FIGHT OR FLIGHT: ENDOGENOUS TIMING IN CONFLICTS

Boris van Leeuwen, Theo Offerman, and Jeroen van de Ven*

Abstract—We study a dynamic game in which players compete for a prize. In a waiting game with two-sided private information about strength levels, players choose fighting, fleeing, or waiting. Players earn a "deterrence value" on top of the prize if their opponent escapes without a battle. We show that this value is a key determinant of the type of equilibrium. For intermediate values, sorting takes place, with weaker players fleeing before others fight. Time then helps to reduce battles. In an experiment, we find support for the key theoretical predictions and document suboptimal predatory fighting.

I. Introduction

FOLLOWING Maynard Smith's (1974) seminal contribution, competition for a prize is often modeled as the war of attrition. In this game, players choose the time at which they intend to flee. Time is costly, and players may differ in their opportunity costs. The player who waits the longest wins the prize and both players pay a cost proportional to the time it takes for the losing player to flee. Maynard Smith refers to this type of interaction as a "display." In a display, no physical contact takes place, although if it does, it does not settle the battle or convey information about which player would win an escalated conflict.

The main contribution of our paper is that we develop and analyze a game in which at any moment, players can not only wait or flee; they also have the option to start a fight. In case of a fight, a battle ensues and the stronger player wins the prize while the losing player incurs a loss. This dynamic fight-or-flight game allows us to make sense of a much wider variety of competitions. It captures the essence of many types of interactions in which the timing of actions plays a crucial role, such as R&D races, litigation, the launch of political or advertising campaigns, and firm acquisitions. It also fits situations in the animal kingdom, where animals fight over territory or prey. In all these examples, players can "flee" (e.g., reduce R&D spending, settle), wait to see if the other gives in, or initiate a fight (e.g., sue the opponent, start a hostile takeover), forcing the other into a battle.

Our dynamic game helps to understand why in some situations, players want to wait and see if the other flees without a battle, while in other circumstances, both want to act as quickly as possible. To illustrate the former type of situation, consider two political candidates who may wait a long time before they officially announce that they are running for office. If the other flees without a battle, they avoid the costs of a costly campaign that is required to win a fight. Male elephant seals that contest the right of exclusive access to a harem usually wait a couple of minutes to allow the other to flee without a bloody fight.

In other instances, players want to act as quickly as possible. A firm that wants to expand its market by acquiring a competitor should act quickly to prevent the prospective target from selling its valuable assets. Another possible interpretation is that compared to letting the other escape, by winning a fight the player sends a stronger signal about its strength to other players, thereby discouraging other players from ever making a challenge. A firm that drives out another firm by force will deter potential future competitors more than if the other firm left voluntarily. In a lawless society without a state monopoly of violence, people may want to rob each other if they can. In an encounter, the stronger player prefers to act as quickly as possible to avoid that the other flees without losing his money.

Notice that both types of examples are not well described by the war of attrition. In the first type of example, it may happen that players fight after a waiting period, which is not a possibility in the war of attrition. The war of attrition also does not capture the essence of the second type of interaction. In particular, the war of attrition does not accommodate that strong players decide to fight in a split-second.

In this paper, we analyze the fight-or-flight game theoretically and experimentally. Theoretically, we identify a key parameter, the deterrence value, that determines how the competition between two players will unfold. The deterrence value is the amount that a player earns on top of the prize if the other player manages to escape. Our theoretical analysis based on standard preferences yields two main novel insights. First, if the deterrence value is negative, all player types will rush and act in a split-second. A negative deterrence value is illustrated by the sale of valuable assets by a fleeing prospective firm in the takeover example. If the deterrence value is positive, players prefer to avoid the costly fight and wait before they act. In the example where two political candidates engage in a battle for office, the costs to organize a campaign represent a positive deterrence value.

