Dominance, Sharing, and Assessment in an Iterated Hawk–Dove Game

Cameron L. Hall[a], Mason A. Porter[b], Marian S. Dawkins[c]

[a]*Department of Engineering Mathematics, University of Bristol, UK*
Department of Mathematics and Statistics, University of Limerick, Ireland
Mathematical Institute, University of Oxford, UK
[b]*Department of Mathematics, University of California Los Angeles, USA*
Mathematical Institute, University of Oxford, UK
[c]*Department of Zoology, University of Oxford, UK*

Abstract

Animals use a wide variety of strategies to reduce or avoid aggression in conflicts over resources. These strategies range from sharing resources without outward signs of conflict to the development of dominance hierarchies, in which initial fighting is followed by the submission of subordinates. Although models have been developed to analyze specific strategies for resolving conflicts over resources, little work has focused on trying to understand why particular strategies are more likely to arise in certain situations. In this paper, we use a model based on an iterated Hawk–Dove game to analyze how resource holding potentials (RHPs) and other factors affect whether sharing, dominance relationships, or other behaviors are evolutionarily stable. We find through extensive numerical simulations that sharing is stable only when the cost of fighting is low and the animals in a contest have similar RHPs, whereas dominance relationships are stable in most other situations. We also explore what happens when animals are unable to assess each other's RHPs without fighting, and we compare a range of strategies for

this problem using simulations. We find (1) that the most successful strategies involve a limited period of assessment followed by a stable relationship in which fights are avoided and (2) that the duration of assessment depends both on the costliness of fighting and on the difference between the animals' RHPs. Along with our direct work on modeling and simulations, we develop extensive software to facilitate further testing; it is available at `https://bitbucket.org/CameronLHall/dominancesharingassessmentmatlab/`.

1. Introduction

When animals are in conflict over resources, the result can be anything from overt fighting (Archer, 1988; Huntingford and Turner, 1987) to settling disputes by signaling (Clutton-Brock and Albon, 1979), formation of dominance hierarchies (Bonabeau, 1999; Beacham and Newman, 1987; Drummond, 2006; Guhl, 1968; O'Donnell, 1998), or even resource sharing without any sign of overt conflict (Wilkinson, 1984). To deal with this diversity of outcomes, a wide variety of game-theoretic models have been developed, including models that involve conditional strategies with assessment (Parker, 1974; Parker and Rubenstein, 1981; Enquist et al., 1990; Enquist and Leimar, 1983; Payne and Pagel, 1996); iterated games, in which animals repeatedly encounter the same individuals (Axelrod and Hamilton, 1981); models based on simple learning rules (Fawcett and Johnstone, 2010; Grewal et al., 2013); and models that include winner and loser effects (where winners are likely to keep winning and losers are likely to keep losing) (Goessmann et al., 2000;

2

Hsu et al., 2005; Kura et al., 2015, 2016; Mesterton-Gibbons and Sherratt, 2016).

This large variety of models often makes it difficult for biologists without detailed mathematical knowledge to understand the differences and similarities between these models. In general, models in which resources are ultimately divided unequally between 'winners' and 'losers' (Eshel and Sansone, 2001; Eshel, 2005; Fawcett and Johnstone, 2010; Hammerstein, 1981; Houston and McNamara, 1991; Kura et al., 2015, 2016; Mesterton-Gibbons et al., 2014, 2016) are often based on the Hawk–Dove framework that was described in Maynard Smith (1979), whereas models that concentrate on the evolution of sharing and other apparently paradoxical acts of cooperation (Baek et al., 2017; Carvalho et al., 2016; Doebeli and Hauert, 2005; Nowak and Sigmund, 1993a,b; Nowak, 2012; Trivers, 2006) are often based on the Iterated Prisoner's Dilemma (IPD) framework that was described in Axelrod (1984). This distinction seems to suggest that different models are needed for different outcomes. Our aim in this paper is to show that, by contrast, a single model — a modified version of the Hawk–Dove model that was developed by Maynard Smith (1982) — can predict the observed diversity of possible outcomes of animal conflict, including overt fighting, resource-sharing, dominance relationships, and other social structures. By making a small number of biologically-realistic modifications to the original Hawk–Dove model, we show that it is possible to derive many of the currently employed game-theoretic models of animal conflict and cooperation (including Conditional Hawk–Dove, Prisoner's Dilemma, and Snowdrift) and to explain the widespread occurrence of dominance hierarchies in animal societies.

3

Our model, and the accompanying software, makes it easy to simulate many different games and to explore the effects of different assumptions and parameter values. With the present work, we hope to facilitate communication between mathematical modelers and field biologists, leading to a better understanding of why different animals resolve their conflicts in such different ways.

Our paper proceeds as follows. We present our model in Section 2. We demonstrate how it shares several features with existing animal interaction models (Eshel and Sansone, 2001; Eshel, 2005; Mesterton-Gibbons et al., 2014) and that, in certain limits, it reduces to a form of the IPD that was described in Axelrod (1984). Our model addresses how differences in resource holding potential (RHP) can affect animals' optimal strategies while also providing a framework for modeling progressive assessment of RHPs by enabling animals' behaviors to change as they learn information about RHP through experience (Maynard Smith and Parker, 1976; Parker and Rubenstein, 1981; Enquist and Leimar, 1983). In Section 3, we investigate evolutionarily stable strategies (ESSs) for cases in which animals begin with complete knowledge of their RHP relative to those of their opponents. This forms the basis for Section 4, where we describe and analyze various strategies by which animals can use fights to learn about their RHP relative to those of their opponents. We conclude and discuss the implications of our results in Section 5. Our software is available is at https://bitbucket.org/CameronLHall/dominancesharingassessmentmatlab/, and we encourage readers to pursue our ideas further and extend them.

4

2. Model development

2.1. An iterated Hawk–Dove game with winners and losers

Consider two animals, A and B, who interact with each other repeatedly in a contest for resources. We construct a model that is based on the classical Hawk–Dove (HD) game (see Figure 1), which was developed to describe a single conflict between animals over a shareable resource (Maynard Smith and Price, 1973; Maynard Smith, 1979, 1982), which we normalize to have a value of 1. In the HD game, each animal has a choice between a Hawk strategy (in which it escalates the conflict) and a Dove strategy (in which it retreats from an escalated conflict). If two Doves encounter each other, they share a resource equally. If a Hawk and a Dove encounter each other, the Hawk takes the entire resource. If two Hawks encounter each other, a fight ensues. Assuming that each animal has an equal chance of winning, each animal receives a payoff of $(1 - c)/2$, where c represents the cost of fighting.

There are various systems for classifying 2×2 games according to the orders of the payoffs to the players (Robinson and Goforth, 2005; Bruns, 2015). For the classical HD game, the classification depends on the value of c. When $c > 1$, fights cost more than the value of the resource, so the classical HD game is an example of the 'Snowdrift' game (also called the 'Chicken' game[1]). When $c \in (0, 1)$, the HD game is an example of a 'Prisoner's Dilemma' (PD). When $c = 1$, the classical HD game is not a strict

---

[1]Although the term 'Chicken' is used more widely than 'Snowdrift' (Robinson and Goforth, 2005; Bruns, 2015), we use the latter to avoid any confusion between game-classification terminology and actual chickens.

	Animal B	
	Dove	Hawk
Dove	$\left(\frac{1}{2}, \frac{1}{2}\right)$	$(0, 1)$
Hawk	$(1, 0)$	$\left(\frac{1-c}{2}, \frac{1-c}{2}\right)$

Animal A

Figure 1: Payoffs to animals A and B in a classical Hawk–Dove game where the value of the resource is 1 and the cost of fighting is c.

ordinal game, as there are multiple ways for a player to obtain a payoff of 0. Such 'games with ties' (Bruns, 2015) are more complicated to analyze than strict ordinal games. They require model parameters to take specific values, so they are 'non-generic' (Broom and Rychtář, 2013) and are thus unlikely to occur in practice. Throughout this paper, we concentrate on ordinal games.

The classical HD game and the IPD game are both inadequate for describing social interactions between animals. One key issue is that both of these games assume that there are no differences in the payoff matrices for the two animals. Without modification, neither game takes into account that one animal may have a larger RHP than the other, or that one animal may place a higher value on a disputed resource. Another issue is that neither game includes any element of assessment, where an animal uses information about its opponent from signals, resource ownership, or past experience to guide its behavior. This is an important issue, as it is well-established that conditional strategies with assessment are far better models of animal conflict than strategies that do not involve assessment (Parker, 1974).

6

In our model, we make three important modifications to the classical HD game:

1. We study an iterated game in which each stage is an HD game.
2. We assume that there is always an explicit winner and loser in any Hawk–Hawk interaction, rather than supposing that each animal obtains an identical payoff.
3. We assume that Hawk–Hawk conflicts have a biased outcome, in that one animal is more likely than its opponent to win the fight.

In Figure 2, we show our modified HD game in normal form.

Our first modification is to consider an iterated game of HD stage games to model repeated interactions of two animals from the same social group. We assume that each HD interaction is identical, such that animals do not get stronger or weaker over time, and that the available resource has the same value in each stage game.

In our second modification, we assume that each Hawk–Hawk interaction has a winner and a loser, instead of each animal receiving an identical payoff. Specifically, we assume that there is a cost $c_W > 0$ of winning a fight and a cost $c_L > 0$ of losing a fight, such that the winner obtains a payoff of $1 - c_W$ and the loser obtains a payoff of $-c_L$. Additionally, $1 - c_W > -c_L$, as the payoff from winning a fight is greater than the payoff from losing it.

In our third modification, we suppose that one animal is more likely to win a Hawk–Hawk fight than its opponent. Specifically, we assume a fixed probability p_A that animal A beats animal B. Consequently, the expected payoff for animal A from a Hawk–Hawk fight is $p_A(1 - c_W + c_L) - c_L$, and the expected payoff for animal B is $1 - c_W - p_A(1 - c_W + c_L)$.

	Animal B	
	Dove	Hawk
Dove	$(\frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, \frac{1}{2})$
Hawk	$(1, 0)$	$(1 - c_W, -c_L)$ or $(-c_L, 1 - c_W)$

Animal A

Figure 2: Payoffs to animals A and B in one stage game of our modified HD model. The value of the resource is 1. When both animals play Hawk, there is a fight that is resolved probabilistically, such that the chance of animal A winning the fight is p_A. The winner obtains a payoff of $1 - c_W$ (the value of the resource minus its cost of fighting), and the loser obtains $-c_L$ (a fixed penalty for losing).

Our modified HD model has three dimensionless parameters: c_W, c_L, and p_A. We assume that c_W and c_L are fixed for a given species and that both animals A and B know the values of c_W and c_L. By contrast, p_A depends on the different fighting abilities (as measured by RHP) of animals A and B.

Let R_A and R_B, respectively, denote the RHPs of animals A and B. The probability that A wins a fight is $p_A = \varphi(R_A, R_B)$, where $\varphi(x, y)$ is some function that satisifes $0 \leq \varphi \leq 1$. We require that $\varphi_x(x, y) > 0$, as an increased RHP implies a higher probability of winning a fight, and that $\varphi(x, y) = 1 - \varphi(y, x)$, so the chance of winning a fight is independent of whether we label an animal as A or B. When RHPs are given by positive real numbers, one suitable choice is

$$\varphi(R_A, R_B) = \frac{R_A}{R_A + R_B}, \tag{1}$$

which enables further simplifications in certain scenarios. If we draw the

RHPs from an exponential distribution, the corresponding p_A comes from a uniform distribution. More generally, if we draw the RHPs from a gamma distribution with shape parameter k (and any rate parameter), $p_A \sim \text{Beta}(k, k)$. In this paper, we draw p_A from such a beta distribution, and we assume that animals A and B have prior knowledge of this distribution, even if they do not have specific knowledge of p_A.

2.2. Terminology and classification of strategies

We are concerned with comparing different strategies for playing the iterated HD (IHD) game that we defined in Section 2.1.

Each interaction of animals A and B (where they play the 'stage game' in Figure 2) is a 'round'. In a round, each animal chooses a 'move' of either Hawk or Dove. As in Houston and McNamara (1991) but unlike Mesterton-Gibbons et al. (2014), Hawk and Dove are the only possible moves in a round. Our model does not include any concept of resource ownership, and we do not consider Bourgeois or anti-Bourgeois strategies. If both animals choose Hawk as their move in a round, we say that a 'fight' occurs.

