fundamentos del Análisis Económico I, Universidad del País Vasco, Avenida Lehendakari Aguirre, 83, E-48015 Bilbao, Spain; elena.inarra@ehu.es.

[‡] *BRiDGE,* Departamento de Fundamentos del Análisis Económico I, Universidad del País Vasco, Avenida Lehendakari Aguirre, 83, E-48015 Bilbao, Spain; a.laruelle@ikerbasque.org.

[§] IKERBASQUE, Basque Foundation of Science, 48011, Bilbao, Spain.

1 Introduction

Marking animals artificially changes phenotypes, since it manipulates physical appearance. Quite a few scientific experiments using animals are conducted by marking some of them. However, if this means that the animals under study undergo some sort of alteration in their behavior, the results of the experiment might be neither accurate nor representative of the whole group.

In an interesting experiment with domestic fowls, Dennis, Newberry, Cheng and Estevez (2008) find evidence of behavioral changes when different proportions of a population are marked on the back of their heads. In this experiment birds do not know whether or not they themselves are marked, but can see the marks on other birds. The two most salient results derived from this work are that marked birds suffer more aggression and have less body mass than their unmarked pen mates.

In game theory it is usually assumed that players know who they are, but they may or may not know the type of their opponents. To the best of our knowledge no situations where individuals lack self-perception but are able to perceive others have ever been modeled[1]. Here we consider this feature, which underlies the experimental work of Denis *et al.* (2008). More precisely we propose a variation of the classical hawk-dove game where individuals are unaware of their own type but see their opponent's type. The concept of "evolutionarily stable strategy" (Maynard Smith and Price, 1973) is used to solve the game.

The paper starts by recalling the results of the classical hawk-dove game in a finite population, which we refer to as the "homogenous game". Then we divide the population into two types: marked and unmarked individuals. The novelty of what we call the "heterogeneous game" is that individuals now meet two types of opponent. Consequently they can play either the same or different actions depending on the type of opponent that they face.

Though the modification of a population by marking a proportion of individuals is not linked to any disparity in capacity, we find that it affects the behavior of members. Playing the strategy that was evolutionarily stable in the homogeneous game against any type of opponent is not evolutionarily stable in the heterogeneous game. Indeed, no individual strategy that treats the two types of individual equally will be evolutionarily stable in heterogeneous games. Interestingly enough, we show that any such game has two evolutionarily stable strategies. We find that in each one, independently of the distribution of the types, one type of individual is always attacked more than the other. This has led us to refer to the type attacked more

[1] Two illustrations of this type of situation are a card game called the "Indian poker game" (See http://en.wikipedia.org/wiki/Blind_man's_bluff_poker or a sequence in Tarantino's movie "Inglourious Basterds" (2009)).

as aliens, with all the negative connotations of the term, and to the other type as locals. This result contradicts the intuition that aggressive behavior toward members of a minority is more probable than aggressive behavior toward members of the majority. In short, tagging generates real discrimination in which the discriminated type, the alien, is systematically treated worse. A comparison of the evolutionarily stable strategies for different proportions of aliens reveals that the higher the proportion of aliens is, the less likely aggressive behavior toward them is.

A comparison of individuals' expected payoffs when the evolutionarily stable strategies are played is then proposed. Not surprisingly, we find that an alien is always worse off than a local, and that an alien's expected payoff increases as the proportion of aliens within the population increases. For a local, however, the expected payoff increases as the proportion of aliens increases up to certain point and decreases thereafter. Furthermore for an individual taken at random the existence of a small proportion of aliens is beneficial whereas a larger proportion is detrimental. We are also able to determine the proportion of individuals that would have to be tagged for a local and a random individual to get their maximal expected payoff.

Our theoretical results support the conclusions obtained in the experimental work conducted by Dennis *et al.* (2008). Considering hawkish behavior as a proxy of the pecking and threatening between birds observed in the experiment, we find that a strategy of more aggressive behavior toward marked birds than toward unmarked ones can be evolutionarily stable. Moreover, the fact that in this experiment marked birds have less body mass than their unmarked pen mates is also supported under the assumption that the body mass of a bird can be evaluated through the expected payoff of an individual in a heterogeneous game.

Now let us compare our results with the relevant literature. In the seminal hawk-dove game only the mixed strategy in which the probability of each individual playing hawk is equal to the ratio between the value of the resource and the cost of fighting is evolutionarily stable. Maynard Smith and Parker (1976) propose a variation in which individuals fight for a territory, each player being either the "owner" or the "intruder". They show that the "bourgeois" strategy, *i.e.* playing hawk when one is the owner and dove when the intruder, is evolutionary stable. Going further, Selten (1980) proves that only pure strategies are evolutionarily stable. However, Binmore and Samuelson (2001a, 2001b) consider the two roles that an individual may play jointly with payoff perturbations, and show that under certain conditions mixed strategies can also be evolutionarily stable in this game. This last result goes along with the findings obtained in our variation. When individuals lack self-perception but recognize the type of the others pure as well as mixed strategies may be evolutionarily stable, depending on the proportion of aliens in a population.

This work can also be linked to the literature of social dynamics, in particular to the inspiring work by Axtell, Epstein and Young (1991). These authors consider the divide-one-dollar game, in which each individual may choose among the following three actions: high (ask for 70 cents), medium (ask for 50 cents) and low (ask for 30 cents) claims. The dollar is divided according to claims whenever they are feasible, otherwise players receive nothing. The dynamics of random bilateral encounters in large populations show that in the long run, any two players in the population tend to demand a medium claim. However, if the population is artificially divided into two groups then a discriminatory norm emerges in society. An equilibrium where the members of one group make a high claim and the members of the other group make a low claim when pairs belonging to different groups meet may persist for substantial periods of time, in which there is intra-group dissension. In a completely different setting we obtain a similar conclusion: random tagging of individuals within a population gives rise to true discrimination against one of the types, the aliens, who moreover behave aggressively among themselves.

The rest of the paper is organized as follows. Section 2 presents the homogeneous game and its unique evolutionarily stable strategy. Section 3 introduces the heterogeneous game and derives the evolutionarily stable strategies and expected payoffs. Section 4 analyzes the experimental work in light of our game theoretical results. Section 5 concludes.

2 The homogenous game

Consider a population of n identical individuals in which any pair face a contested resource of value v and may fight at a cost c. The size of the population n, the value v and the cost c are considered to be fixed with $v < c$. Each individual can be either aggressive and behave as a hawk or passive and behave as a dove. If an individual behaves as a hawk and their opponent as a dove, the aggressive individual gets the resource v while the passive individual gets nothing. If both individuals act like hawks, there is a fight. The winner gets the resource while the loser faces the cost of the fight c. Assuming that the two individuals have the same probability of winning the fight, the expected payoff for each one is half the resource minus the cost of fighting. If the two individuals behave like doves one withdraws and the other gets the resource. Assuming that the two individuals have the same probability of withdrawing the expected payoff for each one is half the resource. This description corresponds to the classical hawk-dove game played by a population of identical individuals, which we refer to as the “homogeneous game” Γ whose payoff matrix may be represented as follows[2]:

	hawk	dove
hawk	$(\frac{v-c}{2}, \frac{v-c}{2})$	$(v, 0)$
dove	$(0, v)$	$(\frac{v}{2}, \frac{v}{2})$

(1)

Let α denote the probability of playing hawk so that an individual can choose either a pure hawk ($\alpha = 1$) or dove ($\alpha = 0$) strategy or a mixed strategy ($0 < \alpha < 1$). Let $u(\alpha, \beta)$ be the expected payoff of an individual that plays α when their opponent plays β. That is,

$$u(\alpha, \beta) = \tfrac{v}{2}(1 - \beta) + \tfrac{c}{2}(\tfrac{v}{c} - \beta)\alpha. \tag{2}$$

Since game Γ is symmetric the opponent's expected payoff is given by $u(\alpha, \beta)$.