The second insight is that if the deterrence value is positive but not too large, sorting will occur in the dynamic fight-or-flight game. That is, the weakest players will flee just before the end. Thus, the dynamic structure helps players to avoid costly fights, in comparison to a static version of the game that

is stripped of its time element. These two results cannot be obtained in a standard war of attrition. In that game, players' waiting correlates positively with their strength, and rushing by all types is never observed in equilibrium. Moreover, the dynamic standard war of attrition does not help players to sort and avoid costly fights in comparison to the static version (Hörisch & Kirchkamp, 2010).

We also investigate what happens in a behavioral model in which players differ in their degree of risk aversion. This model yields two additional testable implications. First, it predicts that sorting will occur in a wider set of circumstances than in the standard model. Second, it predicts that the more risk-averse players flee more frequently before the end.

We test the predictions in an experiment in which we systematically vary the deterrence value and the dynamic or static nature of the game between treatments. Our experimental findings support some of the key features of the theory, at least in terms of its comparative statics. With a negative deterrence value, subjects quickly learn to decide in a split-second. With a positive deterrence value, subjects tend to wait much longer and indeed use time to sort. In agreement with the model of heterogeneous risk aversion, we find that endogenous timing reduces the likelihood of costly battles in a wider set of circumstances than predicted by standard theory. Subjects classified as more risk averse on the basis of an independent task are indeed the ones who tend to flee more often early in the game. Thus, while not all results are consistent with the point predictions of the model, in terms of comparative statics, behavior often moves in the expected direction.

An interesting finding that deviates from the predictions is that a sizable minority of subjects fight early when the deterrence value is positive. This is the case even after ample time to learn. This finding is in stark contrast with some behavioral findings in related dynamic games. For instance, Roth, Murnighan, and Schoumaker (1988) report that the deadline effect, a striking concentration of agreements in the final seconds of the game, is the most robust behavioral finding in a class of games designed to test axiomatic models of Nash bargaining. Roth and Ockenfels (2002) and Ockenfels and Roth (2006) identify substantial last-minute bidding in second-price auctions. They attribute this phenomenon of sniping to both strategic and naive considerations of the bidders. We discuss some potential explanations for the anomaly of early fighting in our contest game at the end of the section IV.

One feature of our experimental design is that time is discrete but with very short time intervals. This makes it hard for subjects to precisely time their actions and could be one of the reasons behind the decrease in costly battles in the dynamic games. In a follow-up experiment, we make it easier for subjects to time their action by making the time intervals longer. Consistent with the theoretical predictions, we no longer observe a decrease in battles compared to the static games when the deterrence value is negative. In other respects, the results closely resemble that of the original experiment.

Our paper contributes to the literature on dynamic games in which players compete for a prize. Several studies compare dynamic with static environments. Hörisch and Kirchkamp (2010) investigate how experimental subjects behave in static and dynamic versions of the war of attrition and some closely related games. Theoretically, the dynamic version of a war of attrition does not help players to sort, and indeed, they do not observe such a difference in their experiments.[1] Theoretically, in an auction with symmetric interdependent valuations, Goeree and Offerman (2003a) do not find that the efficiency of a dynamic English auction is improved compared to the static second-price auction. In contrast, Kirchkamp and Moldovanu (2004) investigate a setup where a bidder's value is determined by his own signal in combination with the signal of his right neighbor. In this setting, bidders can retrieve valuable information in a dynamic auction process. In an experiment, they find that the efficiency of the English auction is higher than in a second-price auction in which no such information can be retrieved, which accords with theory.[2]

The remainder of the paper is organized as follows. Section II introduces the fight-or-flight game and presents the theory. Section III discusses the experimental design and procedures. Section IV provides the experimental results, and section V concludes.

II.   Theory

A.   Dynamic Fight-or-Flight game

We first describe the dynamic version of the fight-or-flight game. In this section, we present a basic version of the game. In section IIC and online appendix B, we discuss several extensions.

Time is discrete, with a finite number of periods $t = 0, 1, \ldots, T$. For each $t < T$, as long as the game has not ended, the two players independently decide to wait, flee (R,

---

[1]There is a large literature on static contest games. Carrillo and Palfrey (2009) study a contest game that is quite close to our static benchmark. They find that subjects compromise more often than in equilibrium, and they discuss some explanations based on cognitive limitations. De Dreu et al. (2016) investigate a game in which a group of attackers competes with a group of defenders. They find that in-group defense is stronger and better coordinated than out-group aggression. Oprea, Henwood, and Friedman (2011) show how the matching protocol affects outcomes in continuous time Hawk-Dove games. Dechenaux, Kovenock, and Sheremeta (2015) provide a survey of the experimental literature on contest games.