The overall game, which consists of a large number of rounds, is a 'contest' between animals A and B. We evaluate the 'total payoff' to each animal by summing the discounted payoffs to animals A and B over all rounds (Fujiwara-Greve, 2015). Taking $\gamma \in (0, 1)$ to be the discount rate and $\rho_k^{(A)}$ to be the payoff to animal A in round j, the total payoff to A at the end of a contest is

$$\varphi_A = \sum_{j=1}^{\infty} \gamma^{j-1} \rho_j^{(A)}. \tag{2}$$

An animal's 'strategy' is a set of rules that enable it to determine the move

that it plays in each round. A strategy specifies a move for the first round and a rule, based on the outcomes of the previous rounds, for determining which move to select in each subsequent round. Strategies can be probabilistic, with a probability between 0 and 1 of playing Dove in a particular round. Animals A and B may use different strategies; for simplicity, we describe strategies from the perspective of animal A.

We distinguish between three types of strategies: 'simple', 'informed', and 'learning'. When animal A pursues a simple strategy, it does not use knowledge of p_A to inform its behavior. When animal A pursues an informed strategy, it begins a contest with perfect knowledge of p_A and uses this information to guide its behavior. Such a strategy is relevant only when animal A is able both to assess its opponent perfectly without any fighting and to exploit that assessment in its choices. When animal A pursues a learning strategy, it begins with limited information about p_A, but it uses information from fights to update its beliefs about p_A and uses its beliefs about p_A to guide its choices. In Table 1, we summarize the three strategy types.

Strategy type	Uses information about p_A to inform behavior	Available to animals with perfect knowledge of p_A at start of contest	Available to animals with no knowledge of p_A at start of contest
Simple	No	Yes	Yes
Informed	Yes	Yes	No
Learning	Yes	No	Yes

Table 1: A summary of the three different strategy types. Because animals that pursue simple strategies do not use p_A to inform their behavior, simple strategies are always available to animals, even if they begin a contest without knowledge of p_A.

In our analysis of simple strategies, we consider only 'memory-1 strategies' (Press and Dyson, 2012; Nowak and Sigmund, 1993b), in which the probability of playing Dove in a round depends only on the moves of the two animals in the previous round. As we discuss in more detail in Section 3.2, one can represent simple memory-1 strategies using $\mathbf{s}$, a vector of probabilities (Nowak and Sigmund, 1993b; Press and Dyson, 2012). Our focus on memory-1 strategies places some restrictions on the variety of simple strategies that we can consider, but many important strategies — including Tit for Tat (Axelrod, 1984), Grim Trigger (Friedman, 1971) (which we call 'Grim'), Pavlov (Nowak and Sigmund, 1993b), and extortionate strategies (Press and Dyson, 2012) — are examples of memory-1 strategies.

Our analysis of informed and learning strategies also focuses on memory-1 strategies. In an informed memory-1 strategy, the probability of playing Dove in a round depends both on p_A and on the moves from the previous round. That is, an informed memory-1 strategy is one in which an animal uses knowledge of p_A to choose an element from a set of memory-1 'substrategies', each of which can be represented in the standard vector format of a simple strategy (see Section 3.2). For a learning memory-1 strategy, the probability of playing Dove in a round depends both on an animal's current beliefs about p_A and on the moves from the previous round. That is, the results of all previous fights inform an animal's beliefs about p_A, and it uses these beliefs to determine the substrategy that it employs in a round

An animal needs a learning strategy only if it does not begin with perfect information about p_A. An animal with knowledge of p_A is able to use a simple strategy or an informed strategy, whereas an animal without this knowledge

11

can use either a simple strategy or a learning strategy. Determining optimal strategies for animals with knowledge of p_A is the focus of Section 3, and determining optimal strategies for animals without knowledge of p_A is the focus of Section 4.

Even with our consideration of only memory-1 strategies, the set of possible strategies for this game is infinite. Additionally, there are well-established difficulties with assessing the quality of strategies for iterated games. For discount rates that are sufficiently close to 1, there are 'folk theorems' that imply that there is a very large class of strategies that are subgame-perfect Nash equilibria (Friedman, 1971; Fujiwara-Greve, 2015).

Our approach for assessing the success of a strategy is analogous to the methods that were proposed in Maynard Smith and Price (1973) to compare strategies for a single multi-stage conflict between two animals. We consider a limited set of plausible strategies and use simulations to evaluate the outcomes of contests for every possible pair of strategies. Following Axelrod (1984), we refer to the set of all contests between strategies as a 'tournament'. As in Maynard Smith and Price (1973), our contests involve probabilistic elements; to account for this, we run multiple simulations for each pair of strategies and report the mean payoff to each animal. We then present these results in a matrix, from which we evaluate which strategies are evolutionarily stable.

An 'evolutionarily stable strategy' (ESS) is based on the following idea: if all members of a population adopt an ESS, a mutant strategy cannot successfully invade the population (Maynard Smith, 1982). In mathematical terms, let $E(T, S)$ represent the expected payoff to animal A when it pursues

12

strategy T and animal B pursues strategy S. A strategy S is an ESS if $E(T, S) \leq E(S, S)$ for all strategies $T \neq S$, with the additional condition that $E(S, T) > E(T, T)$ whenever $E(T, S) = E(S, S)$. Not all games have ESSs, some games have multiple ESSs, and sometimes probabilistic combinations of strategies can be ESSs; these are called 'mixed ESSs'. (Maynard Smith and Price, 1973; Fujiwara-Greve, 2015). In many cases, it is possible to use the matrix of payoffs from a tournament to compute the set of all possible ESSs (Haigh, 1975; Broom and Rychtář, 2013).

2.3. A Bayesian approach to learning

An animal that pursues a learning strategy begins a contest with complete knowledge of c_W and c_L, but with no information about p_A beyond the fact that $p_A \sim \text{Beta}(k, k)$ for a specified k. The animal uses information from its interactions with its opponent to improve its estimate of p_A. Various methods have been developed for implementing learning in repeated games with incomplete information (Aumann and Maschler, 1995; Sastry et al., 1994; Watkins and Dayan, 1992; Jordan, 1995). Many of these approaches are very general and sophisticated, but they depend on the assumption that each player is rational and can therefore exploit the assumed rationality of their opponent to obtain information.

In the present work, by contrast, we use a Bayesian approach to incorporate the information obtained from fights into an animal's beliefs about p_A, and we assume that its beliefs about p_A do not change if there is no fight. We assume that p_A is constant in time, so we can treat fights as independent Bernoulli trials with a fixed probability of success. Using Bayes' rule, we update estimates of p_A as an animal acquires information from its

13

fights (Gelman et al., 2004). The beta distribution is a conjugate prior of the Bernoulli distribution, so if one can represent an animal's initial beliefs about its probability of winning a fight using a beta distribution, then these beliefs remain a beta distribution (but with altered parameter values) as it obtains information from its fights (Gelman et al., 2004; Trimmer et al., 2011).

There has been some discussion in the biological literature about whether animals (including humans) are capable of Bayesian updating (Trimmer et al., 2011; Valone, 2006; McNamara et al., 2006). The consensus of such work is that observations of vertebrate behavior (especially with foraging) are consistent with Bayesian updating. The main challenges of taking a Bayesian approach are determining the parameters for the prior distribution and deciding how animals should use their current estimate of the distribution of p_A to inform their behavior. In Section 4, we propose and compare a variety of learning strategies that uninformed animals can use to assess their probability of winning a fight, and we discuss how this information influences which strategy they adopt.

3. Analysis of strategies for animals with knowledge of p_A

When cost-free or low-cost observations are reliable indicators of RHPs, it is reasonable to assume that all animals begin contests with knowledge of p_A. In this section, we seek ESSs among the strategies that are available to an animal with knowledge of p_A. Finding optimal informed strategies will also guide the development of plausible learning strategies in Section 4.

	Animal B	
	Dove	Hawk
Dove	$(\frac{1}{2}, \frac{1}{2})$	$(\frac{1}{2}, \frac{1}{2})$
Hawk	$(1, 0)$	(μ_A, μ_B)

Figure 3: Expected values of the payoffs to animals A and B in one round of our modified HD game. We give the expected payoffs from fighting, μ_A and μ_B in equations (3) and (4). The character of the game (and the corresponding best strategies) depend on the values of μ_A and μ_B.

3.1. Stage-game classification and its implications

Suppose that two animals are playing against each other in the modified HD game of Figure 2 and that they both know p_A. Replacing the random outcome of a fight with the expected value of the interaction, we obtain the game in Figure 3, where the expected payoffs μ_A and μ_B to animals A and B when they both play Hawk are

$$\mu_A = -c_L + p_A(1 - c_W + c_L), \tag{3}$$

$$\mu_B = 1 - c_W - p_A(1 - c_W + c_L). \tag{4}$$

The optimal strategies for informed animals A and B in an iterated game depend on the ordering of the payoffs of the different interactions in the stage game; these, in turn, depend on the values of μ_A and μ_B. The stage-game payoffs associated with Dove–Dove and Hawk–Dove plays by the animals are $(\frac{1}{2}, \frac{1}{2})$ and $(1, 0)$, respectively. We can thus classify the different stage games according to the values of μ_A and μ_B relative to each other and to 0, $\frac{1}{2}$,

15

and 1. Because $\mu_A < 1$, $\mu_B < 1$, and $\mu_A + \mu_B < 1$, there are 10 possible orderings of the payoffs. Assuming without loss of generality that $\mu_A > \mu_B$, we enumerate the 5 remaining orderings in Table 2. We obtain the other 5 orderings by swapping μ_A and μ_B.

As one can see in Table 2, we name the classifications of the stage game based on the ordering of payoffs to animals A and B. For example, if animal A is faced with a 'Snowdrift' situation and animal B is faced with a 'Deadlock' situation, we have a 'Snowdrift–Deadlock' game.

All possible stage games other than the Snowdrift game have a unique Nash equilibrium. In the Snowdrift game, for which μ_B and μ_A are both negative, there are three Nash equilibria: (Hawk, Dove), (Dove, Hawk), and a mixed Nash equilibrium in which animal A plays Dove with probability $2|\mu_A|/(1+2|\mu_A|)$ and animal B plays Dove with probability $2|\mu_B|/(1+2|\mu_B|)$.

Ordering	Name	Nash Equilibria
$\mu_B < \mu_A < 0 < \frac{1}{2} < 1$	Snowdrift (A beats B)	(D,H), (H,D), or mixed
$\mu_B < 0 < \mu_A < \frac{1}{2} < 1$	PD–Snowdrift	(H,D)
$\mu_B < 0 < \frac{1}{2} < \mu_A < 1$	Deadlock–Snowdrift	(H,D)
$0 < \mu_B < \mu_A < \frac{1}{2} < 1$	PD (A beats B)	(H,H)
$0 < \mu_B < \frac{1}{2} < \mu_A < 1$	Deadlock–PD	(H,H)

Table 2: Classification of the stage game according to the values of μ_A and μ_B, with $\mu_A > \mu_B$. We base the nomenclature on the ordering of payoffs to the animals. For example, a 'PD–Snowdrift' game is one in which the ordering of payoffs to animal A is identical to that in the classical Prisoner's Dilemma game, whereas the ordering of payoffs to animal B is identical to that in the classical Snowdrift game.

The different classifications of the stage game correspond to different

regions in (μ_A, μ_B)-space (see Figure 4). For fixed c_W and c_L, changing p_A leads to changes in μ_A and μ_B, which in turn correspond to changes in the classification of the stage game. A fixed choice of c_W and c_L defines a unique line segment in Figure 4 that connects $(-c_L, 1 - c_W)$ to $(1 - c_W, -c_L)$. The center of each line segment occurs when $\mu_A = \mu_B = \frac{1}{2}(1 - c_W - c_L)$, for which $p_A = \frac{1}{2}$. The blue lines in Figure 4 show three examples for different values of the parameters c_W and c_L. As one increases p_A, one moves along such a line in Figure 4 from the top left to the bottom right.

As we can see from the example line segments in Figure 4, the sequence of stage games as p_A increases from 0 to 1 depends on the values of c_W and c_L. Considering all allowable possibilities in (c_W, c_L)-space (see Section 2.1), there are six different parameter regimes, which are associated with different sequences of stage-game classifications. We show the parameter regimes in Figure 5, and we detail the associated sequences of stage-game classifications in Table 3.