The concept of “evolutionarily stable strategy” introduced by Maynard Smith and Price (1973) is applied to solve the hawk-dove game. This notion captures the resilience of a given strategy against any other strategy in the following sense: Consider a population where most members play an evolutionarily stable strategy while a small fraction of mutants choose a different strategy. In this situation every mutant's expected payoff is smaller than the expected payoff of a "normal" individual, so that the mutants are driven out from the population[3].

An evolutionarily stable strategy may be formally determined as follows. Let $\mathcal{B}(\beta)$ denote the set of an individual's best responses to an opponent playing strategy β. Recall that a best response is a strategy that yields the highest payoff given the opponent's strategy. The two conditions for a strategy α^* to be evolutionarily stable are: *(i)* $\alpha^* \in \mathcal{B}(\alpha^*)$ and *(ii)* for any $\beta \neq \alpha^*$ such that $\beta \in \mathcal{B}(\alpha^*)$ we have $u(\alpha^*, \beta) > u(\beta, \beta)$. Condition *(i)* states that α^* has to be a best response to itself. That is, the pair of strategies (α^*, α^*) is a symmetric Nash equilibrium (Nash, 1951). Condition *(ii)* states that if the opponent plays a best response to α^* (other than α^*) then the payoff of playing α^* is strictly greater than the payoff of playing that best response.

For game Γ the set of an individual's best responses to an opponent playing β is

$$\mathcal{B}(\beta) = \begin{cases} \{1\} & \text{if } \beta < \frac{v}{c} \\ \{\alpha \mid \alpha \in [0,1]\} & \text{if } \beta = \frac{v}{c} \\ \{0\} & \text{if } \beta > \frac{v}{c}. \end{cases}$$

Thus, if the probability of the opponent playing hawk is smaller than the ratio between the resource and the cost, the unique best response is to play hawk, while if it is greater than

[2] If we had $v > c$ then the structure of the game would be equivalent to a prisoner's dilemma, while if we had $v = c$ then it would be equivalent to a coordination game.

[3] See Maynard Smith (1982), Chapter 2, and Weibull (1995), Chapter 2, for a detailed explanation of this notion. A good introduction can be found in Osborne (2004), Chapter 13.

that ratio the unique best response is to play dove. If the probability of the opponent playing hawk is equal to that ratio then any strategy is a best response. For this game strategy $\frac{v}{c}$ is the only evolutionarily stable strategy. It is the only strategy that is a best response to itself, $\frac{v}{c} \in \mathcal{B}(\frac{v}{c})$, and it satisfies Condition *(ii)*: $u(\frac{v}{c}, \beta) - u(\beta, \beta) = \frac{c}{2}(\frac{v}{c} - \beta)^2 > 0$ for $\beta \neq \frac{v}{c}$.

3 The heterogenous game

3.1 The model

Consider that a proportion x $(0 < x < 1)$ of a population of n individuals is artificially marked. Then there are two types of individual: the marked (M), and the unmarked (U). Assume that a pair of individuals is selected at random. This is equivalent to assuming that an individual is selected at random from a group of n individuals, and then the opponent is selected at random from the remaining $n-1$ individuals. Thus, at any bilateral encounter between two individuals four cases are possible: both individuals are unmarked (U, U); the first individual is unmarked and the second is marked (U, M); the first is marked and the second is unmarked (M, U); or both are marked (M, M). The probabilities of these four possible encounters, which we denote respectively by $p(U, U)$, $p(U, M)$, $p(M, U)$ and $p(M, M)$, are given by

$$p(M,M) = \frac{x(nx-1)}{n-1},\ p(U,U) = \frac{(1-x)(n-nx-1)}{n-1} \text{ and } p(U,M) = p(M,U) = \frac{x(1-x)n}{n-1}. \tag{3}$$

Since marks are made at random, they do not reflect genetic differences between individuals. Consequently the probability of winning or losing a fight is not determined by the presence or absence of marks. Hence the same hawk-dove game, whose payoff matrix given by (1), is played in each state of nature.

The key feature of this model is that individuals fail to recognize their own type but do recognize their opponent's type. This implies that the first individual does not distinguish between states (U, U) and (M, U) or between states (U, M) and (M, M) while the second individual does not distinguish between states (U, U) and (U, M) or between states (M, U) and (M, M).

What is a strategy in this context? Obviously individuals find themselves in a position of choosing a probability of playing hawk for each type of opponent, marked or unmarked. A *strategy* can thus be represented by $\boldsymbol{\alpha} = (\alpha_U, \alpha_M)$ where α_U gives the probability of behaving as a hawk to an unmarked individual and α_M gives the probability of behaving as a hawk to a marked one. Of course any strategy α played by an individual in a homogeneous game can be played against either type of opponent in a heterogeneous one, *i.e.* $\alpha_U = \alpha_M = \alpha$. Such a strategy is referred to here as homogeneous by contrast to a heterogeneous strategy where $\alpha_U \neq \alpha_M$. Thus there are two pure homogeneous strategies, which we represent as follows: hawk against either type of opponent $(1,1)$, and dove against either type of opponent $(0,0)$; and two pure heterogeneous strategies: $(0,1)$ playing dove against unmarked individuals and hawk against marked ones, and $(1,0)$ playing hawk against unmarked individuals and dove against marked ones.

The *expected payoff of an individual* playing $\boldsymbol{\alpha} = (\alpha_U, \alpha_M)$ while the opponent plays $\boldsymbol{\beta} = (\beta_U, \beta_M)$ is the sum of the expected payoffs she would obtain in every distinct encounter weighted by its probability of occurrence. For instance, in the encounter (U, M) the first individual recognizes her opponent as marked and plays hawk with probability α_M while the later recognizes the former as an unmarked and plays hawk with probability β_U, being $u(\alpha_M, \beta_U)$ the individual's expected payoff derived from such bilateral encounter. This expected payoff is multiplied by $p(U, M)$. The expected payoffs in the remaining encounters are defined analogously. Therefore the expected payoff of an individual playing $\boldsymbol{\alpha}$ against an opponent playing $\boldsymbol{\beta}$ is given by $U(\boldsymbol{\alpha}, \boldsymbol{\beta})$. That is,

$$U(\boldsymbol{\alpha},\boldsymbol{\beta}) = p(U,U)u(\alpha_U,\beta_U)+p(U,M)u(\alpha_M,\beta_U)+p(M,U)u(\alpha_U,\beta_M)+p(M,M)u(\alpha_M,\beta_M).$$

Using (2) and (3) we rewrite $U(\boldsymbol{\alpha}, \boldsymbol{\beta})$ as

$$\begin{aligned} U(\boldsymbol{\alpha},\boldsymbol{\beta}) &= \tfrac{v}{2}\left[1-(1-x)\beta_U - x\beta_M\right] \\ &+\tfrac{c}{2(n-1)}(1-x)\left[(n-1)\tfrac{v}{c} - (n-nx-1)\beta_U - nx\beta_M\right]\alpha_U \\ &+\tfrac{c}{2(n-1)}x\left[(n-1)\tfrac{v}{c} - (n-nx)\beta_U - (nx-1)\beta_M\right]\alpha_M. \end{aligned} \tag{4}$$

In addition, the individual's expected payoff can be decomposed into an unmarked individual's expected payoff ($U_U(\boldsymbol{\alpha},\boldsymbol{\beta})$) multiplied by the probability of being unmarked $(1-x)$ and a marked individual's expected payoff ($U_M(\boldsymbol{\alpha},\boldsymbol{\beta})$) multiplied by the probability of being marked (x). That is, $U(\boldsymbol{\alpha},\boldsymbol{\beta})$ can be written as

$$U(\boldsymbol{\alpha},\boldsymbol{\beta}) = (1-x)U_U(\boldsymbol{\alpha},\boldsymbol{\beta}) + xU_M(\boldsymbol{\alpha},\boldsymbol{\beta})$$

where

$$\begin{aligned} U_U(\boldsymbol{\alpha},\boldsymbol{\beta}) &= \tfrac{v}{2}(1-\beta_U) + \tfrac{c}{2}(\tfrac{v}{c}-\beta_U)\tfrac{n-nx-1}{n-1}\alpha_U + \tfrac{c}{2}(\tfrac{v}{c}-\beta_U)\tfrac{nx}{n-1}\alpha_M, \\ U_M(\boldsymbol{\alpha},\boldsymbol{\beta}) &= \tfrac{v}{2}(1-\beta_M) + \tfrac{c}{2}(\tfrac{v}{c}-\beta_M)\tfrac{n-nx}{n-1}\alpha_U + \tfrac{c}{2}(\tfrac{v}{c}-\beta_M)\tfrac{nx-1}{n-1}\alpha_M. \end{aligned} \tag{5}$$

Note the similarity between (2) and (5).