[2]The war of attrition has been applied to various settings, including versions with private information (Fudenberg & Tirole, 1986; Ponsati & Sákovics, 1995) and applications to public good provision (Bliss & Nalebuff, 1984; Weesie, 1993). Oprea, Wilson, and Zillante (2013) experimentally study war of attrition games with two-sided private information (as in Fudenberg & Tirole, 1986) and observe behavior close to theoretical predictions. More generally, the study of dynamic games reveals novel insights that significantly surpass what we know from the study of static games. Recent contributions include Potters, Sefton, and Vesterlund (2005), Levin and Peck (2008), Ivanov, Levin, and Peck (2009), Kolb (2015), and Agranov and Elliott (2017). The recent experimental literature on continuous time experiments shows that outcomes in continuous time may substantially differ from outcomes in discrete time (Friedman & Oprea, 2012; Oprea, Charness, & Friedman, 2014; Bigoni et al., 2015; Calford & Oprea, 2017).

for "retreat"), or fight (F). In the final period, players can no longer wait and have to choose F or R. The game ends with at least one player choosing F or R, at which point the action set becomes null. At the start, each player i is privately informed of her fighting ability a_i. It is common knowledge that a_i is independently drawn from a uniform distribution over the unit interval. A player's strategy lists for every ability the number of periods in which she chooses to wait and her choice if play reaches the period in which she wants to act. A player type's strategy $s(a_i)$ is described as (t, A), where $A \in \{F, R\}$. This means that player i with ability a_i will choose action A (fight or flee) in period t if the other player did not fight or flee earlier.

The game ends as soon as one of the players decides to fight or flee. The outcome can be a battle or an escape. A battle occurs if the player with the shortest waiting time chooses to fight or if both choose to fight at the same time. An escape occurs if the player with the shorter waiting time chooses to flee or if they both choose to flee at the same time. If one of the players chooses to fight and the other chooses to flee at the same time, an escape occurs with probability p and a battle with probability $1 - p$.

Payoffs. In case of a battle, the player with the higher ability receives $v^h > 0$ (the prize), and the other earns $-v^l$, where $v^h, v^l > 0$. In case of an escape, the player who chose to flee earns 0 while the other earns $v^h + k$, the prize plus a deterrence payoff k. This deterrence value can be positive or negative. A positive deterrence value captures situations where fighting is costly, so that players prefer to get the prize without fighting for it. A negative deterrence value captures situations in which beating the other generates a higher value compared to when the other escapes. We restrict the analysis to $k > -v^h$, so that if the other escapes, this always gives a higher payoff than escaping. As tie-breaking rules, we assume that if there is a battle between equally strong players, it is randomly determined (with equal probability) which player receives v^h and which player receives $-v^l$. If both players decided to flee at the same time, it is randomly determined (with equal probability) who earns 0 and who earns $v^h + k$. Alternatively, players could be allowed to share the prize equally in case both flee. This would not affect the theoretical analysis if players are risk neutral.

We assume that players maximize their expected utility and do not discount the future. In online appendix B, we analyze the case with discounting, but here our aim is to show how time per se affects the ability of players to sort themselves according to their strength. The case without discounting is also relevant for many cases, such as when the cost of waiting is small compared to the prize, the maximum duration of the game is short, or the consumption of the prize happens at a fixed point in time.[3]

We allow for the possibility that players are risk averse. To keep the model parsimonious, we assume that each player's utility function is piecewise linear in the payoff x and given by

$$U(x) = \begin{cases} x & \text{if } x \geq 0 \\ \lambda x & \text{if } x < 0. \end{cases} \tag{1}$$

Here, $\lambda > 0$ captures the degree of risk aversion (for $\lambda > 1$) or risk seeking ($0 < \lambda < 1$). Naturally, when $\lambda > 1$, this specification is also consistent with loss aversion. Our approach does not distinguish between loss and risk aversion.