Parameter regimes IV, V, and VI encompass a narrower range of stage games than parameter regimes I, II, and III. For example, we see from Table 3 that the stage-game classifications that occur in parameter regime IV as we change p_A from 0 to $\frac{1}{2}$ are a subset of those that occur in parameter regime II for the same values of p_A. For example, the transition from Snowdrift–Deadlock to Snowdrift–PD at $p_A = (\frac{1}{2} - c_W)/q$ does not occur in parameter regime IV, as $(\frac{1}{2} - c_W)/q < 0$ in this case. Similarly, the stage-game classifications that occur for $p_A \in [0, 1/2]$ in parameter regime V are a subset of those that occur for $p_A \in [0, 1/2]$ in parameter regime III; and the stage-game classifications for $p_A \in [0, 1/2]$ in parameter regime VI are a subset of those
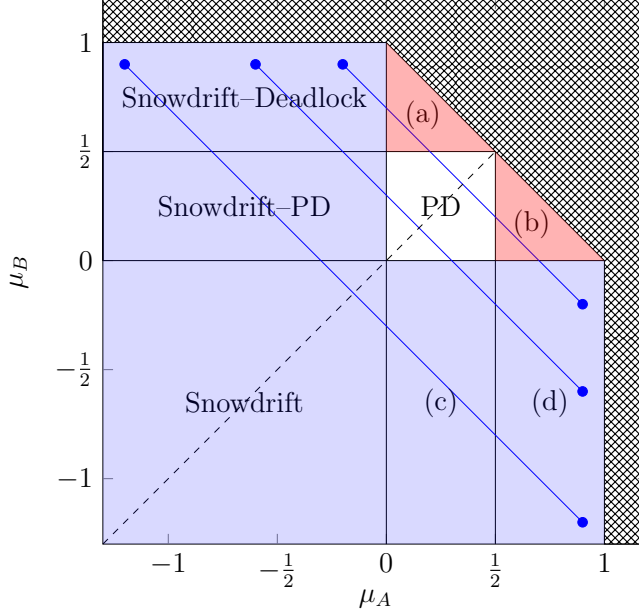
Figure 4: The classification of the stage game depends on μ_A and μ_B, which are the expected payoffs of fighting for animals A and B, respectively. We mark the case with $p_A = 0.5$ (in which animals A and B are evenly matched) with a dashed line, and we separate the different stage-game classifications with solid lines. We shade the region that is inaccessible when c_L and c_W are positive. In addition to the stage-game classifications that we label in the diagram, (a) designates a PD–Deadlock game, (b) designates Deadlock–PD, (c) designates PD–Snowdrift, and (d) designates Deadlock–Snowdrift. The blue lines show the possible values of μ_A and μ_B from varying p_A for three different choices of c_W and c_L. From left to right, the blue lines correspond to the parameter pairs $(c_W, c_L) = (0.1, 1.2)$, $(c_W, c_L) = (0.1, 0.6)$, and $(c_W, c_L) = (0.1, 0.2)$. In Section 3.4, we discuss the behavior of two informed animals that both pursue the optimal strategy from Section 3.3. This behavior depends on the stage-game classification; we illustrate this dependence by coloring regions of the figure. We use pink when informed animals who pursue an optimal strategy play Hawk against each other in all but the first round, white when they play Dove against each other in every round, and light blue when the animal with a larger RHP plays Hawk in every round and the animal with a smaller RHP plays Dove in every round.

for $p_A \in [0, 1/2]$ in parameter regime V (and therefore also a subset of those for $p_A \in [0, 1/2]$ in parameter regime III).

To focus on cases in which optimal strategy can vary significantly as p_A changes, we concentrate our analysis on parameter regimes I, II, and III.
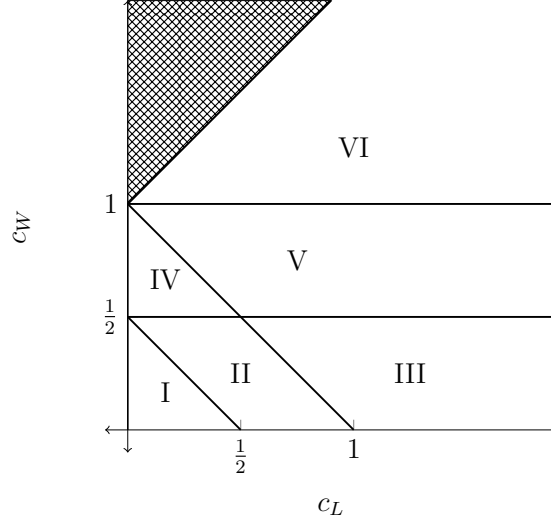

Figure 5: As p_A changes, the classification of the stage game changes. We show the different regimes in (c_W, c_L)-space that lead to different sequences of stage-game classifications. We shade the region that is inaccessible due to the requirement that $1 - c_W > -c_L$. We outline the stage games in Table 3, and one can also see them in Figure 4.

3.2. Defining informed strategies

We consider seven informed strategies: 'Bully', 'Nash', 'Mixed Nash', 'Selfish', 'Snowdrift TfT', 'PD TfT', and 'Informed TfT'. In Table 4, we describe the substrategies that are pursued by these informed strategies for different stage games. Supposing that animal A is the one that is playing the strategy and that animal B is its opponent, we summarize the seven informed strategies as follows:

19

- 'Bully' always plays Hawk if $p_A > \frac{1}{2}$ (i.e., when it is more likely than not to beat its opponent) and always plays Dove if $p_A < \frac{1}{2}$.

- 'Nash' pursues a Nash-equilibrium substrategy for each move. In a Snowdrift game, 'Nash' uses the difference in RHP to decide which pure Nash equilibrium to pursue; it always plays Hawk if $p_A > \frac{1}{2}$ and always plays Dove if $p_A < \frac{1}{2}$.

- 'Mixed Nash' pursues a Nash-equilibrium substrategy for each move. In a Snowdrift game, 'Mixed Nash' pursues a mixed substrategy in which its probability of playing Dove is $2|\mu_A|/(1 + 2|\mu_A|)$.

- 'Selfish' chooses a substrategy based on its expected payoff from a fight. It always plays Hawk if $\mu_A > \frac{1}{2}$, pursues a Tit for Tat substrategy if $0 < \mu_A < \frac{1}{2}$, and always plays Dove if $\mu_A < 0$.

- 'Snowdrift TfT' pursues a Tit for Tat substrategy when the stage game is Snowdrift or Snowdrift–PD, but it otherwise follows the Nash equilibrium.

- 'PD TfT' pursues a Tit for Tat substrategy when the stage game is PD or PD–Deadlock, but it otherwise follows the Nash equilibrium.

- 'Informed TfT' pursues a Tit for Tat substrategy, except when this is clearly unsuitable. Specifically, it always plays Hawk in a Deadlock–Snowdrift or Deadlock–PD stage game and always plays Dove in a Snowdrift–Deadlock stage game, but it otherwise pursues Tit for Tat.

We do not specify substrategies for the non-generic games that occur when p_A is at a 'critical' value at a boundary between stage-game classifications

20

(see Table 3), because the probability that this occurs in a simulation is vanishingly small. If such a critical situation occurs, we assume that each animal chooses the most aggressive of its available strategies. That is, each animal chooses the strategy that it would pursue if its RHP were larger by an arbitrarily small amount for a fixed value of its opponent's RHP. Our software also returns a warning in such a critical case; no such warning occurred in any of our tournaments.

We now discuss our motivation behind the above strategies. One approach that animals can take is to pursue a Nash-equilibrium strategy of the stage game in each round. This involves a small complication for the Snowdrift game (where there are three Nash equilibria), but otherwise each animal has a unique strategy in each round. Pursuing the stage-game Nash equilibrium in each round of an iterated game is not necessarily optimal. See Axelrod and Hamilton (1981) for the most famous example, which arises in the IPD. Nevertheless, cooperative strategies, such as Tit for Tat and Grim (in which each animal plays Dove until the opposing animal has played Hawk), can still be very successful and are evolutionarily stable (except to invasion from other cooperative strategies) for sufficiently large values of the discount parameter γ (Axelrod and Hamilton, 1981; Fujiwara-Greve, 2015).

We examine informed strategies in which animals frequently follow a Nash-equilibrium strategy, but they may deviate from this for certain stage-game classifications (for example, by playing Tit for Tat when faced with a PD stage game). Not all deviations from a Nash-equilibrium strategy are appropriate, as a Nash-equilibrium strategy is clearly optimal in certain stage-game classifications. For example, if animal A is playing a Deadlock–

21

PD or Deadlock–Snowdrift stage game, its best strategy is to play Hawk, regardless of whether animal B plays Hawk or Dove. Consequently, all of our informed strategies pursue an Always Hawk substrategy when faced with a Deadlock–PD or Deadlock–Snowdrift stage game. Similarly, all of our informed strategies pursue an Always Dove substrategy when faced with a Snowdrift–Deadlock stage game.

For other stage-game classifications, the choice of optimal strategy is less obvious. For simplicity, we assume for each type of stage game that an informed animal pursues one of four different substrategies: (1) always play Hawk, (2) always play Dove, (3) play Tit for Tat, or (4) play a mixed Hawk–Dove strategy. For a given type of stage game, only a strict subset of these strategies is a reasonable choice for a rational animal. For example, the mixed Hawk–Dove strategy can be a Nash equilibrium only for a Snowdrift game, and Always Dove is not a rational option (compared to Always Hawk or Tit for Tat) for an animal in a PD–Snowdrift game, because the lack of punishment (in the sense described by Abreu, 1988) implies that 'Always Dove' cannot be a subgame-perfect equilibrium.

3.3. Tournaments for comparing informed strategies

We use tournaments, which consist of many contests between different pairs of strategies, to compare the performance of our seven informed strategies against each other and against nine simple memory-1 strategies, where the term 'simple' is used as defined in Section 2.2: animals pursuing simple strategies do not take advantage of the information that is available from p_A. simple memory-1 strategies, where the term 'simple' signifies that they do not take advantage of the information that is available from p_A. We repre-

sent each simple memory-1 strategy using **s**, a vector of six probabilities. The first five elements give the probabilities of an animal playing Dove when the previous moves by it and its opponent are Dove–Dove, Dove–Hawk, Hawk–Dove, Hawk–Hawk (and the animal loses), and Hawk–Hawk (and the animal wins). The sixth element gives the probability that the animal plays Dove on its first move. The vector **s** is analogous to the strategy vectors in Nowak and Sigmund (1993b) and Press and Dyson (2012).

We consider the following simple memory-1 strategies: Always Dove, 25% Dove, 50% Dove, 75% Dove, Always Hawk, Tit for Tat (Axelrod and Hamilton, 1981), Grim (Friedman, 1971), Pavlov (Nowak and Sigmund, 1993b), and Modified Pavlov. In Table 5, we summarize the probabilities of playing Dove (the components of **s**) that are associated with these strategies.

An animal that uses the Pavlov strategy plays Dove with very high probability (which we take to be 0.99 in our tournaments) in response to a previous (Dove, Dove) or (Hawk, Hawk) round; otherwise, it plays Hawk. It plays Hawk with very high probability (which we again take to be 0.99 in practice) in response to a previous (Hawk, Dove) or (Dove, Hawk) round. Modified Pavlov follows the same 'win–stay, lose–shift' philosophy of Pavlov (see Nowak and Sigmund, 1993b), but instead of treating 'Hawk–Hawk' as a loss in all circumstances (as in the original Pavlov), Modified Pavlov treats 'Hawk–Hawk' as a loss only if the associated fight is a loss. We assume that an animal that plays Pavlov or Modified Pavlov is equally likely to play Hawk or Dove in its first move. In tournaments that we do not discuss in this paper, we examined versions of Pavlov and Modified Pavlov in which the first move is either definitely Hawk or definitely Dove. These other choices

23

for the first move had little effect on the overall payoffs.

Regime	Range of p_A values	Classification
I	$0 < p_A < c_L/q$	Snowdrift–Deadlock
	$c_L/q < p_A < (\frac{1}{2} - c_W)/q$	PD–Deadlock
	$(\frac{1}{2} - c_W)/q < p_A < \frac{1}{2}$	PD (B beats A)
II	$0 < p_A < (\frac{1}{2} - c_W)/q$	Snowdrift–Deadlock
	$(\frac{1}{2} - c_W)/q < p_A < c_L/q$	Snowdrift–PD
	$c_L/q < p_A < \frac{1}{2}$	PD (B beats A)
III	$0 < p_A < (\frac{1}{2} - c_W)/q$	Snowdrift–Deadlock
	$(\frac{1}{2} - c_W)/q < p_A < (1 - c_W)/q$	Snowdrift–PD
	$(1 - c_W)/q < p_A < \frac{1}{2}$	Snowdrift (B beats A)
IV	$0 < p_A < c_L/q$	Snowdrift–PD
	$c_L/q < p_A < \frac{1}{2}$	PD (B beats A)
V	$0 < p_A < (1 - c_W)/q$	Snowdrift–PD
	$(1 - c_W)/q < p_A < \frac{1}{2}$	Snowdrift (B beats A)
VI	$0 < p_A < \frac{1}{2}$	Snowdrift (B beats A)

Table 3: Changes to the classification of the stage game as p_A increases from 0 to $\frac{1}{2}$ in each of the parameter regimes from Figure 5. For convenience, we define the notation $q = 1 - c_W + c_L$. We obtain the classifications of the stage game for $p_A > \frac{1}{2}$ by symmetry. For $p_A = \frac{1}{2}$, animals A and B are evenly matched, and the game is a classical PD (in regimes I, II, and IV) or Snowdrift game (in regimes III, V, and VI), where both animals have the same expected payoff from a fight. In our numerical experiments, we concentrate on examples with $c_W < \frac{1}{2}$, so we are in one of regimes I–III.