We have modeled a population formed by two types of individual who play a hawk-dove game. The main characteristic is that individuals fail to perceive their own type but recognize the type of their opponents. The probabilities of the different types of encounter, the strategies and expected payoffs are defined. Thus, we have all the ingredients of a game, hereafter referred to as a *heterogeneous game* and denoted by Γ_x, where $0 < x < 1$ is the proportion of marked individuals, which is the key parameter in this paper.

We proceed to solve the heterogeneous game Γ_x by applying the concept of evolutionarily stable strategy as for game Γ. Let $\mathcal{B}_x(\boldsymbol{\beta})$ be the set of an individual's best responses to an opponent playing $\boldsymbol{\beta}$. Strategy $\boldsymbol{\alpha}^*$ is *evolutionarily stable* if and only if *(i)* $\boldsymbol{\alpha}^* \in \mathcal{B}_x(\boldsymbol{\alpha}^*)$, and *(ii)* for any $\boldsymbol{\beta} \in \mathcal{B}_x(\boldsymbol{\beta})$ such that $\boldsymbol{\beta} \neq \boldsymbol{\alpha}^*$ we have $U(\boldsymbol{\alpha}^*, \boldsymbol{\beta}) > U(\boldsymbol{\beta}, \boldsymbol{\beta})$.

3.2 Best responses

First, we determine the set of an individual's best responses given their opponent's strategy. If her opponent plays strategy $\boldsymbol{\beta} = (\beta_U, \beta_M)$ the best response of an individual is to choose $\boldsymbol{\alpha} = (\alpha_U, \alpha_M)$ such that $U(\boldsymbol{\alpha}, \boldsymbol{\beta})$ is maximized. This best choice appears more clearly if we rewrite (4) as follows

$$U(\boldsymbol{\alpha}, \boldsymbol{\beta}) = f_0(\boldsymbol{\beta}) + f_U(\boldsymbol{\beta})\ \alpha_U + f_M(\boldsymbol{\beta})\ \alpha_M$$

where

$$\begin{aligned}
f_0(\boldsymbol{\beta}) &= \tfrac{v}{2}\left[1 - (1-x)\beta_U - x\beta_M\right] \\
f_U(\boldsymbol{\beta}) &= \tfrac{c}{2(n-1)}(1-x)nx\left[\tfrac{n-1}{nx}\tfrac{v}{c} - \tfrac{n-nx-1}{nx}\beta_U - \beta_M\right] && (6) \\
f_M(\boldsymbol{\beta}) &= \tfrac{c}{2(n-1)}x(nx-1)\left[\tfrac{n-1}{nx-1}\tfrac{v}{c} - \tfrac{n-nx}{nx-1}\beta_U - \beta_M\right]. && (7)
\end{aligned}$$

The best choice of an individual is $\alpha_U = 1$ whenever $f_U(\boldsymbol{\beta}) > 0$, $\alpha_U = 0$ whenever $f_U(\boldsymbol{\beta}) < 0$ and any α_U whenever $f_U(\boldsymbol{\beta}) = 0$. Similarly the choice of α_M depends on the sign of $f_M(\boldsymbol{\beta})$. Thus $\mathcal{B}_x(\boldsymbol{\beta})$, the set of an individual's best responses to an opponent playing $\boldsymbol{\beta}$, is given by

$$\mathcal{B}_x(\boldsymbol{\beta}) = \begin{cases}
\{(\zeta_U, \zeta_M) \mid \zeta_U, \zeta_M \in [0,1]\} & \text{if } [f_U(\boldsymbol{\beta}) = 0 \text{ and } f_M(\boldsymbol{\beta}) = 0] \\
\{(\alpha_U, 1) \mid \alpha_U \in [0,1]\} & \text{if } [f_U(\boldsymbol{\beta}) = 0 \text{ and } f_M(\boldsymbol{\beta}) > 0] \\
\{(0,1)\} & \text{if } [f_U(\boldsymbol{\beta}) < 0 \text{ and } f_M(\boldsymbol{\beta}) > 0] \\
\{(0, \alpha_M) \mid \alpha_M \in [0,1]\} & \text{if } [f_U(\boldsymbol{\beta}) < 0 \text{ and } f_M(\boldsymbol{\beta}) = 0] \\
\{(1, \gamma_M) \mid \gamma_M \in [0,1]\} & \text{if } [f_U(\boldsymbol{\beta}) > 0 \text{ and } f_M(\boldsymbol{\beta}) = 0] \\
\{(1,0)\} & \text{if } [f_U(\boldsymbol{\beta}) > 0 \text{ and } f_M(\boldsymbol{\beta}) < 0] \\
\{(\gamma_U, 0) \mid \gamma_U \in [0,1]\} & \text{if } [f_U(\boldsymbol{\beta}) = 0 \text{ and } f_M(\boldsymbol{\beta}) < 0] \\
\{(1,1)\} & \text{if } [f_U(\boldsymbol{\beta}) > 0 \text{ and } f_M(\boldsymbol{\beta}) > 0] \\
\{(0,0)\} & \text{if } [f_U(\boldsymbol{\beta}) < 0 \text{ and } f_M(\boldsymbol{\beta}) < 0].
\end{cases} \qquad (8)$$

Second, we determine the strategies that are best responses to themselves. As will be proven in the next theorem only three strategies satisfy this property. The first one is independent of the proportion of marked individuals (x), while for the other two the probability of behaving as a hawk depends on x. We denote them by $\frac{\mathbf{v}}{\mathbf{c}}$, $\boldsymbol{\alpha}_x^*$ and $\boldsymbol{\gamma}_x^*$, respectively. They are given by

$$\begin{aligned}
\frac{\mathbf{v}}{\mathbf{c}} &= (\tfrac{v}{c}, \tfrac{v}{c}) \\
\boldsymbol{\alpha}_x^* &= \begin{cases} (\frac{v}{c} - (1-\frac{v}{c})\frac{nx}{n-nx-1}, 1) & \text{if } x \leq \underline{w} \\ (0,1) & \text{if } \underline{w} < x < \overline{w} \\ (0, \frac{n-1}{nx-1}\frac{v}{c}) & \text{if } x \geq \overline{w} \end{cases} \qquad (9) \\
\boldsymbol{\gamma}_x^* &= \begin{cases} (\frac{n-1}{n(1-x)-1}\frac{v}{c}, 0) & \text{if } x \leq 1-\overline{w} \\ (1,0) & \text{if } 1-\overline{w} < x < 1-\underline{w} \\ (1, \frac{v}{c} - (1-\frac{v}{c})\frac{n(1-x)}{n-n(1-x)-1}) & \text{if } x \geq 1-\underline{w}. \end{cases} \qquad (10) \\
\text{where } \underline{w} &= \tfrac{v}{c}(1-\tfrac{1}{n}) \text{ and } \overline{w} = \tfrac{v}{c}(1-\tfrac{1}{n}) + \tfrac{1}{n}.
\end{aligned}$$

We can now state and prove the following result:

Theorem 1 *In any heterogeneous game Γ_x, only strategies $\frac{\mathbf{v}}{\mathbf{c}}$, $\boldsymbol{\alpha}_x^*$ and $\boldsymbol{\gamma}_x^*$ are best responses to themselves.*

Proof. Using (7) and (8) we check that these three strategies satisfy $\boldsymbol{\beta} \in \mathcal{B}_x(\boldsymbol{\beta})$ and that no other strategy does.