B. Equilibrium

We look for pure-strategy Bayesian Nash equilibria. In this section, we derive equilibria under the assumption that players have threshold strategies, where types below a certain threshold flee and types above that threshold fight. Intuitively, stronger types have more to gain from fighting. We also assume that no type acts after the period in which the stronger type acts. In appendix A, we show that all equilibrium profiles satisfy these properties.

Negative deterrence value. $-v^h < k < 0$. For a negative deterrence value, the payoff of winning a battle exceeds that of allowing the other to escape. In this case, there is a unique equilibrium outcome in which all players fight or flee immediately. The very strong types will want to fight, and very weak types will want to flee. If the weakest types would flee after $t = 0$, the strongest types have an incentive to fight before that so the opponent does not escape. But then the weakest types would deviate to fleeing earlier. This implies that the strongest types fight immediately, and the weakest types flee immediately. Any other type will then act immediately as well. Acting later is costly, because it does not result in fewer battles with stronger types who fight and gives weaker types the possibility of escaping.

With all types acting immediately, let type $\tilde{a}$ be indifferent between fighting and fleeing. All stronger types fight and all weaker types flee. Suppose type $\tilde{a}$ flees. If the opponent is weaker, the expected payoff is $(v^h + k)/2$, and this happens with probability $\tilde{a}$. If the opponent is stronger, a battle results with probability $1 - p$, and this will always be lost by type $\tilde{a}$, giving a payoff $-\lambda v^l$. The expected payoff of fleeing is therefore given by

$$\tilde{a}\frac{1}{2}(v^h + k) + (1 - \tilde{a})(1 - p)(-\lambda v^l). \tag{2}$$

Suppose type $\tilde{a}$ fights. A weaker opponent escapes with probability p, giving a payoff $v^h + k$, and otherwise there is a battle that will be won by type $\tilde{a}$, giving a payoff v^h. If the opponent is stronger, there will always be a battle that will be lost by $\tilde{a}$. The expected utility of fighting is then given by

$$\tilde{a}[p(v^h + k) + (1 - p)v^h] + (1 - \tilde{a})(-\lambda v^l). \tag{3}$$

---

[3]By design, discounting also cannot play a role in the experiment.

Since type $\tilde{a}$ is indifferent between fleeing and fighting, it follows that

$$\tilde{a} = \frac{p\lambda v^l}{\frac{1}{2}v^h + p\lambda v^l + k\left(p - \frac{1}{2}\right)}. \tag{4}$$

The threshold $\tilde{a}$ is increasing in the probability of an escape p. As p increases, fighting against weaker types becomes less attractive since they become more likely to escape. More types will then flee in equilibrium. The effect of k on $\tilde{a}$ depends on the value of p. For $p < \frac{1}{2}$, an increase in k has a larger impact on the fleeing payoff than on the fighting payoff. This means fleeing becomes more attractive, and more types will flee in equilibrium. For $p > \frac{1}{2}$, the reverse is true.

Positive deterrence value. $k > 0$. With a positive deterrence value, players are better off when the other manages to escape than when they win a battle. In this case, all the action will be concentrated in the final two periods of the game. Intuitively, sufficiently strong players will wait until the last period to give other players the option to escape. Fighting should take place only in the last period. Weaker types will then also prefer to wait until at least the penultimate period, since waiting until then gives opponents the option to escape without the risk of ending up in a fight.