Stage game	Bul	N	MN	Slf	SDT	PDT	InT
Deadlock–Snowdrift	H	H	H	H	H	H	H
Deadlock–PD	H	H	H	H	H	H	H
PD–Snowdrift	H	H	H	TfT	H	H	TfT
PD (A beats B)	H	H	H	TfT	H	TfT	TfT
PD (B beats A)	D	H	H	TfT	H	TfT	TfT
PD–Deadlock	D	H	H	TfT	H	TfT	TfT
Snowdrift (A beats B)	H	H	Mix	D	TfT	H	TfT
Snowdrift (B beats A)	D	D	Mix	D	TfT	D	TfT
Snowdrift–PD	D	D	D	D	TfT	D	TfT
Snowdrift–Deadlock	D	D	D	D	D	D	D

Table 4: Summary of substrategies when animal A play each of our informed strategies: Bully ('Bul'), Nash ('N'), Mixed Nash ('MN'), Selfish ('Slf'), Snowdrift TfT ('SDT'), PD TfT ('PDT'), and Informed TfT ('InT'). The entries in the table indicate the substrategy that is pursued by animal A when the stage-game classification is the one in the left column. The strategies in the table are Hawk ('H'), Dove ('D'), Tit for Tat ('TfT'), and mixed substrategies ('Mix'). See the text for details.

Strategy name	Probability of playing Dove (otherwise playing Hawk)					
	After DD	After DH	After HD	After HH (loss)	After HH (win)	On first move
Always Hawk	0	0	0	0	0	0
25% Dove	0.25	0.25	0.25	0.25	0.25	0.25
50% Dove	0.5	0.5	0.5	0.5	0.5	0.5
75% Dove	0.75	0.75	0.75	0.75	0.75	0.75
Always Dove	1	1	1	1	1	1
Tit for Tat	1	0	1	0	0	1
Grim	1	0	0	0	0	1
Pavlov	0.99	0.01	0.01	0.99	0.99	0.5
Modified Pavlov	0.99	0.01	0.01	0.99	0.01	0.5

Table 5: Our memory-1 strategies. Each row gives the components of **s** of the specified memory-1 strategy. We show the probabilities that an animal plays Dove (otherwise, it plays Hawk) after different move pairs from it and its opponent (listed in that order) in the previous round. We abbreviate Dove as 'D' and Hawk as 'H', so the numbers in the 'After DH' column give the probability of playing Dove when the animal played Dove and its opponent played Hawk in the previous round. The final column gives the probability that an animal plays Dove in the first round.

In each tournament, we specify the parameters for the cost of fighting (c_W and c_L), the discount rate (γ), and the shape parameter of the symmetric beta distribution of win probabilities (k) at the outset, and we use the same values throughout the tournament. In all of our tournaments, we take $c_W = 0.1$, $\gamma = 0.995$, and $k = 1$. In this section, we describe the results of tournaments with three different values of c_L (0.2, 0.6, and 1.2). Combined with taking $c_W = 0.1$, these values of c_L correspond to different parameter regimes (see Section 3.1 and Figure 4). Specifically, $(c_W, c_L) = (0.1, 0.2)$ corresponds to regime I, $(c_W, c_L) = (0.1, 0.6)$ corresponds to regime II, and $(c_W, c_L) = (0.1, 1.2)$ corresponds to regime III. Our code (in MATLAB) for running our tournaments is available at https://bitbucket.org/CameronLHall/dominancesharingassessmentmatlab/.

We now discuss our parameter choices and the probability distribution from which we select p_A (i.e., the probability that animal A wins a fight). Taking $k = 1$ corresponds to choosing values of p_A from a uniform distribution between 0 and 1. We also perform numerical simulations (which we do not present in the paper) using the arcsine distribution $p_A \sim \text{Beta}(\frac{1}{2}, \frac{1}{2})$ and using a unimodal beta distribution with a reasonably large variance by taking $p_A \sim \text{Beta}(2, 2)$. We obtain qualitatively similar results for all three of these choices.

We choose the discount rate γ to be close to 1 so that the overall payoff from a contest is not dominated by payoffs from the first few rounds. Because of this, we use a large number of rounds (1000 of them in our simulations) in each contest to minimize numerical error from the finite nature of the contest. Because we take $\gamma = 0.995$ and the maximum (respectively, minimum)

28

payoff from each round is 1 (respectively, $-c_L$), the overall payoffs from our simulations are no more than $\frac{\gamma^{1000}}{1-\gamma} \approx 1.33$ smaller than and no more than $-1.33c_L$ larger greater than for contests with infinitely many rounds.

We perform strategy-pair comparisons by assigning one strategy to animal A and the other strategy to animal B and then calculating the mean payoff to animal A from a large number of simulations (using a wide range of values of p_A). The resulting mean payoff to animal A approximates the expected payoff to animal A in the given strategy pair. We use variance-reduction techniques (Owen, 2013) in order to improve the efficiency and accuracy of our estimates of expected payoff; the methods used to compute our results are described in Appendix A.

With 16 different strategies (seven informed ones and nine simple ones), we anticipate a tournament to consist of 256 strategy-pair comparisons. In practice, we can reduce the number of strategy-pair comparisons, because certain informed strategies are identical in some parameter regimes. For example, the 'Bully' and 'Nash' strategies are identical except when the stage game is PD or PD–Deadlock. Because neither of these stage games occur in parameter regime III, the 'Bully' and 'Nash' strategies correspond to identical behaviors in this parameter regime, so we can combine them. In our tabulated results (see Appendix B), we indicate where we combine informed strategies that are identical in a given parameter regime.

In Appendix B, we give a full description of our results in the form of three tables (see Tables B.1, B.2, and B.3) that show the mean payoff to animal A when it pursues the strategy in the row and animal B pursues the strategy in the column. To make ESSs easier to identify, we color the cells in Tables

B.1, B.2, and B.3 based on the value of $E(T,S) - E(S,S)$. We show cells as red when $E(T,S) - E(S,S) > 0$ and blue when $E(T,S) - E(S,S) < 0$; deeper shades correspond to larger differences. Therefore, a red cell indicates that the strategy in the row is able to invade the strategy in the column. If a strategy is a strong ESS, all cells in the corresponding column are blue, except for the one on the main diagonal.

3.4. Discussion of tournament results for informed strategies

From the tables in Appendix B, we see that no simple strategy is a pure ESS and that at least one informed strategy is able to invade each simple strategy (except the Pavlov strategy in parameter regime III). The only strategy that is a pure ESS in all parameter regimes is PD Tit for Tat (PD TfT); see column L in Table B.1, column N in Table B.2, and column J in Table B.3.

Interpreting Tables B.1, B.2, and B.3 as payoff matrices for the overall contests, one can use the method of Haigh (1975) to determine whether there are any mixed ESSs. Using this approach, we find that no mixed ESSs are possible in parameter regimes I or II. In parameter regime III, there is a mixed ESS in which 87% of the animals pursue a Pavlov strategy and 13% of animals pursue a Modified Pavlov strategy. This ESS corresponds to an expected payoff to the animals of 94.3 (which is less than the expected payoff of 99.3 of the PD TfT ESS). We concentrate on PD TfT in the rest of our discussion, but it is worth highlighting that a combination of simple strategies can be evolutionarily stable when fighting is very costly.

Consider the case in which animals A and B both pursue PD TfT strategies. We use the definition of PD TfT from Table 4 to describe how the

behavior of these animals depends on the model parameters c_W, c_L, and p_A. We give the details of this analysis in Appendix C. There are three possible pair behaviors (a dominance relationship, continual fights, and sharing of resources) in parameter regime I, two possible pair behaviors (a dominance relationship and sharing of resources) in parameter regimes II and IV, and only one possible pair behavior (a dominance relationship) in parameter regimes III, V, and VI. We illustrate this in Figure 4, which we color to indicate the values of μ_A and μ_B that are associated with the different pair behaviors.

4. Analysis of strategies for animals with no knowledge of p_A

The pair behaviors from Section 3.4 occur only if both animals in a contest pursue a PD TfT strategy. Because PD TfT is an informed strategy, the observation of such behavior depends on the ability of the two animals to perfectly assess each other's RHPs (and hence the value of p_A) before a contest begins. However, perfect assessment of RHP is unlikely without paying some cost; animals obtain some information from conventional signals, but there is evolutionary pressure towards costliness of signals and assessment to ensure their accuracy (Dawkins and Guilford, 1991; Arnott and Elwood, 2009). We incorporate a costly assessment process into our model by assuming that animals start a contest with no information about p_A but use (potentially costly) Hawk–Hawk fights to obtain information about p_A. An animal's beliefs about p_A guide its behavior.

31

4.1. Defining learning strategies

We model assessment using learning strategies with Bayesian updating (see Section 2.3). Suppose that animal A pursues a learning strategy. At the outset of a contest, we represent animal A's initial beliefs about p_A as a beta distribution Beta(α_0, β_0). We need to specify the values of α_0 and β_0 as part of the definition of the learning strategy. After each Hawk–Hawk fight, we update animal A's beliefs about p_A. If animal A has won m fights and lost n fights since the beginning of a contest, we represent its beliefs about p_A by the beta distribution Beta($\alpha_0 + m, \beta_0 + n$).

In each round, animal A decides its move using a memory-1 substrategy **s** (see Section 3.3) that depends on its current beliefs about p_A. We need to specify the relationship between animal A's beliefs (represented by the probability distribution Beta(α, β)) about p_A in a given round and the memory-1 substrategy **s** that it uses in that round as part of the learning strategy.

We base the relationship between animal A's current beliefs and its current substrategy on the PD Tit for Tat informed strategy. As an animal becomes more certain about p_A, its behavior should converge to that of an informed animal that uses a PD TfT strategy. Therefore, as α and β approach infinity with $\frac{\alpha}{\alpha+\beta} \to p_A$, the substrategy that is associated with Beta(α, β) should converge to the PD TfT strategy for that value of p_A.

There are numerous plausible learning strategies that achieve this, and there are numerous possible choices for α_0 and β_0. As with our informed strategies (see Section 3.2), we define a small set of plausible learning strategies and compare them using tournaments. We consider four types of learning strategies:

(1) weighted learning strategies,

(2) mean-based learning strategies,

(3) median-based learning strategies, and

(4) high-quantile learning strategies.

For fixed c_W and c_L, recall that the substrategy that is associated with PD TfT changes abruptly as p_A passes through the certain critical values. As we describe in Appendix C, we define $\tilde{p}_1$ and $\tilde{p}_2$ such that if animal A uses PD TfT, it pursues Always Dove when $0 \leq p_A < \tilde{p}_1$, Tit for Tat when $\tilde{p}_1 < p_A < \tilde{p}_2$, and Always Hawk when $\tilde{p}_2 < p_A \leq 1$. These critical values of p_A are important for all of the different learning strategies; we give them for each parameter regime in Table C.1.

In a weighted learning strategy, we construct the substrategy vector $\mathbf{s}$ by taking a weighted mean of substrategy vectors for the different substrategies of PD TfT. The weighting for Always Dove is given by the probability that $p_A \in (0, \tilde{p}_1)$, the weighting for Tit for Tat is given by the probability that $p_A \in (\tilde{p}_1, \tilde{p}_2)$, and the weighting for Always Hawk is given by the probability that $p_A \in (\tilde{p}_2, 1)$.

In a mean-based learning strategy, $\mathbf{s}$ is the substrategy of PD TfT that is associated with the mean of the distribution of an animal's current beliefs about p_A. The mean of $\mathrm{Beta}(\alpha, \beta)$ is $\frac{\alpha}{\alpha+\beta}$, so animal A's strategy is

$$
\mathbf{s} = \begin{cases} \mathbf{s}_{\mathrm{D}}, & 0 \leq \frac{\alpha}{\alpha+\beta} < \tilde{p}_1, \\[2mm] \mathbf{s}_{\mathrm{TfT}}, & \tilde{p}_1 \leq \frac{\alpha}{\alpha+\beta} < \tilde{p}_2, \\[2mm] \mathbf{s}_{\mathrm{H}}, & \tilde{p}_2 \leq \frac{\alpha}{\alpha+\beta} \leq 1, \end{cases} \tag{5}
$$

where $\mathbf{s}_{\mathrm{D}}$, $\mathbf{s}_{\mathrm{TfT}}$, and $\mathbf{s}_{\mathrm{H}}$ denote the substrategy vectors of 'Always Dove',

'Tit for Tat', and 'Always Hawk', respectively. To maximize the amount of information that an animal learns, we make the most aggressive available choice for **s** when $\frac{\alpha}{\alpha+\beta}$ is at a critical value. For example, when $\frac{\alpha}{\alpha+\beta} = \tilde{p}_2$, we use the Always Hawk substrategy, rather than Tit for Tat or Always Dove.