1. We have that $\frac{\mathbf{v}}{\mathbf{c}} \in \mathcal{B}_x(\frac{\mathbf{v}}{\mathbf{c}})$ since $f_U((\zeta_U, \zeta_M)) = f_M((\zeta_U, \zeta_M)) = 0$ iff $\zeta_U = \zeta_M = \frac{v}{c}$.

2. For strategy $\boldsymbol{\alpha}_x^*$, we first check that $f_U((\alpha_U, 1)) = 0$ iff $\alpha_U = \frac{v}{c} - (1-\frac{v}{c})\frac{nx}{n-nx-1}$, jointly with $\frac{v}{c} - (1-\frac{v}{c})\frac{nx}{n-nx-1} \geq 0$ for $x \leq \underline{w}$, and $f_M((\frac{v}{c} - (1-\frac{v}{c})\frac{nx}{n-nx-1}, 1)) > 0$. Second, we have $f_U((0,1)) < 0$ if $x > \underline{w}$ and $f_M((0,1)) > 0$ if $x < \overline{w}$. Third, $f_M((0, \alpha_M)) = 0$ iff $\alpha_M = \frac{n-1}{nx-1}\frac{v}{c}$, jointly with $\frac{n-1}{nx-1}\frac{v}{c} \leq 1$ if $x \geq \overline{w}$ and $f_U((0, \frac{n-1}{nx-1}\frac{v}{c})) < 0$. Hence $\boldsymbol{\alpha}_x^* \in \mathcal{B}_x(\boldsymbol{\alpha}_x^*)$.

3. For strategy $\boldsymbol{\gamma}_x^*$, we first check that $f_U((\gamma_U, 0)) = 0$ iff $\gamma_U = \frac{n-1}{n(1-x)-1}\frac{v}{c}$, jointly with $\frac{n-1}{n(1-x)-1}\frac{v}{c} \leq 1$ if $x \leq 1-\overline{w}$, and $f_M((\frac{n-1}{n(1-x)-1}\frac{v}{c}, 0)) < 0$. Second, we have $f_U((1,0)) > 0$ if $x > 1-\overline{w}$ while $f_M((1,0)) < 0$ if $x < 1-\underline{w}$. Third, $f_M((1,\gamma_M)) = 0$ iff $\gamma_M = \frac{v}{c} - (1-\frac{v}{c})\frac{n(1-x)}{n-n(1-x)-1}$, jointly with $\frac{v}{c} - (1-\frac{v}{c})\frac{n(1-x)}{n-n(1-x)-1} \geq 0$ for $x \geq 1-\underline{w}$ and $f_U((0, \frac{v}{c} - (1-\frac{v}{c})\frac{n(1-x)}{n-n(1-x)-1})) > 0$. Hence $\boldsymbol{\gamma}_x^* \in \mathcal{B}_x(\boldsymbol{\gamma}_x^*)$.

4. It remains to show that no other strategy can be a best response to itself, which is done by checking that $f_U((1,1)) < 0$, and $f_U((0,0)) > 0$.

■

Thus, we have one homogeneous and two heterogeneous strategies that are best responses to themselves. That is, we have three symmetric Nash equilibria. The homogeneous strategy $\frac{\mathbf{v}}{\mathbf{c}}$ corresponds to the evolutionarily strategy $\frac{v}{c}$ in the homogeneous game. The other two strategies are heterogeneous strategies. Note that the strategy is adopted by all individuals (unmarked and marked). Obviously it could not be otherwise since in our game individuals do not know their own type.

3.3 Evolutionarily stable strategies

The next question is whether the strategies that are best responses to themselves are evolutionarily stable. As the following result shows, strategy $\frac{\mathbf{v}}{\mathbf{c}}$ (where individuals have a probability of playing hawk against any type of opponent of $\frac{v}{c}$) does not satisfy this property.

Theorem 2 *In any heterogeneous game Γ_x, strategy $\frac{\mathbf{v}}{\mathbf{c}}$ is not evolutionarily stable.*

Proof. It is immediately apparent that $f_U(\frac{\mathbf{v}}{\mathbf{c}}) = f_M(\frac{\mathbf{v}}{\mathbf{c}}) = 0$, and by (8) any strategy is a best response to strategy $\frac{\mathbf{v}}{\mathbf{c}}$. To show that $\frac{\mathbf{v}}{\mathbf{c}}$ is not evolutionarily stable, choose a strategy $\boldsymbol{\beta}$ such that the difference $U(\frac{\mathbf{v}}{\mathbf{c}}, \boldsymbol{\beta}) - U(\boldsymbol{\beta}, \boldsymbol{\beta})$ is negative. Using (4) this difference can be written as

$$U(\tfrac{\mathbf{v}}{\mathbf{c}}, \boldsymbol{\beta}) - U(\boldsymbol{\beta}, \boldsymbol{\beta})) = \tfrac{c}{2(n-1)} \left[(\tfrac{v}{c} - \beta_U)^2 (n - nx - 1)(1 - x) + \right. \\ \left. 2(\tfrac{v}{c} - \beta_U)(\tfrac{v}{c} - \beta_M) n(1 - x)x + (\tfrac{v}{c} - \beta_M)^2 (nx - 1)x \right].$$

(i) If $x = \frac{n-1}{n}$ the difference above is reduced to

$$U(\tfrac{\mathbf{v}}{\mathbf{c}}, \boldsymbol{\beta}) - U(\boldsymbol{\beta}, \boldsymbol{\beta}) = \tfrac{c}{2n}(\tfrac{v}{c} - \beta_M) \left[2(\tfrac{v}{c} - \beta_U) + (\tfrac{v}{c} - \beta_M)(n - 2) \right]$$

which turns out to be negative if strategy $\boldsymbol{\beta}$ is chosen such that

$$0 < \beta_M < \tfrac{v}{c} \text{ and } \tfrac{v}{c} < \beta_U < 1 \text{ so that } \beta_U - \tfrac{v}{c} > \tfrac{n-2}{2}(\tfrac{v}{c} - \beta_M). \tag{11}$$

(ii) If $x \neq \frac{n-1}{n}$, denote $Z = \frac{\frac{v}{c} - \beta_U}{\beta_M - \frac{v}{c}}$. In this case the difference is a quadratic equation in Z:

$$U(\tfrac{\mathbf{v}}{\mathbf{c}}, \boldsymbol{\beta}) - U(\boldsymbol{\beta}, \boldsymbol{\beta}) = \tfrac{c}{2(n-1)}(\beta_M - \tfrac{v}{c})^2 \left[(n - nx - 1)(1 - x)Z^2 - 2n(1 - x)xZ + (nx - 1)x \right]$$

whose discriminant is $\Delta = 4(1 - x)x(n - 1) > 0$. Thus, the difference under study is negative for any Z such that $\frac{2n(1-x)x - \sqrt{\Delta}}{2(n-nx-1)(1-x)} < Z < \frac{2n(1-x)x + \sqrt{\Delta}}{2(n-nx-1)(1-x)}$ and, in particular, for a strategy $\boldsymbol{\beta} = (\beta_U, \beta_M)$ such that the following equality is satisfied:

$$\frac{\frac{v}{c} - \beta_U}{\beta_M - \frac{v}{c}} = \frac{nx}{n - nx - 1}. \tag{12}$$

■

Thus, while probability $\frac{v}{c}$ of playing hawk is evolutionarily stable in the homogeneous game, playing this strategy against any type of opponent is not evolutionarily stable in the heterogeneous game. At first sight, this result may appear surprising as the differentiation introduced within the members of the population is merely artificial. But from the proof of Theorem 2 it can be easily understood why strategy $\frac{\mathbf{v}}{\mathbf{c}}$ is not evolutionarily stable. The strategies $\boldsymbol{\beta}$ that satisfy (11) or (12) perform better against $\frac{\mathbf{v}}{\mathbf{c}}$ than strategy $\frac{\mathbf{v}}{\mathbf{c}}$ does against itself. These strategies are such that the probability of aggressive behavior toward one type is higher than $\frac{v}{c}$ while the probability of aggressive behavior toward the other type is smaller than $\frac{v}{c}$.