A median-based learning strategy is identical to a mean-based learning strategy, except that we use the median of the distribution Beta(α, β) instead of the mean. High-quantile learning strategies also use the same principle, but they are based on some other quantile of the distribution of an animal's beliefs about p_A. By basing such a learning strategy on a high quantile, an animal can behave aggressively until it has enough information to have a prescribed amount of confidence that aggressive behavior is suboptimal. For example, if animal A pursues a learning strategy that is based on the 0.95 quantile, it uses an Always Hawk substrategy until the 0.95 quantile of Beta(α, β) is below the critical value $\tilde{p}_2$. Therefore, the animal plays Hawk during every move until it is 95% confident that $p_A < \tilde{p}_2$.

In addition to specifying the relationship between current beliefs and current substrategy, it is important to specify each animal's initial beliefs about p_A (as represented by the parameters α_0 and β_0). Because we assume that animals initially have no information about p_A, apart from the fact that $p_A \sim$ Beta(k, k), it may appear natural to choose $\alpha_0 = \beta_0 = k$. However, there are potential advantages for an animal to use a prior that reflects its "optimism" about its chances of winning fights (as reflected by a large value of α_0 and/or a small value of β_0). Such an animal engages in more fights in the early rounds of a contest to gain information about its RHP relative to that of its opponent. One possible result is weak animals sustaining avoidable

costs that they would not obtain if using an unbiased prior, but these costs can be outweighed by the benefits to stronger animals that might otherwise have "given up" following an unlucky loss in an early fight. Consequently, the expected payoff (averaged over all values of p_A) may be higher for an animal that uses a biased prior with $\alpha_0 > \beta_0$, instead of an unbiased prior.

For the mean-based, median-based, and high-quantile learning strategies, we introduce two additional features to deal with problems that can arise when changing substrategies. The first feature ensures that sharing develops if both animals adopt a TfT substrategy after a period of being more aggressive. The second feature introduces some hysteresis that prevents frequent changes between the Always Hawk and TfT substrategies. We describe these features in detail in Appendix D.

4.2. Tournament for comparing learning strategies

As in Section 3.3, we use tournaments to compare the performance of several learning strategies against each other and against the 9 simple memory-1 strategies from Table 5. Our results are tabulated in Appendix B. We consider 16 different learning strategies. For each type of learning strategy (weighted, mean-based, median-based, and high-quantile), we consider 4 different strategies, which entail different levels of aggression. A more aggressive animal is willing to pay a higher cost for information about RHPs, and it requires more evidence (in the form of defeats in fights) before it adopts a less aggressive substrategy. In Table E.1, we outline the 25 learning strategies in our tournaments. We take $k = 1$, so we draw p_A from a uniform distribution. As with our tournaments in Section 3.3, we obtain the same qualitative results when we instead use $k = 1/2$ or $k = 2$.

For weighted, mean-based, and median-based learning strategies, we encode aggressiveness through the values of the shape parameters (α_0, β_0) that determine an animal's initial beliefs about its chances of winning a fight. We interpret these parameters biologically by comparing them with the true distribution $p_A \sim \text{Beta}(k, k)$ for the probability of winning a fight. When the distribution $\text{Beta}(\alpha_0, \beta_0)$ determines an animal's prior beliefs about winning a fight, the animal acts as if it believes that it has already won $\alpha_0 - k$ fights and lost $\beta_0 - k$ fights at the beginning of a contest. In our learning strategies, we fix $\beta_0 = k$ and modify the level of aggression by using different values of α_0. Specifically, we consider $\alpha_0 = k$, $\alpha_0 = k + 4$, $\alpha_0 = k + 8$, and $\alpha_0 = k + 12$.

As we described in Section 4.1, mean-based and median-based learning strategies can include a secondary prior $\text{Beta}(\bar{\alpha}_0, \bar{\beta}_0)$ to incorporate hysteresis into the process of changing substrategies. We take $\bar{\alpha}_0 = \bar{\beta}_0 = k$ in all mean-based and median-based strategies. For $k = 1$, this choice entails a flat prior when making decisions about increasing the level of aggression.

We use the same prior for all high-quantile strategies, for which we examine different aggression levels by using different quantiles of the distribution to reflect an animal's beliefs about its chances of winning a fight. For the four different aggression levels, we use the 0.8, 0.9, 0.95, and 0.98 quantiles. In the most aggressive strategy, which is based on the 0.98 quantile, an animal uses an Always Hawk substrategy until it has a 2% or lower confidence that p_A is in the range where Always Hawk is the best substrategy. We use a secondary quantile that is lower than the main quantile (see Appendix D) for deciding when to change from a less aggressive substrategy to a more aggressive one. We use 0.7, 0.8, 0.9, and 0.95 as the secondary quantiles for

0.8, 0.9, 0.95, and 0.98, respectively.

In Tables E.2, E.3, and E.4, we present our results for our tournaments in parameter regimes I, II, and III, respectively. We indicate strategies using our notation from Table E.1. We color the squares in Tables E.2–E.4 according to the values of $Q(T, S) = [E(T, S) - E(S, S)]/[E(\text{AllH}, \text{AllD})]$. See the supplementary material for the values of the mean payoffs $E(T, S)$. (They are also available at https://bitbucket.org/CameronLHall/dominancesharingassessmentmatlab/.) Each contest consists of 1000 rounds, and we use the parameter values $k = 1$, $\gamma = 0.995$, and $c_W = 0.1$. To examine results in the three main parameter regimes, we consider $c_L = 0.2$, $c_L = 0.6$, and $c_L = 1.2$ in the three tournaments. In all tournaments, we use the variance-reduction techniques from Appendix A.

Unlike in our tournaments in Section 3.3, we do not observe any strong ESSs. This is not surprising, because learning strategies with similar parameters (or ones with similar structures, like the mean-based and median-based strategies) are likely to perform similarly, and differences may be smaller than the 'noise' from simulations. Consequently, it is unlikely that a single learning strategy appears to be an unambiguous ESS based on our simulations. Additionally, it is difficult to apply the method of Haigh (1975) to find mixed ESSs for this tournament, both because the game that is defined by the mean payoffs is nongeneric and because the time that is required for Haigh's algorithm grows exponentially in the number of strategies. Instead, we draw qualitative conclusions about the relative performance of different learning strategies by examining Tables E.2, E.3, and E.4.

For the most part, we observe that the learning strategies (J–Y) are

resistant to invasion by the simple strategies (A–I). The least aggressive (i.e., with $\alpha_0 = \beta_0 = 1$) mean-based and median-based strategies are less successful than the other learning strategies. Weighted strategies are more successful than the least aggressive mean-based and median-based strategies, but they are less successful than other learning strategies, especially in parameter regime III. Additionally, the most aggressive mean-based, median-based, and high-quantile strategies are less successful (in the sense of being vulnerable to invasion) than some other learning strategies in parameter regime III.

We also observe many white and light-colored cells (especially in parameter regime I) when a pair of learning strategies interact with each other. Therefore, it appears that a variety of different learning strategies may be able to coexist successfully. In a population in which animals use results from fights to inform their long-term behavior, there may not be a single optimal strategy for using such information. Mean-based, median-based, and high-quantile learning strategies all have the potential to be effective if the aggression parameters (i.e., α_0 for the mean-based and median-based strategies and the quantile for the high-quantile strategies) are not too large or too small. Parameter regime I appears to have the broadest range of parameter values that correspond to effective learning strategies, possibly because the penalty for fighting is relatively small.

One prominent feature of the results in Tables E.2–E.4 is that every cell in column G is either dark blue or white. This illustrates that the simple 'Grim' strategy is very successful, although it is not a strong ESS, as evidenced by the white off-diagonal cells. These white cells indicate that mutant animals who use some other strategies (e.g., Always Dove or Tit for Tat) obtain the

same payoff as animals in a population who use the Grim strategy, enabling the mutants to invade the population because of genetic drift. A population of animals that play Grim all do very well, as they share their resources during each round. However, a mutant animal that pursues a learning strategy in a population of animals who play Grim is very unsuccessful. When a mutant plays Hawk to obtain information about its probability of winning fights, the response of a Grim is to punish the mutant by playing Hawk in every subsequent round. Even if the mutant has a relatively large RHP, this is likely to lead to an overall payoff that is lower than what it would obtain by playing Dove in every round.

We also find that the columns corresponding to strategies O, P, T, W, and X contain only white and blue cells in all parameter regimes. This indicates that these strategies are also difficult to invade; like Grim, they can only be invaded via genetic drift. Strategies O, P, T, W, and X all share the property that animals pursuing these strategies begin by playing Hawk at every round and then modify their behavior if they are unsuccessful in fights. This is consistent with the idea that successful animals without prior knowledge of p_A will use costly assessment (represented by fighting) in order to establish their strength relative to their opponent, and then adopt a long-term behavior that is consistent with their estimate of their own strength. From simulations of contests between animals pursuing these strategies (not shown here), we find that the assessment process takes longer (involving more rounds of fighting) when p_A is close to a critical threshold value. This indicates that assessment will be more costly when animals are evenly matched.

There are some statements in the abstract (points (1) and (2) in the

39

5. Conclusions and Discussion

We have developed a model of animal conflict based on an iterated Hawk–Dove game to analyze how different factors affect whether the outcome of a contest between animals is resource sharing, overt aggression, or the formation of a dominance relationship. We have shown that the same model can explain the conditions under which very different outcomes are evolutionarily stable and that key factors that determine which outcomes occur are costs of fighting, degree of asymmetry between animals in a contest, and the ability to learn from experience.

Although the Iterated Prisoners Dilemma and modifications of it (e.g., see Axelrod, 1984; Nowak, 2006) have been cited extensively as evidence for the evolution of cooperation and resource sharing, our findings demonstrate that sharing is stable only under rather limited conditions — specifically, when the cost of fighting is low and the animals in a contest have similar RHPs. As the asymmetries between animals become larger, resource sharing becomes less stable. By contrast, dominance hierarchies are stable for a much wider range of conditions. The explanation appears to be that progressively larger differences in RHPs make it progressively more beneficial both for the stronger animal in a contest to fight (rather than share)

40

and for the weaker animal to demur without even displaying or attempting to fight. Consequently, dominance hierarchies develop readily in the face of asymmetries and become progressively more stable for progressively larger asymmetries (Dawkins, 2010). This provides an explanation for why dominance hierarchies are widespread in so many species (Huntingford and Turner, 1987; Bonabeau, 1999; Braddock and Braddock, 1955; Beacham and Newman, 1987; Chase, 1982; Drummond, 2006; Guhl, 1968; O'Donnell, 1998).

When animals do not know RHPs at the outset of a contest, we found that the most successful strategies involve a limited period of assessment followed by a longer period in which they avoid fights and a stable dominance hierarchy arises. This is consistent with observations (Chase, 1982; Pagel and Dawkins, 1997). We also found that the duration of assessment depends both on the costliness of fighting and on the difference between the RHPs of the animals in a contest. Although animals that used assessment tended to outperform animals that used simple strategies, we also observed that Grim is resistant to invasion from mutants who use learning strategies. An interesting open problem is to determine whether it is possible for a mixture of simple strategies to be an ESS when some animals are capable of using assessment.

We have explored only a small number of the myriad strategies that are possible in our iterated asymmetric Hawk–Dove game. To promote future work, on we have developed software (see https://bitbucket.org/CameronLHall/dominancesharingassessmentmatlab/) to allow readers to pursue our ideas further and extend them to a wider range of questions. For example, our informed strategies from Section 3 are examples of the

memory-1 Markovian strategies in Nowak and Sigmund (1993b) and Press and Dyson (2012). An interesting extension is to consider memory-1 Markovian strategies for iterated asymmetric games more generally, following a similar approach to those in Nowak and Sigmund (1990) and Nowak and Sigmund (1993b). Other valuable extensions include the consideration of memory-s strategies for $s \geq 2$, analysis of the impact of bifurcations in the dynamics of tournaments with noise, and the study of tournaments of animals on networks (Szabó, 2007). We hope that the simplicity of our model and the ease of use of our software will enable the development of answers to these and many other questions about animal conflict.