By contrast, the other two strategies, $\boldsymbol{\alpha}_x^*$ and $\boldsymbol{\gamma}_x^*$, which are best responses to themselves, are evolutionarily stable strategies. This is the main result of the paper.

Theorem 3 *In any heterogeneous game Γ_x, only strategies $\boldsymbol{\alpha}_x^*$ and $\boldsymbol{\gamma}_x^*$ are evolutionarily stable.*

Proof. By Theorem 1 we know that the only strategies that are best responses to themselves are $\boldsymbol{\alpha}_x^*$, $\boldsymbol{\gamma}_x^*$ and $\frac{\mathbf{v}}{\mathbf{c}}$. By Theorem 2 it is known that $\frac{\mathbf{v}}{\mathbf{c}}$ is not evolutionarily stable. Hence, it remains only to analyze strategies $\boldsymbol{\alpha}_x^*$ and $\boldsymbol{\gamma}_x^*$.

Using (8) the set of an individual's best responses to an opponent playing $\boldsymbol{\alpha}_x^*$ is

$$\mathcal{B}_x(\boldsymbol{\alpha}_x^*) = \begin{cases} \{(\alpha_U, 1) \mid \alpha_U \in [0,1]\} & \text{if } x \leq \underline{w} \\ \{(0,1)\} & \text{if } \underline{w} < x < \overline{w} \\ \{(0, \alpha_M) \mid \alpha_M \in [0,1]\} & \text{if } \overline{w} \leq x. \end{cases}$$

For $\underline{w} < x < \overline{w}$ we have that $\boldsymbol{\alpha}_x^* = (0,1)$ is the *only* best response to itself, hence Condition *(ii)* of the evolutionarily stable strategy definition does not need to be checked. For the other values of x, however, we must check whether that the difference $U(\boldsymbol{\alpha}_x^*, \boldsymbol{\beta}) - U(\boldsymbol{\beta}, \boldsymbol{\beta})$ is strictly positive for any $\boldsymbol{\beta} \in \mathcal{B}_x(\boldsymbol{\alpha}_x^*)$, with $\boldsymbol{\beta} \neq \boldsymbol{\alpha}_x^*$. Using (4) we obtain that

$$U(\boldsymbol{\alpha}_x^*, \boldsymbol{\beta}) - U(\boldsymbol{\beta}, \boldsymbol{\beta}) = \begin{cases} \frac{c}{2}\frac{(1-x)(n-nx-1)}{n-1}(\frac{v}{c} - (1-\frac{v}{c})\frac{nx}{n-nx-1} - \alpha_U)^2 & \text{if } x \leq \underline{w} \\ \frac{c}{2}\frac{x(nx-1)}{n-1}(\frac{n-1}{nx-1}\frac{v}{c} - \alpha_M)^2 & \text{if } \overline{w} \leq x \end{cases}$$

This difference is strictly positive if $\boldsymbol{\beta} \neq \boldsymbol{\alpha}_x^*$. Therefore strategy $\boldsymbol{\alpha}_x^*$ is evolutionarily stable.

The proof that strategy $\boldsymbol{\gamma}_x^*$ is evolutionarily stable is omitted because it is similar to the previous one. ■

So in general the evolutionarily stable strategies are mixed strategies, although they may be pure strategies for some specific proportions of marked individuals. This is the case when the proportion of marked individuals x is equal to $\frac{v}{c}$, *i.e.* for game $\Gamma_{v/c}$. In this case an interesting comparison between $\frac{v}{c}$, the evolutionarily stable strategy in game Γ, and $(0,1)$, the evolutionarily stable strategy in game $\Gamma_{v/c}$, can be made. In game $\Gamma_{v/c}$ an individual who plays strategy $(0,1)$ behaves as a hawk whenever her opponent is marked which occurs with frequency $\frac{v}{c}$. Therefore strategies $\frac{v}{c}$ in Γ and $(0,1)$ in $\Gamma_{v/c}$ are similar in the sense that a probability in game Γ is substituted by a frequency in game $\Gamma_{v/c}$. Furthermore strategy $(0,1)$ is also evolutionarily stable in game Γ_x for values of x close to $\frac{v}{c}$, *i.e.* $\underline{w} < x < \overline{w}$. Note, however, that for values of x smaller than $\underline{w}$, an individual that plays strategy $(0,1)$ plays dove so often that their opponent is better off playing hawk at probabilities greater than $\frac{v}{c}$ while the opposite occurs for values of x higher than $\overline{w}$. Analogously an individual who plays strategy $(1,0)$ in game $\Gamma_{1-v/c}$ plays hawk with a frequency of $1-x=v/c$. The same comparison can be made between $\frac{v}{c}$ in game Γ, and strategy $(1,0)$ in game $\Gamma_{1-v/c}$.

Now let us move on to the expected payoffs for the evolutionarily stable strategies. Let us start with $\boldsymbol{\alpha}_x^*$. Given game Γ_x, plugging (9) into (4) we obtain

$$U(\boldsymbol{\alpha}_x^*, \boldsymbol{\alpha}_x^*) = \begin{cases} \frac{v}{2}(1-\frac{v}{c}) + \frac{c^2-v^2}{2c}\frac{x}{n-nx-1} & \text{if } x \leq \underline{w} \\ \frac{v}{2} - \frac{nx-1}{n-1}\frac{cx}{2} & \text{if } \underline{w} < x < \overline{w} \\ \frac{v}{2}(1-\frac{v}{c}) - \frac{v^2}{2c}\frac{1-x}{nx-1} & \text{if } x \geq \overline{w}. \end{cases}$$

Observe that the payoff of an individual is larger for smaller proportions of marked individuals than for large proportions of individuals. In the next section we come back to this result and explain why this is so.

If we decompose the payoff according to the type of individuals we have that (5) yields to

$$U_U(\boldsymbol{\alpha}_x^*, \boldsymbol{\alpha}_x^*) = \begin{cases} \frac{v}{2}(1-\frac{v}{c}) + \frac{c-v}{c}\frac{vnx}{n-nx-1} & \text{if } x \leq \underline{w} \\ \frac{v}{2}(1-\frac{v}{c}) + \frac{v}{2c}\frac{v(n-1)+cnx}{n-1} & \text{if } \underline{w} < x < \overline{w} \\ \frac{v}{2}(1-\frac{v}{c}) + \frac{v^2}{2c}\frac{2nx-1}{nx-1} & \text{if } x \geq \overline{w}, \end{cases}$$

and

$$U_M(\boldsymbol{\alpha}_x^*, \boldsymbol{\alpha}_x^*) = \begin{cases} \frac{v}{2}(1-\frac{v}{c}) - \frac{c-v}{2c}\frac{2nv(1-x)-c-v}{n-nx-1} & \text{if } x \leq \underline{w} \\ \frac{v}{2}(1-\frac{v}{c}) - \frac{c-v}{2c}\frac{(n-1)v+(nx-1)c}{n-1} & \text{if } \underline{w} < x < \overline{w} \\ \frac{v}{2}(1-\frac{v}{c}) - \frac{v}{c}\frac{n(1-x)v}{nx-1} & \text{if } x \geq \overline{w}. \end{cases}$$

Clearly an unmarked individual always obtains a larger payoff than a marked individual. The reverse holds in the second evolutionarily stable strategy. Moreover, by plugging (10) into (4) and (5) we can check that the following relations hold:

$$
\begin{aligned}
U(\boldsymbol{\gamma}_x^*, \boldsymbol{\gamma}_x^*) &= U(\boldsymbol{\alpha}_{1-x}^*, \boldsymbol{\alpha}_{1-x}^*) && (13)\\
U_U(\boldsymbol{\gamma}_x^*, \boldsymbol{\gamma}_x^*) &= U_M(\boldsymbol{\alpha}_{1-x}^*, \boldsymbol{\alpha}_{1-x}^*) && (14)\\
U_M(\boldsymbol{\gamma}_x^*, \boldsymbol{\gamma}_x^*) &= U_U(\boldsymbol{\alpha}_{1-x}^*, \boldsymbol{\alpha}_{1-x}^*). && (15)
\end{aligned}
$$

3.4 Interpretation of the results in terms of aliens and locals

The two evolutionarily stable strategies, $\boldsymbol{\alpha}_x^*$ and $\boldsymbol{\gamma}_x^*$, are the two faces of a single coin, the only difference being the norm that determines the type of individual treated worse.

Let us start by examining strategy $\boldsymbol{\alpha}_x^*$, given by (9). Note that, whatever x, the probability of aggression toward marked individuals is always greater than toward unmarked ones. The reverse holds for strategy $\boldsymbol{\gamma}_x^*$ given by (10). So in each strategy one type is treated worse than the other: the marked individuals in strategy $\boldsymbol{\alpha}_x^*$ and the unmarked individuals in strategy $\boldsymbol{\gamma}_x^*$. Let us refer to the worse-treated individuals as *aliens* and to the better-treated individuals as *locals*. A closer look at (9) and (10) reveals that the probabilities of aggression toward aliens/locals are identical in both strategies if those probabilities are expressed as a function of the proportion of aliens.

That is, the two strategies are identical but refer to different norms that define who are the aliens and who are the locals. Strategy $\boldsymbol{\alpha}_x^*$ corresponds to the norm that defines aliens as marked individuals, in proportion x, while strategy $\boldsymbol{\gamma}_x^*$ corresponds to the norm that defines aliens as unmarked individuals, in proportion $1 - x$.

More precisely let y denote the proportion of aliens and let $g_A(y)$ be the probability of aggression toward aliens and $g_L(y)$ be the probability of aggression toward locals. If we define

$$
\begin{aligned}
g_L(y) &= \begin{cases} \frac{v}{c} - (1 - \frac{v}{c})\frac{ny}{n-ny-1} & \text{if } y \le \underline{w} \\ 0 & \text{otherwise} \end{cases} \\
g_A(y) &= \begin{cases} 1 & \text{if } y < \overline{w} \\ \frac{n-1}{ny-1}\frac{v}{c} & \text{otherwise,} \end{cases} && (16)
\end{aligned}
$$

we can rewrite (9) and (10) as $\boldsymbol{\alpha}_x^* = (g_L(x), g_A(x))$ and $\boldsymbol{\gamma}_x^* = (g_A(1-x), g_L(1-x))$ clearly showing that the two strategies are basically identical. We can speak of $g_A(y)$ as the probability of hawkish behavior toward an alien, and of $g_L(y)$ as the probability of hawkish behavior toward a local if one evolutionarily stable strategy is played by all individuals. In *Figure* 1 we plot $g_A(y)$ and $g_L(y)$ for $n = 10$ and $v/c = 1/3$.

Figure 1: Probabilities of a hawk behavior toward a local (g_L) and toward an alien (g_A) as a function of the proportion of aliens (y)

The following comments can be made on the trend in the probability of aggression. The probability of aggression toward aliens is 1 for any proportion of aliens smaller than $\overline{w}$, and then it decreases toward $\frac{v}{c}$. The probability of aggression toward locals, however, decreases from $\frac{v}{c}$ to 0 for a proportion of aliens smaller than $\underline{w}$ and remains at 0 thereafter. Furthermore, aggression toward aliens is always greater than the aggression suffered by individuals in a homogeneous game ($g_A(y) > \frac{v}{c}$) while the reverse holds for locals ($g_L(y) < \frac{v}{c}$).

Similarly we may wonder what happens to the trend in payoffs assuming that all individuals play an evolutionarily stable strategy. Since the equality (13) holds we can speak of the payoff as a function of the proportion of aliens y. We denote it by $U^*(y)$, and define it as $U^*(y) = U(\boldsymbol{\alpha}_y^*, \boldsymbol{\alpha}_y^*)$. To facilitate the interpretation of the trend in $U^*(y)$ we take as reference the payoff obtained in the homogeneous game Γ if all individuals play the evolutionarily stable strategy. We denote this by $\bar{U}^*$. From (2) we determine $\bar{U}^* = u(\frac{v}{c}, \frac{v}{c}) = \frac{v}{2}(1 - \frac{v}{c})$. In *Figure 2* we represent $U^*(y)$ and $\bar{U}^*$ graphically for $n = 10$, $v/c = 1/3$.

Figure 2: Payoff in the homogeneous game ($\bar{U}^*$) and payoff of a random individual in the heterogeneous game ($U^*(y)$) as a function of the proportion of aliens (y)

This figure shows that introducing aliens is beneficial for individuals as long as their proportion is smaller than $\underline{w}$. For proportions of aliens greater than $\overline{w}$ the contrary effect arises. The maximal payoff is obtained for a proportion of aliens for a proportion of $\underline{w}$.

We can also study the trend in the payoff of marked and unmarked individuals. Once more, since the equality (14) holds we can speak of the payoffs of a local, $U_L^*(y)$, given as a function of the proportion of aliens y. Define this as $U_L^*(y) = U_U(\boldsymbol{\alpha}_y^*, \boldsymbol{\alpha}_y^*)$. Similarly equality (15) allows us to define $U_A^*(y) = U_M(\boldsymbol{\alpha}_y^*, \boldsymbol{\alpha}_y^*)$ as being the payoff of an alien. In *Figure 3* we plot $U_L^*(y)$ and $\bar{U}^*$ while in *Figure 4* we plot $U_A^*(y)$ and $\bar{U}^*$ for $n = 10$ and $v/c = 1/3$.

Figure 3: Payoff in the homogeneous game ($\bar{U}^*$) and payoff of a local in the heterogeneous game ($U_L^*(y)$) as a function of the proportion of aliens (y)

Figure 4: Payoff in the homogeneous game ($\bar{U}^*$) and payoff of an alien in the heterogeneous game ($U_A^*(y)$) as a function of the proportion of aliens (y)

At first glance these figures immediately reveal that the situation is as expected: aliens are worse off than individuals in a homogeneous game, while locals are better off. The maximal payoff for a local is obtained for a proportion of aliens of $\overline{w}$.

A closer analysis enables the trend in these payoffs according to the proportion of aliens to be explained. Two effects are involved: suffering more or less aggression and, in case of constant aggression, the response to the level of aggression suffered.

The less aggression suffered, the greater the well-being. This property partially explains the trend in the payoffs. *Figure* 1 shows that $g_L(y)$ decreases for values of y smaller than $\underline{w}$ and $g_A(y)$ decreases for values of y larger than $\overline{w}$. Consequently for values of y smaller than $\underline{w}$ a local's payoff increases, and for values of y larger than $\overline{w}$ an alien's payoff also increases (see *Figures* 3 and 4). This increase in payoffs for the same intervals is also observed for a random individual (see *Figure* 2).