Acknowledgements

References

Abreu, D., 1988. On the theory of infinitely repeated games with discounting. Econometrica 56 (2), 383–396.

Archer, J., 1988. The Behavioural Biology of Aggression. Cambridge University Press, Cambridge, UK.

Arnott, G., Elwood, R. W., 2009. Assessment of fighting ability in animal contests. Animal Behaviour 77 (5), 991–1004.

Aumann, R. J., Maschler, M. B., 1995. Repeated Games with Incomplete Information. MIT Press, Cambridge, MA, USA.

Axelrod, R., 1984. The Evolution of Cooperation. Basic Books Inc., New York, NY, USA.

Axelrod, R., Hamilton, W., 1981. The evolution of cooperation. Science 211 (4489), 1390–1396.

Baek, S. K., Yi, S. D., Jeong, H.-C., 2017. Duality between cooperation and defection in the presence of tit-for-tat in replicator dynamics. Journal of Theoretical Biology 430, 215–220.

Beacham, J. L., Newman, J. A., 1987. Social experience and the formation of dominance relationships in the pumpkinseed sunfish, *Lepomis gibbosus*. Animal Behaviour 35 (5), 1560–1563.

Bonabeau, E., 1999. Dominance orders in animal societies: The self-organization hypothesis revisited. Bulletin of Mathematical Biology 61 (4), 727–757.

Braddock, J. C., Braddock, Z. I., 1955. Aggressive behavior among females of the siamese fighting fish, betta splendens. Physiological Zoology 28 (2), 152–172.

Broom, M., Rychtář, J., 2013. Game-Theoretical Models in Biology. CRC Press, Boca Raton, FL, USA.

Bruns, B., 2015. Names for games: Locating 2×2 games. Games 6 (4), 495–520.

Carvalho, A. L. C., Rocha, H. P., Amaral, F. T., Guimaraes, F. G., 2016. Iterated prisoner's dilemma: An extended analysis. In: Anais do 11. Congresso Brasileiro de Inteligência Computacional. SBIC, pp. 1–6.

Chase, I. D., 1982. Dynamics of hierarchy formation: The sequential development of dominance relationships. Behaviour 80 (3), 218–239.

Clutton-Brock, T., Albon, S., 1979. The roaring of red deer and the evolution of honest advertisement. Behaviour 69 (3), 145–170.

Dawkins, M. S., 2010. Do asymmetries destabilize the Prisoner's Dilemma and make reciprocal altruism unlikely? Animal Behaviour 80 (2), 339–341.

Dawkins, M. S., Guilford, T., 1991. The corruption of honest signalling. Animal Behaviour 41 (5), 865–873.

Doebeli, M., Hauert, C., 2005. Models of cooperation based on the Prisoner's Dilemma and the Snowdrift game. Ecology Letters 8 (7), 748–766.

Drummond, H., 2006. Dominance in vertebrate broods and litters. The Quarterly Review of Biology 81 (1), 3–32.

Enquist, M., Leimar, O., 1983. Evolution of fighting behaviour: Decision rules and assessment of relative strength. Journal of Theoretical Biology 102 (3), 387–410.

Enquist, M., Leimar, O., Ljundberg, T., Mallner, Y., Segerdahl, N., 1990. A test if the sequential assessment game — Fighting in the cichlid fish *Nannacara anomala*. Animal Behaviour 40 (1), 1–14.

Eshel, I., 2005. Asymmetric population games and the legacy of Maynard Smith: From evolution to game theory and back? Theoretical Population Biology 68 (1), 11–17.

Eshel, I., Sansone, E., 2001. Multiple asymmetry and concord resolutions of a conflict. Journal of Theoretical Biology 213 (2), 209–222.

Fawcett, T. W., Johnstone, R. A., 2010. Learning your own strength: Winner and loser effects should change with age and experience. Proceedings of the Royal Society B: Biological Sciences 277 (1686), 1427–1434.

Friedman, J. W., 1971. A non-cooperative equilibrium for supergames. The Review of Economic Studies 38 (1), 1–12.

Fujiwara-Greve, T., 2015. Non-Cooperative Game Theory. Springer-Verlag, Heidelberg, Germany.

Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B., 2004. Bayesian Data Analysis, 2nd Edition. Chapman & Hall / CRC, Boca Raton, FL, USA.

Goessmann, C., Hemelrijk, C., Huber, R., 2000. The formation and maintenance of crayfish hierarchies: behavioral and self-structuring properties. Behavioral Ecology and Sociobiology 48 (6), 418–428.

Grewal, J. K., Hall, C. L., Porter, M. A., Dawkins, M. S., 2013. Formation of dominance relationships via strategy updating in an asymmetric Hawk–Dove game. arXiv:1308.5358.

Guhl, A. M., 1968. Social inertia and social stability in chickens. Animal Behaviour 16 (2–3), 219–232.

Haigh, J., 1975. Game theory and evolution. Advances in Applied Probability 7 (1), 8–11.

Hammerstein, P., 1981. The role of asymmetries in animal contests. Animal Behaviour 29 (1), 193–205.

Houston, A. I., McNamara, J. M., 1991. Evolutionarily stable strategies in the repeated hawk–dove game. Behavioral Ecology 2 (3), 219–227.

Hsu, Y., Earley, R. L., Wolf, L. L., 2005. Modulation of aggressive behaviour by fighting experience: mechanisms and contest outcomes. Biological Reviews 81 (01), 33.

Huntingford, F. A., Turner, A. K., 1987. Animal Conflict. Springer-Verlag, Heidelberg, Germany.

Jordan, J. S., 1995. Bayesian learning in repeated games. Games and Economic Behavior 9 (1), 8–20.

Kura, K., Broom, M., Kandler, A., 2015. Modelling dominance hierarchies under winner and loser effects. Bulletin of Mathematical Biology 77 (6), 927–952.

Kura, K., Broom, M., Kandler, A., 2016. A game-theoretical winner and loser model of dominance hierarchy formation. Bulletin of Mathematical Biology 78 (6), 1259–1290.

Maynard Smith, J., 1979. Game theory and the evolution of behaviour. Proceedings of the Royal Society B: Biological Sciences 205 (1161), 475–488.

Maynard Smith, J., 1982. Evolution and the Theory of Games. Cambridge University Press, Cambridge, UK.

Maynard Smith, J., Parker, G., 1976. The logic of asymmetric contests. Animal Behaviour 24 (1), 159–175.

Maynard Smith, J., Price, G. R., 1973. The logic of animal conflict. Nature 246 (5427), 15–18.

McNamara, J. M., Green, R. F., Olsson, O., 2006. Bayes' theorem and its applications in animal behaviour. Oikos 112 (2), 243–251.

Mesterton-Gibbons, M., Dai, Y., Goubault, M., 2016. Modeling the evolution of winner and loser effects: A survey and prospectus. Mathematical Biosciences 274, 33–44.

Mesterton-Gibbons, M., Karabiyik, T., Sherratt, T. N., 2014. The iterated hawk–dove game revisited: The effect of ownership uncertainty on bourgeois as a pure convention. Dynamic Games and Applications 4 (4), 407–431.

Mesterton-Gibbons, M., Sherratt, T. N., 2016. How residency duration affects the outcome of a territorial contest: Complementary game-theoretic models. Journal of Theoretical Biology 394, 137–148.

Nowak, M., Sigmund, K., 1990. The evolution of stochastic strategies in the Prisoner's Dilemma. Acta Applicandae Mathematicae 20 (3), 247–265.

Nowak, M., Sigmund, K., 1993a. Chaos and the evolution of cooperation.

Proceedings of the National Academy of Sciences of the United States of America 90 (11), 5091–5094.

Nowak, M., Sigmund, K., 1993b. A strategy of win-stay, lose-shift that outperforms tit-for-tat in the Prisoner's Dilemma game. Nature 364 (6432), 56–58.

Nowak, M. A., 2006. Five rules for the evolution of cooperation. Science 314 (5805), 1560–1563.

Nowak, M. A., 2012. Evolving cooperation. Journal of Theoretical Biology 299, 1–8.

O'Donnell, S., 1998. Dominance and polyethism in the eusocial wasp *Mischocyttarus mastigophorus* (Hymenoptera: Vespidae). Behavioral Ecology and Sociobiology 43 (4-5), 327–331.

Owen, A. B., 2013. Monte Carlo Theory, Methods and Examples.
URL https://statweb.stanford.edu/~owen/mc/

Pagel, M., Dawkins, M. S., 1997. Peck orders and group size in laying hens: 'Futures contracts' for non-aggression. Behavioural Processes 40 (1), 13–25.

Parker, G. A., 1974. Assessment strategy and the evolution of fighting behaviour. Journal of Theoretical Biology 47 (1), 223–243.

Parker, G. A., Rubenstein, D. I., 1981. Role assessment, reserve strategy, and acquisition of information in asymmetric animal conflicts. Animal Behaviour 29 (1), 221–240.

Payne, R. J. H., Pagel, M., 1996. When is false modesty a false economy? An optimality model of escalating signals. Proceedings of the Royal Society B: Biological Sciences 263 (1376), 1545–1550.

Press, W. H., Dyson, F. J., 2012. Iterated Prisoner's Dilemma contains strategies that dominate any evolutionary opponent. Proceedings of the National Academy of Sciences of the United States of America 109 (26), 10409–10413.

Robinson, D., Goforth, D., 2005. The Topology of the 2×2 Games: A New Periodic Table. Routledge.

Sastry, P. S., Phansalkar, V. V., Thathachar, M. A. L., 1994. Decentralized learning of Nash equilibria in multi-person stochastic games with incomplete information. IEEE Transactions on Systems, Man, and Cybernetics 24 (5), 769–777.

Szabó, G. & Fath, G., 2007. Spatial structure favors cooperative behavior in the snowdrift game with multiple interactive dynamics. Physics Reports 446, 97–216.

Trimmer, P. C., Houston, A. I., Marshall, J. A. R., Mendl, M. T., Paul, E. S., McNamara, J. M., 2011. Decision-making under uncertainty: Biases and Bayesians. Animal Cognition 14 (4), 465–476.

Trivers, R., 2006. Reciprocal altruism: 30 years later. In: Cooperation in Primates and Humans. Springer-Verlag, Heidelberg, Germany, pp. 67–83.

Valone, T. J., 2006. Are animals capable of bayesian updating? An empirical review. Oikos 112 (2), 252–259.

Watkins, C. J. C. H., Dayan, P., 1992. Q-learning. Machine Learning 8 (3-4), 279–292.

Wilkinson, G. S., 1984. Reciprocal food sharing in the vampire bat. Nature 308 (5955), 181–184.

Appendices

A. Computation of tournament results

To maximize the accuracy of our approximations of expected payoffs and ensure that it is reasonable to compare them between strategy pairs, we employ variance-reduction techniques that are common in Monte Carlo methods (Owen, 2013). We begin each tournament by generating a set of values for p_A, and we use the same set of values of p_A for each strategy-pair comparison. This is equivalent to using the method of common numbers for variance reduction (Owen, 2013); it increases the likelihood that the differences between mean payoffs for different strategy pairs result from differences in the expected payoffs, rather than from random fluctuations. We also use the method of antithetic sampling (Owen, 2013) to minimize any bias in favor of animals A or B in the calculated mean payoffs. We implement the method of antithetic sampling by ensuring, for each value of p_A, that we also use its complement $1 - p_A$ as a value for p_A.

In the tournaments in our paper, we draw 250 values for p_A from the uniform distribution $U(0, 1) = \text{Beta}(1, 1)$, and we then generate an additional 250 values of p_A by taking the complements of the first set of values. For each strategy-pair comparison, we run 20 contests with each of the 500 values of

p_A. This gives a total of $10,000$ contests for each strategy comparison; we report the mean payoff to animal A from these $10,000$ contests. The purpose of performing multiple contests with each value of p_A is to reduce the noise in our plots of mean payoff as a function of p_A. (We do not show these plots in this paper, but one can generate them using the code at https://bitbucket.org/CameronLHall/dominancesharingassessmentmatlab/.)

B. Tournament results for animals with knowledge of p_A

In Tables B.1, B.2, and B.3, we describe our results for simulations in parameter regimes I, II, and III, respectively. In these tables, we show the mean payoff to animal A when it pursues the strategy in the row and animal B pursues the strategy in the column.