When the probability of an aggression is constant, the trend in the payoff may be explained by the response to the level of aggression. *Figure* 1 shows that function $g_L(y)$ is 0 for values of y greater than $\underline{w}$, meaning that locals receive no aggression. In that case their best response is to play hawk. But for y greater than $\overline{w}$ they play hawk less often ($g_A(y)$ decreases and $g_L(y) = 0$)). Therefore in this interval the payoff of locals decreases. Analogously function $g_A(y)$ is 1 for values of y smaller than $\overline{w}$, meaning that aliens receive maximum aggression. In that case their best response is to play dove. For y smaller than $\underline{w}$ they play hawk less often ($g_L(y)$ decreases and $g_A(y) = 1$). Therefore in this interval the payoff of aliens increases.

It remains to explain the trend in the payoffs for values of y lying between $\underline{w}$ and $\overline{w}$ where locals receive no aggression and aliens receive maximum aggression. The best response for a local is to play hawk and for an alien it is to play dove. As the proportion of aliens increases in this interval, hawkish behavior increases, which turns out to be beneficial for locals and harmful for aliens. For a random individual the overall effect is negative.

The following proposition summarizes the main features that are observed in the foregoing figures. The proof is omitted because of its simplicity.

Proposition 1 *For any proportion of aliens $0 < y < 1$ we have that:*

(i) $U_A^*(y) < \bar{U}^* < U_L^*(y)$, (17)

(ii) $U^*(y) > \bar{U}^*$ *if* $y < \underline{w}$ *and* $U^*(y) < \bar{U}^*$ *if* $y > \overline{w}$,

(iii) $U_L^*(y)$ *is maximal for* $y = \overline{w}$ *and* $U^*(y)$ *is maximal for* $y = \underline{w}$.

4 The experiment

Dennis *et al.* (2008) have conducted several experiments with groups of domestic fowls. They consider group sizes of 10 and 50 birds in which different proportions (20%, 50% and 100% respectively) are marked. They study the birds' aggressive behavior measured by the number of pecks and threats in the encounters between them. The most significant results of this experiment are: *(i)* Marked domestic fowls receive more pecks than their unmarked pen mates. *(ii)* Marked domestic fowls in the 20% group receive significantly more threats than domestic fowls in the 100% marked group. *(iii)* There is no significant difference in the aggression received by marked fowls in the 20 and 50% marked groups. *(iv)* Aggressiveness toward marked fowls in populations with 100% of marked birds is lower than in any mixed population. *(v)* Marked fowls have a lower body mass than their unmarked pen mates.

To evaluate these experimental results in light of our model we assume that behaving as a hawk is a good proxy for the pecking and threatening between birds observed in the experiment. We also assume that the strategy played by the population is the evolutionarily stable strategy $\boldsymbol{\alpha}_x^* = (g_L(x), g_A(x))$. It seems reasonable to assume that the norm is for the aliens to be the marked fowls. In addition we consider that the expected payoff of an individual can be used as a proxy for a bird's body mass. With these assumptions we find that some of our theoretical results are consistent with the experimental ones:

(i) Marked domestic fowls receive more pecks than their unmarked pen mates. By (16) we have $g_L(x) < g_A(x)$.

(ii) Marked domestic fowls in the 20% group receive significantly more threats than domestic fowls in the 100% marked group. By (16) we have $g_A(0.2) = 5\frac{(n-1)}{n-5}\ \frac{v}{c} > \frac{v}{c}$.

(*iii*) *There is no significant difference in the aggression received by marked birds in the 20 and 50% marked groups.* By (16) we have $g_A(0.5) = g_A(0.2) = 1$ if $0.5 \leq \frac{v}{c}(1 - \frac{1}{n})$.

Therefore this empirical result is supported by our theoretical findings under the assumption that the value of the resource v is basically greater than half the cost c.

(iv) Aggressiveness toward marked birds in populations with 100% of marked birds is lower than any other mixed population. By (16) we have that for any $0 < x < 1$, $g_A(x) > \frac{v}{c}$.

(v) Marked birds have a lower body mass than their unmarked pen mates. By (17) we have $U_M(\boldsymbol{\alpha}_x^*, \boldsymbol{\alpha}_x^*) < U_U(\boldsymbol{\alpha}_x^*, \boldsymbol{\alpha}_x^*)$.

When we determine the body masses of the two types of bird with their corresponding equilibrium payoffs we indeed obtain that the payoff of a marked bird is smaller than the payoff of an unmarked one.

5 Concluding comments

The contribution of this study can be summarized as follows: We introduce a variation of the hawk-dove game in which there is a population formed by two types of individual who do not perceive their own type but do recognize the type of their opponent. Although the difference between the two types is "artificial" it is not innocuous. Our game has two evolutionarily stable strategies in which the probability of being aggressive toward one type of individual is always higher than the probability of being aggressive toward the other type. It is worth stressing that the type of individual treated worse may not be the minority group. This contradicts the intuition according to which the type which constitutes the minority of the population seems likely to be discriminated against. The probability of aggression toward aliens does however decrease with the proportion of aliens. Increasing the proportion of aliens also decreases the probability of aggressive behavior toward locals. For a random individual the effect is positive for small proportion of aliens, and negative for large proportions.

Finally, we would like to point out that although our research was inspired by a biological experiment, the approach might also serve to explain other social situations. In particular, it should be emphasized that our results are similar to those obtained in the paper by Axtell *et al.* (1991) in which an artificial division of a group of individuals into two subgroups generates real discrimination.

References

[1] Axtell R. L., Epstein J. M. and H. P. Young (2001) The Emergence of Classes in a Multi-Agent Bargaining Model. Social Dynamics edited by S. N. Durlauf and H. P. Young.

[2] Binmore K. and Samuelson L. (2001a), Evolution and Mixed Strategies. *Games and Economic Behavior* **34**, 200–226

[3] Binmore K. and Samuelson L. (2001b), Can Mixed Strategies be Stable in Asymmetric Games. *Journal of Theoretical Biology* **210**, 1–14

[4] Dennis, R. L., R. C. Newberry, H. W. Cheng, and I. Estevez, (2008), Appearance Matters: Artificial Marking Alters Aggression and Stress, *Poultry Science* **87**, 1939–1946.

[5] Maynard Smith J., (1982), *Evolution and the theory of games,* Cambridge University Press; Cambridge, UK.

[6] Maynard Smith J., and Price G. R, (1973) The Logic of Animal Conflict, *Nature* **246**, 15–18

[7] Maynard Smith J. and Parker G A. (1976), The Logic of Asymmetric Contests. *Animal Behavior* **24**, 159–75.

[8] Nash, J. F., (1951), Non-Cooperative Games, *Annals of Mathematics* **54**, 286–295.

[9] Osborne, M. J., (2004), *An Introduction to Game Theory,* Oxford University Press; New York, USA.

[10] Selten, R., (1980), A note on Evolutionarily Stable Strategies in Asymmetric Animal Contests, *Journal of Theoretical Biology* 84, 93–101.

[11] Weibull, J. W., (1995), *Evolutionary Game Theory,* MIT Press; Cambridge, Massachusetts, USA.