The colors indicate whether a mutant with a different strategy can successfully invade a population of animals who are using a single strategy. In these tables, a mutant pursues a strategy in the row and the population strategy is in the column. The color of a cell indicates whether the expected payoff to the mutant is more than, the same as, or less than the expected payoff for an animal using the population strategy against another animal using the population strategy. A red cell indicates that a mutant has a higher payoff than animals in the population, so the row strategy can successfully invade the column strategy. By contrast, a blue cell indicates that the mutant has a lower payoff than animals in the population, so the column strategy is resistant to invasion by that row strategy. A white cell indicates that the mutant has a similar or identical payoff to the population, so genetic drift may lead to a mixed population with animals who pursue both strategies.

51

The intensity of the colors indicates the size of the difference between the expected payoff to a mutant and the expected payoff to an animal that uses the population strategy. Specifically, we color cells according to $Q(T,S) = [E(T,S) - E(S,S)]/[E(\text{AllH}, \text{AllD})]$, where T is the strategy that is pursued by animal A and S is the strategy that is pursued by animal B. Cells are dark when $|Q(T,S)| > 10^{-1}$, of a medium shade when $10^{-2} < |Q(T,S)| \leq 10^{-1}$, and light when $10^{-3} < |Q(T,S)| < 10^{-2}$.

	A	B	C	D	E	F	G	H	I	J	K	L
(A) AllH	69.5	101.8	134.1	166.4	198.7	70.1	70.2	134.2	102.2	108.3	75.5	75.8
(B) 25%D	52.2	82.6	113.0	143.4	173.8	83.0	53.0	113.0	84.6	93.7	60.5	73.7
(C) 50%D	34.8	63.3	91.9	120.4	149.0	92.2	35.9	91.9	67.3	79.0	45.5	71.2
(D) 75%D	17.4	44.1	70.8	97.5	124.2	97.6	19.5	70.8	50.2	64.3	30.4	67.9
(E) AllH	0.0	24.8	49.7	74.5	99.3	99.3	99.3	49.8	49.6	49.7	15.3	64.2
(F) TfT	69.2	82.3	91.6	97.4	99.3	99.3	99.3	92.3	79.2	58.7	60.1	65.0
(G) Grim	69.3	101.3	133.2	164.6	99.3	99.3	99.3	127.6	101.2	58.7	60.0	65.0
(H) Pav	34.7	63.3	91.9	120.4	148.7	92.8	46.3	98.6	90.2	79.2	45.5	71.3
(I) MPav	55.5	83.3	111.1	138.5	149.0	96.0	63.6	104.6	87.5	80.6	56.0	71.7
(J) Bul	60.5	82.6	104.8	126.9	149.0	110.4	110.4	104.5	101.0	99.3	66.6	81.3
(K) N, MN, SDT	72.7	100.3	128.1	155.7	183.4	88.5	88.4	127.8	110.2	111.5	78.7	79.0
(L) Slf, PDT, InT	72.5	92.6	109.7	123.6	134.5	112.6	112.6	110.2	103.2	97.0	78.6	87.5

Table B.1: Overall payoffs to animal A in parameter regime I when it pursues the strategy in the row and animal B pursues the strategy in the column. Our parameter values are $c_W = 0.1$, $c_L = 0.2$, $\gamma = 0.995$, and $k = 1$. We color cells according to $Q(T,S) = [E(T,S) - E(S,S)]/[E(\text{AllH}, \text{AllD})]$, where T is the strategy that is pursued by animal A (indicated by the row) and S is the strategy that is pursued by animal B (indicated by the column). Cells are red when $Q(T,S) > 10^{-3}$, blue when $Q(T,S) < -10^{-3}$, and white when $|Q(T,S)| \leq 10^{-3}$. Cells are dark when $|Q(T,S)| > 10^{-1}$, of a medium shade when $10^{-2} < |Q(T,S)| \leq 10^{-1}$, and light when $10^{-3} < |Q(T,S)| < 10^{-2}$. The column strategies correspond to the row strategies with the same letter. We assign a common letter to strategies that are identical in parameter regime I. We abbreviate strategies as follows: AllH = Always Hawk; X%D = X% Dove and otherwise Hawk; AllD = Always Dove; TfT = Tit for Tat; Pav = Pavlov; MPav = Modified Pavlov; Bul = Bully; N = Nash; PDT = PD Tit for Tat; MN = Mixed Nash; Slf = Selfish; SDT = Snowdrift Tit for Tat; and InT = Informed Tit for Tat.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
(A) AllH	29.7	72.0	114.2	156.4	198.7	30.6	30.7	114.2	70.8	77.4	62.5	62.9	43.2	62.7	43.7
(B) 25%D	22.4	60.2	98.1	136.0	173.8	60.7	23.3	98.1	60.9	70.5	56.9	71.9	43.1	64.3	58.1
(C) 50%D	14.9	48.4	81.9	115.5	149.0	82.3	16.3	81.9	51.3	63.5	51.3	77.0	42.6	64.3	68.2
(D) 75%D	7.5	36.7	65.7	95.0	124.2	95.2	9.7	66.0	42.2	56.6	45.7	78.1	41.6	62.6	73.9
(E) AllH	0.0	24.8	49.7	74.5	99.3	99.3	99.3	49.5	49.7	49.7	40.1	75.1	40.1	59.2	75.1
(F) TfT	29.7	60.1	81.8	94.9	99.3	99.3	99.3	83.3	58.8	27.8	22.5	54.7	22.6	35.9	54.7
(G) Grim	29.6	71.7	113.7	154.8	99.3	99.3	99.3	111.1	75.5	27.9	22.5	54.7	22.5	35.9	54.7
(H) Pav	14.9	48.4	82.0	115.5	148.8	83.5	29.7	97.7	84.9	63.2	51.1	77.7	42.7	64.4	68.9
(I) MPav	32.0	66.0	99.5	132.9	148.9	84.0	44.3	103.7	72.0	62.8	51.7	74.3	42.2	61.9	64.6
(J) Bul	51.7	76.0	100.3	124.6	149.0	101.6	101.6	100.2	95.7	99.3	84.5	110.0	65.2	94.1	90.7
(K) N, MN	53.2	79.6	105.9	132.2	158.5	93.8	93.7	106.0	97.8	100.8	86.0	102.0	66.7	86.1	82.8
(L) Slf	53.1	73.8	92.5	109.1	123.6	110.2	110.2	93.0	90.5	75.4	70.0	99.3	70.1	83.4	99.3
(M) SDT	50.4	81.0	109.1	135.0	158.5	93.7	93.6	109.3	97.6	98.0	83.1	102.0	63.9	83.3	82.7
(N) PDT	53.1	77.3	99.7	120.4	139.5	106.9	106.9	99.9	95.6	91.3	85.9	115.2	66.7	99.3	96.0
(O) InT	50.2	75.3	95.8	111.9	123.6	110.2	110.2	96.6	90.1	72.5	67.2	99.3	67.2	80.6	99.3

Table B.2: Overall payoffs to animal A in parameter regime II when it pursues the strategy in the row and animal B pursues the strategy in the column. Our parameter values are $c_W = 0.1$, $c_L = 0.6$, $\gamma = 0.995$, and $k = 1$. All other details are identical to those for Table B.1.

	A	B	C	D	E	F	G	H	I	J	K	L	M	N
(A) AllH	−29.7	27.2	84.5	141.5	198.7	−28.7	−28.7	84.5	25.8	33.0	25.1	50.1	−19.0	−18.7
(B) 25%D	−22.2	26.8	75.8	124.8	173.8	27.4	−21.2	75.7	27.4	37.2	30.1	68.0	5.6	22.3
(C) 50%D	−14.9	26.1	67.0	108.0	149.0	67.5	−13.3	67.1	29.3	41.3	35.4	79.3	26.2	52.5
(D) 75%D	−7.5	25.5	58.4	91.3	124.2	91.4	−4.9	58.4	31.3	45.6	40.6	83.9	43.2	72.3
(E) AllD	0.0	24.9	49.6	74.5	99.3	99.3	99.3	49.8	49.8	49.7	45.8	81.9	56.4	81.9
(F) TfT	−29.7	26.6	66.9	91.2	99.3	99.3	99.3	69.7	31.2	−16.3	−17.9	47.1	−6.4	47.0
(G) Grim	−29.6	27.3	84.1	140.4	99.3	99.3	99.3	86.4	37.6	−16.2	−17.7	47.0	−6.6	47.1
(H) Pav	−14.8	26.1	67.1	107.9	149.2	70.2	6.0	96.4	80.1	41.4	35.3	80.3	26.9	54.2
(I) MPav	−5.0	38.2	81.0	123.7	148.8	64.7	13.9	101.0	48.7	37.6	31.9	76.0	18.8	45.2
(J) Bul, N, PDT	36.5	64.6	92.7	120.9	149.0	86.5	86.4	92.8	86.8	99.3	90.2	131.5	53.8	79.4
(K) MN	34.5	64.2	93.6	123.3	152.8	83.7	77.7	93.7	85.5	96.4	88.7	135.4	51.2	76.6
(L) Slf	37.4	59.6	80.2	99.3	116.8	106.4	106.4	80.5	79.1	67.2	63.3	99.3	73.9	99.3
(M) SDT	5.2	52.9	91.4	121.3	142.2	94.1	94.1	93.0	74.4	61.4	59.8	124.8	33.4	87.0
(N) InT	5.3	48.8	81.9	104.6	116.8	106.3	106.4	83.8	67.7	36.1	34.3	99.3	45.8	99.3

Table B.3: Overall payoffs to animal A in parameter regime III when it pursues the strategy in the row and animal B pursues the strategy in the column. Our parameter values are $c_W = 0.1$, $c_L = 1.2$, $\gamma = 0.995$, and $k = 1$. All other details are identical to those for Table B.1.

C. Pair behaviors associated with the PD TfT strategy

Suppose that animals A and B both pursue PD TfT strategies. For fixed values of c_L and c_W, animal A's substrategy (Always Dove, Tit for Tat, or Always Hawk) depends on the value of p_A. To determine the parameter values that are associated with different pair behaviors, we seek critical values of p_A that separate the different substrategies of PD TfT. We define $\tilde{p}_1$ and $\tilde{p}_2$ such that animal A pursues Always Dove when $0 \leq p_A < \tilde{p}_1$, Tit for Tat when $\tilde{p}_1 < p_A < \tilde{p}_2$, and Always Hawk when $\tilde{p}_2 < p_A \leq 1$. From the definition of PD TfT in Table 4 and the boundaries between stage-game classifications (see Table 3), we calculate $\tilde{p}_1$ and $\tilde{p}_2$ in the different parameter regimes and obtain the results in Table C.1.

Parameter regimes	Critical values of p_A	
	$\tilde{p}_1$	$\tilde{p}_2$
I	$c_L/(1 - c_W + c_L)$	$(\frac{1}{2} + c_L)/(1 - c_W + c_L)$
II and IV	$c_L/(1 - c_W + c_L)$	$(1 - c_W)/(1 - c_W + c_L)$
III, V, and VI	$\frac{1}{2}$	$\frac{1}{2}$

Table C.1: Critical values, $\tilde{p}_1$ and $\tilde{p}_2$, of p_A in the PD TfT strategy. Animal A pursues Always Dove when $0 \leq p_A < \tilde{p}_1$, Tit for Tat when $\tilde{p}_1 < p_A < \tilde{p}_2$, and Always Hawk when $\tilde{p}_2 < p_A \leq 1$. Tit for Tat is not used in parameter regimes III, V, and VI, so $\tilde{p}_1 = \tilde{p}_2$ in those regimes.

If both animals in a contest pursue a PD TfT strategy, there are three different pair behaviors (depending on the values of μ_A and μ_B); we show these using colored shadings in Figure 4. If the stage game is PD, both

animals use a TfT substrategy, and we observe that each animal plays Dove in every round. If the stage game is PD–Deadlock or Deadlock–PD, one animal uses a TfT substrategy, and the other uses an Always Hawk substrategy. Consequently, both animals play Hawk in every round except for the first one. For any other stage game, the animal with the larger RHP uses an Always Hawk substrategy, and the animal with the smaller RHP uses Always Dove. This corresponds to a dominance relationship.

All three of the above possible pair behaviors can occur in parameter regime I. If there is a sufficiently large difference in the RHPs of the two animals (specifically, if $|p_A - \frac{1}{2}| > \frac{1 - c_W - c_L}{2(1 - c_W + c_L)}$), fighting is too costly for the animal with the smaller RHP, so there is a dominance relationship. If there is a moderate difference in the RHPs of the two animals (specifically, if $\frac{1 - c_W - c_L}{2(1 - c_W + c_L)} < |p_A - \frac{1}{2}| < \frac{c_W + c_L}{2(1 - c_W + c_L)}$), fighting is not too costly, and the two animals fight in every round. The weaker animal pursues a strategy of Tit for Tat, but this is not a sufficient threat to prevent the stronger animal from playing Hawk in every round, leading to continual fighting. If there is a sufficiently small difference in the RHPs of the two animals (specifically, if $|p_A - \frac{1}{2}| < \frac{c_W + c_L}{2(1 - c_W + c_L)}$), sharing resources gives a better payoff for the stronger animal than continual fighting. Therefore, Tit for Tat is a credible threat, and both animals play Dove in every round to avoid the cost of fighting.

In parameter regime II (and also in regime IV), the possible pair behaviors are dominance relationships (when $|p_A - \frac{1}{2}| > \frac{1 - c_W - c_L}{2(1 - c_W + c_L)}$) and resource sharing (when $|p_A - \frac{1}{2}| < \frac{1 - c_W - c_L}{2(1 - c_W + c_L)}$). Sharing, in which both animals pursue a Tit for Tat substrategy, occurs if and only if the expected payoff from a fight is positive for both animals. In this situation, the animal with the smaller

RHP expects to do better from a fight than it would from playing Dove against its opponent's Hawk, and the animal with the larger RHP expects to do better from sharing than it would from a fight. Therefore, the animal with the smaller RHP can credibly threaten the animal with the larger RHP, and Tit for Tat becomes an effective substrategy for both animals.

In parameter regime III (and also in regimes V and VI), the only possible pair behavior is a dominance relationship. The animal with the larger RHP always plays Hawk, and the animal with the smaller RHP always plays Dove. Because fighting is costly (especially for the loser) in parameter regime III, dominance relationships are favored, even when the difference between the animals' RHPs is arbitrarily small.

We see from the above results that the qualitative types of pair behaviors in our observations differ across the parameter regimes. In parameter regime I, we observe all three pair behaviors; in parameter regimes II and IV, we observe only dominance relationships and resource sharing; in parameter regimes III, V, and VI, we observe only dominance relationships. From Figure 5, it follows that $c_W + c_L$ (i.e., the sum of the fighting cost for the two animals) is a key parameter for determining possible pair behaviors. If $c_W + c_L < \frac{1}{2}$, a dominance relationship, continual fights, or sharing of resources can occur; if $\frac{1}{2} < c_W + c_L < 1$, sharing of resources or dominance relationships can occur; and if $c_W + c_L > 1$, there are only dominance relationships.

D. Incorporation of Tit for Tat into learning strategies

The first is a delay in adopting Tit for Tat. We specify that any animal that shifts from Always Hawk to Tit for Tat first plays Dove for two rounds

58

before beginning the Tit for Tat process. To understand why this is important, suppose that animal B pursues a Tit for Tat strategy and that animal A changes its substrategy from Always Hawk to Tit for Tat after round m. Because animal A has been playing Always Hawk up to this point and animal B is imitating this choice in its Tit for Tat strategy, animal B's move during round m is Hawk. If animal A immediately switches to Tit for Tat for round $m + 1$, the two animals continue to play Hawk during every move. If animal A plays Dove during round $m + 1$ and then starts copying animal B's previous move starting from round $m + 2$, the two animals will alternate Dove–Hawk and Hawk–Dove in all rounds starting from $m + 1$. The only way to create the possibility of Dove–Dove interactions in all future rounds is for animal A to play Dove for two rounds before pursuing a Tit for Tat substrategy.

The second feature is hysteresis, which prevents overly frequent changes in substrategies. For mean-based and median-based strategies, we specify two priors. One (the main prior, with parameters α_0 and β_0) determines whether to change to a less aggressive substrategy, and the other (the secondary prior, with parameters $\bar{\alpha}_0$ and $\bar{\beta}_0$) determines whether to change to a more aggressive substrategy. We require the secondary prior to be less optimistic about p_A than the main prior by demanding that $\bar{\alpha}_0 \leq \alpha_0$ and $\bar{\beta}_0 \geq \beta_0$.

Suppose that animal A pursues a mean-based learning strategy in which the secondary prior is different from the main prior. We use the main prior to decide the substrategy for the animal's first move of the contest. If $\frac{\alpha_0}{\alpha_0 + \beta_0} \geq \tilde{p}_2$, the animal's substrategy in the first round is Always Hawk. As the contest progresses, if the mean $\frac{\alpha}{\alpha + \beta}$ of the main distribution goes from above

$\tilde{p}_2$ to below $\tilde{p}_2$, then $\mathbf{s}$ changes from $\mathbf{s}_{\mathrm{H}}$ to $\mathbf{s}_{\mathrm{TfT}}$. The main prior thereby determines whether to change from a more-aggressive substrategy (Always Hawk) to a less-aggressive substrategy (Tit for Tat). If $\frac{\alpha}{\alpha+\beta}$ subsequently goes from below $\tilde{p}_2$ to above $\tilde{p}_2$, an animal does not immediately change substrategy again. If the true value of p_A is near $\tilde{p}_2$, there may be frequent changes between Always Hawk and Tit for Tat as $\frac{\alpha}{\alpha+\beta}$ fluctuates around $\tilde{p}_2$. Because an animal that changes from Always Hawk to Tit for Tat plays Dove for at least two rounds, it is suboptimal for it to make frequent changes in its substrategy. Instead, it changes from $\mathbf{s}_{\mathrm{TfT}}$ back to $\mathbf{s}_{\mathrm{H}}$ only if the mean $\frac{\bar{\alpha}}{\bar{\alpha}+\bar{\beta}}$ of the secondary distribution increases above $\tilde{p}_2$. Because $\bar{\alpha}_0 \leq \alpha_0$ and $\bar{\beta}_0 \geq \beta_0$ (where equality cannot occur simultaneously in these two inequalities when the main and secondary priors differ from each other), it follows that $\frac{\bar{\alpha}}{\bar{\alpha}+\bar{\beta}} < \frac{\alpha}{\alpha+\beta}$, so $\frac{\bar{\alpha}}{\bar{\alpha}+\bar{\beta}}$ increases above $\tilde{p}_2$ only after $\frac{\alpha}{\alpha+\beta}$ increases above $\tilde{p}_2$.

We also introduce hysteresis for high-quantile strategies, but now we do so by specifying a main quantile and a secondary quantile, rather than by specifying a main prior and a secondary prior. We require the secondary quantile to be lower than the main quantile. In a quantile-based strategy, an animal changes from Always Hawk to Tit for Tat if the main quantile of $\mathrm{Beta}(\alpha, \beta)$ goes from above $\tilde{p}_2$ to below $\tilde{p}_2$; it changes from Tit for Tat to Always Hawk if the secondary quantile of $\mathrm{Beta}(\alpha, \beta)$ goes from below $\tilde{p}_2$ to above $\tilde{p}_2$. Similarly, an animal changes from Tit for Tat to Always Dove if the main quantile of $\mathrm{Beta}(\alpha, \beta)$ goes from above $\tilde{p}_1$ to below it.

E. Tournament results for animals with no knowledge of p_A

In Table E.1, we list the 25 different strategies (which we label from 'A' to 'Y') in our tournament for animals with no knowledge of p_A. We show the results of our tournaments in Tables E.2 (for parameter regime I), E.3 (for regime II), and E.4 (for regime III). We color the cells of Tables E.2, E.3, and E.4 to indicate whether a mutant who pursues a strategy that is indicated by the row can invade a population of animals who pursue the strategy that is indicated by the column. This coloring system is identical to the coloring that we described in Appendix B for Tables B.1, B.2, and B.3.

Because of space limitations (and in contrast to Tables B.1, B.2, and B.3), we do not give numerical results for the mean payoff to animal A when it pursues the strategy in the row and animal B pursues the strategy in the column. These results are available at `https://bitbucket.org/CameronLHall/dominancesharingassessmentmatlab/`.

Strategies N–U are examples of mean-based and median-based learning strategies. As we described in Appendix D, we introduce hysteresis to the transitions between Tit for Tat and Always Hawk using a 'secondary prior' for deciding when to change from a less-aggressive substrategy to a more-aggressive one. For all of these strategies, we use a flat prior as the secondary prior (which entails that $\bar{\alpha}_0 = \bar{\beta}_0 = 1$).

map: please check the accuracy in the rephrasing of the paragraph above I made one minor alteration of phrasing, but otherwise all is fine.

Strategies V–Y are high-quantile strategies. For these strategies, we introduce hysteresis using a 'secondary quantile' for decisions that concern

changing from a less-aggressive substrategy to a more-aggressive one. For strategy V, we use a secondary quantile of 0.7; for W, we use 0.8; for X, we use 0.9; for Y, we use 0.95. In order to determine appropriate values for these secondary quantiles, we ran a series of simulations where Animal A used a high-quantile strategy, Animal B used an Always Hawk strategy, and p_A was close to $\tilde{p}_2$. We chose the secondary quantiles to be sufficiently close to the primary quantiles to enable animals to change substrategy from Always Hawk to Tit for Tat and back again, but far enough away from the primary quantiles that there were exceptionally few occasions when an animal would make three or more changes in substrategy (i.e., from Always Hawk to Tit for Tat to Always Hawk to Tit for Tat).

map: I don't understand the last sentence above; more than twice in total? in a certain window? need to clarify further

Is this clearer?

Label	Strategy type	Description/details of main prior
A	Simple	Always Hawk
B	Simple	25% Dove
C	Simple	50% Dove
D	Simple	75% Dove
E	Simple	Always Dove
F	Simple	Tit for Tat
G	Simple	Grim
H	Simple	Pavlov
I	Simple	Modified Pavlov
J	Weighted learning	Prior: $\text{Beta}(1,1)$ (flat)
K	Weighted learning	Prior: $\text{Beta}(5,1)$
L	Weighted learning	Prior: $\text{Beta}(9,1)$
M	Weighted learning	Prior: $\text{Beta}(13,1)$
N	Mean-based learning	Prior for decreasing aggression: $\text{Beta}(1,1)$ (flat)
O	Mean-based learning	Prior for decreasing aggression: $\text{Beta}(5,1)$
P	Mean-based learning	Prior for decreasing aggression: $\text{Beta}(9,1)$
Q	Mean-based learning	Prior for decreasing aggression: $\text{Beta}(13,1)$
R	Median-based learning	Prior for decreasing aggression: $\text{Beta}(1,1)$ (flat)
S	Median-based learning	Prior for decreasing aggression: $\text{Beta}(5,1)$
T	Median-based learning	Prior for decreasing aggression: $\text{Beta}(9,1)$
U	Median-based learning	Prior for decreasing aggression: $\text{Beta}(13,1)$
V	High-quantile learning	Flat prior; Quantile for decreasing aggression: 0.8
W	High-quantile learning	Flat prior; Quantile for decreasing aggression: 0.9
X	High-quantile learning	Flat prior; Quantile for decreasing aggression: 0.95
Y	High-quantile learning	Flat prior; Quantile for decreasing aggression: 0.98

Table E.1: Our 25 strategies in our tournaments of learning strategies. The first column indicates the label for the strategy (see Tables E.2–E.4), the second column indicates the strategy type, and the third column gives some further details about the strategy.

Table E.2: Overall payoffs to animal A in parameter regime I when it pursues the strategy in the row and animal B pursues the strategy in the column. The letters in the rows and columns correspond to the strategies in Table E.1. The parameter values are $c_W = 0.1$, $c_L = 0.2$, $\gamma = 0.995$, and $k = 1$. We color cells according to $Q(T, S) = [E(T, S) - E(S, S)]/[E(\text{AllH}, \text{AllD})]$, where T is the strategy that is pursued by animal A (indicated by the row) and S is the strategy that is pursued by animal B (indicated by the column). Cells are red when $Q(T, S) > 10^{-3}$, blue when $Q(T, S) < -10^{-3}$, and white when $|Q(T, S)| \leq 10^{-3}$. Cells are dark when $|Q(T, S)| > 10^{-1}$, of a medium shade when $10^{-2} < |Q(T, S)| \leq 10^{-1}$, and light when $10^{-3} < |Q(T, S)| < 10^{-2}$.

		Simple									Weighted				Mean				Median				Quantile			
		A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Table E.3: Overall payoffs to animal A in parameter regime II when it pursues the strategy in the row and animal B pursues the strategy in the column. The letters in the rows and columns correspond to the strategies in Table E.1. The parameter values are $c_W = 0.1$, $c_L = 0.6$, $\gamma = 0.995$, and $k = 1$. We color cells as in Table E.2.

Table E.4: Overall payoffs to animal A in parameter regime III when it pursues the strategy in the row and animal B pursues the strategy in the column. The letters in the rows and columns correspond to the strategies in Table E.1. The parameter values are $c_W = 0.1$, $c_L = 1.2$, $\gamma = 0.995$, and $k = 1$. We color cells as in Table E.2.