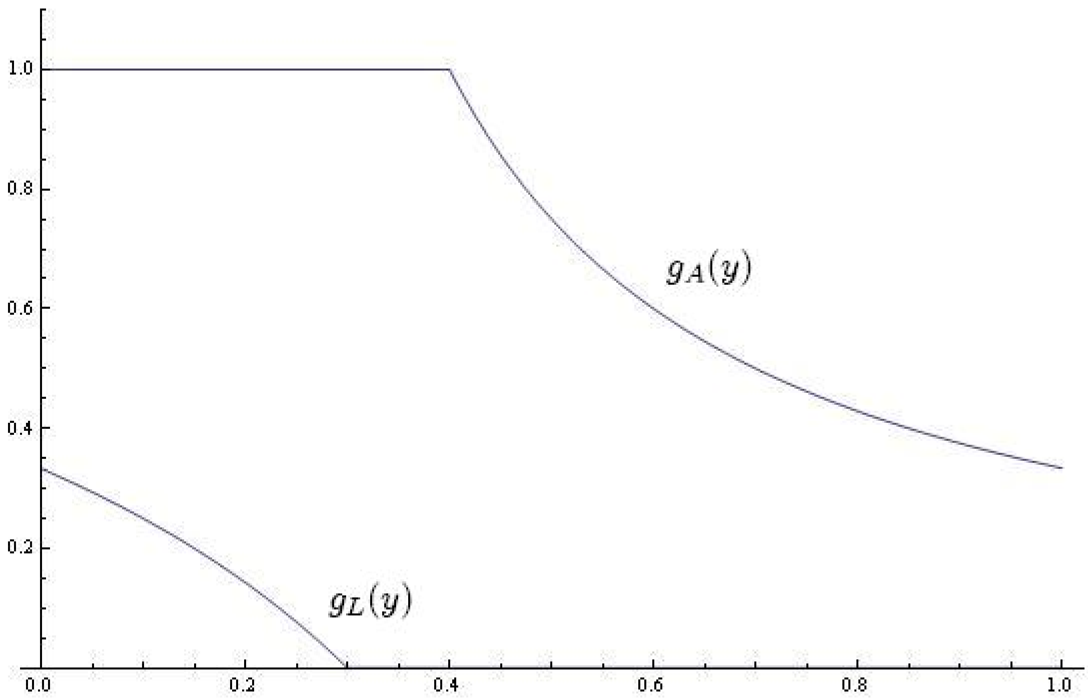

Figure 1: Probabilities of a hawk behavior toward local (g_L) and toward aliens (g_A) as a function of the proportion of aliens (y)

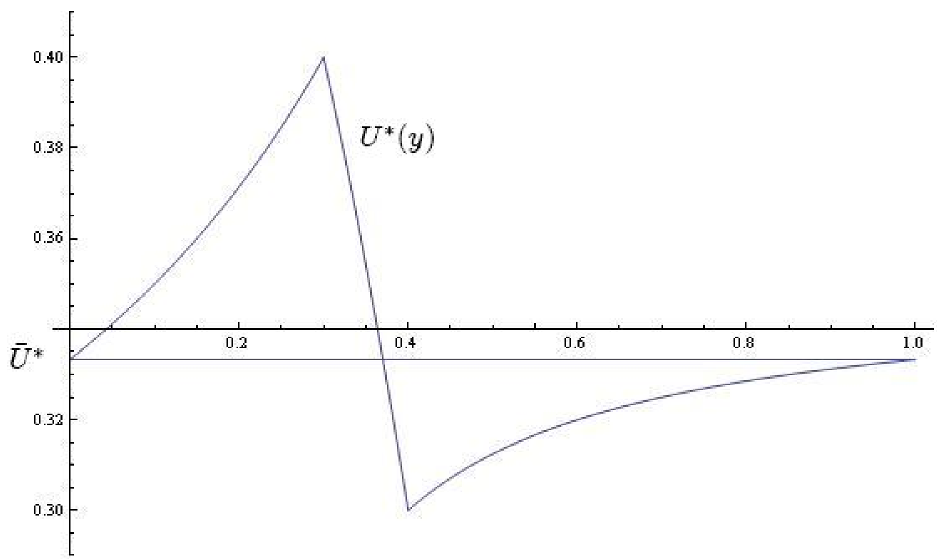

Figure 2: Payoff in the homogeneous game ($\bar{U}^*$) and payoff of a random individual in the heterogeneous game ($U^*(y)$) as a function of the proportion of aliens (y)

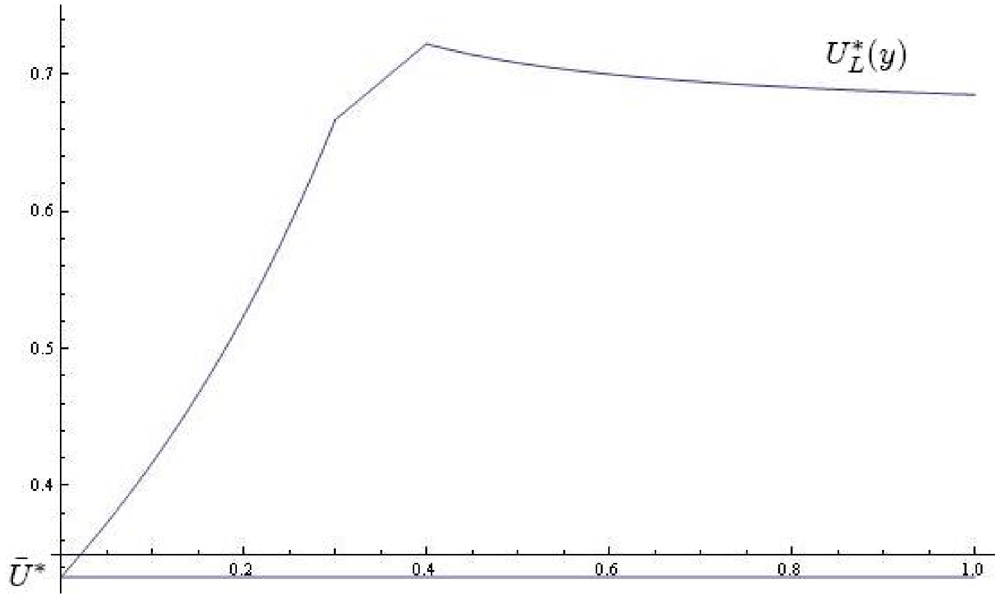

Figure 3: Payoff in the homogeneous game ($\bar{U}^*$) and payoff of a local in the heterogeneous game ($U_L^*(y)$) as a function of the proportion of aliens (y)

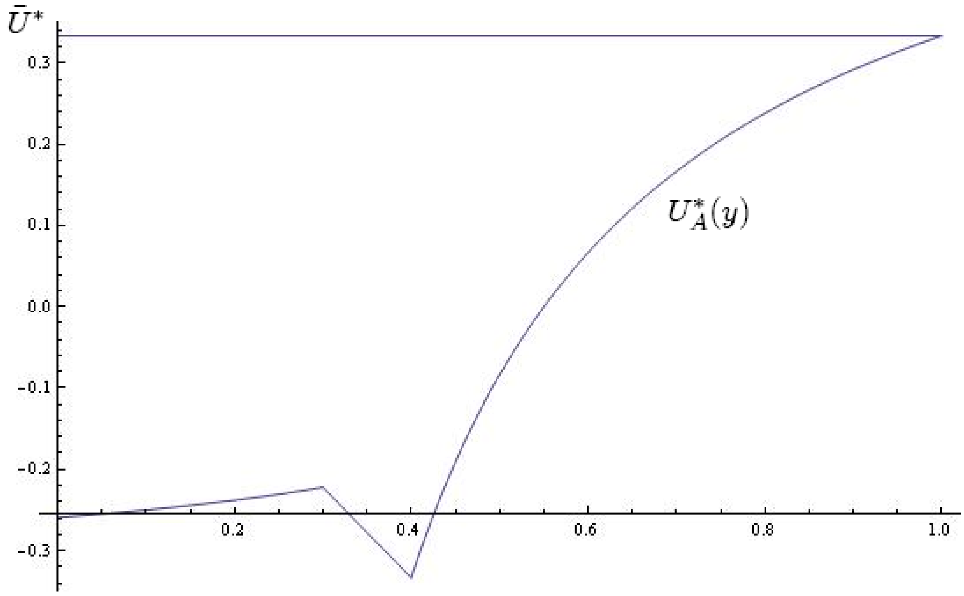

Figure 4: Payoff in the homogeneous game ($\bar{U}^*$) and payoff of an alien in the heterogeneous game ($U_A^*(y)$) as a function of the proportion of aliens (y)