Consequently, for $k > 0$, there is a fraction of types that flees at $T - 1$ and a fraction that flees at T. The remaining fraction fights at T. All types that flee have the same payoff independent of the moment that they flee; they always lose a battle with a type that fights, and their payoff when the opponent flees is independent of their fighting ability. The equilibrium therefore does not pin down which types flee first, only the fraction. To determine the fraction of types that flee, we can assume without loss of generality that the weakest types flee at $T - 1$. The equilibrium can then be characterized by two threshold levels, $\hat{a}_1$ and $\hat{a}_2 > \hat{a}_1$. Type $\hat{a}_1$ is indifferent between fleeing at $T - 1$ and fleeing at T. Type $\hat{a}_2$ is indifferent between fleeing at T and fighting at T. A fraction of types $\hat{a}_1$ flees at $T - 1$, and a fraction of types $\hat{a}_2 - \hat{a}_1$ flees at T. Types above $\hat{a}_2$ fight at T. The values of $\hat{a}_1$ and $\hat{a}_2$ are given by

$$\hat{a}_1 = \frac{\lambda v^l[(v^h - k)(1 - 2p) - 2kp^2]}{(v^h + k + 2(1-p)\lambda v^l)\left(\frac{1}{2}v^h - \left(\frac{1}{2} - p\right)k\right)},$$

$$\hat{a}_2 = \frac{2(1-p)\lambda v^l}{v^h + k + 2(1-p)\lambda v^l}. \tag{5}$$

The fraction of types fleeing at $T - 1$ is positive for values of k below $\hat{k}$, where

$$\hat{k} = \frac{1 - 2p}{1 - 2p + 2p^2}v^h. \tag{6}$$

For larger values of k, all types wait until the final period. Intuitively, if k is large, it always pays off to wait and give

others the option to escape, even if that implies risking a battle with stronger types. The same is true for larger values of p. If the probability of an escape is large, it becomes more attractive to wait, even if the opponent fights.

Thus there can be three types of equilibrium outcomes. If $k < 0$, there is a *rushing equilibrium* in which all types immediately fight or flee. For intermediate positive values of k, there is a *timing equilibrium* in which some types wait until the penultimate period and then flee, while all others wait until the final period and then fight or flee. For high values of k, there is a *waiting equilibrium* in which all types wait until the last period and then fight or flee. While we derived these equilibria under the assumption that players have threshold strategies, in appendix A, we show that no other equilibria exist. The equilibrium outcome is generically unique, except for $k = 0$ or $k = \hat{k}$.

Proposition 1 (Equilibrium).

(i) *If $k < 0$, the unique equilibrium outcome is a rushing equilibrium in which all players act immediately. Players with abilities $[0, \tilde{a}]$ flee at $t = 0$, and players with abilities $(\tilde{a}, 1]$ fight at $t = 0$.*

(ii) *If $0 < k < \hat{k}$, the unique equilibrium outcome is a timing equilibrium in which a fraction $\hat{a}_1$ of types flee in period $T - 1$, a fraction $\hat{a}_2 - \hat{a}_1$ of types flee in period T, and all types above $\hat{a}_2$ fight in period T.*

(iii) *If $k > \max\{\hat{k}, 0\}$, the unique equilibrium outcome is a waiting equilibrium in which types $[0, \tilde{a}]$ flee in period T and types $(\tilde{a}, 1]$ fight in period T, and $\tilde{a} = 1$ for any $v^h < (1 - 2p)k$.*

Proof. All proofs are in appendix A.

Figure 1 illustrates the equilibrium outcomes. Figure 1a shows equilibrium outcomes for different combinations of the probability of an escape (p) and the deterrence value (k). Figure 1b shows how the threshold values change with k. For $k < 0$, fewer types fight as k increases. A higher k makes letting the other escape relatively more attractive, and such an escape becomes less likely by fighting. This reverses for positive values of k, with more types fighting as k increases. For higher values of k, fewer types flee early. Fighting becomes relatively more attractive with more weaker types still around. The figure also illustrates how these thresholds change with an increase in p.

To shed light on whether the dynamic time element of the fight-or-flight game decreases costly battles, we use a static version of the game as benchmark. In the static game, players choose simultaneously between fight and flee, and the same payoffs result as when players reach the final period of the dynamic game. The Bayesian Nash equilibrium of the static game coincides with the equilibrium of the dynamic game for parameters where all players act in the same period (that is, either case i or case iii described in proposition 1).

An interesting feature of the timing equilibrium of the dynamic game is that sorting takes place over time, resulting in fewer battles compared to what happens in the static game.
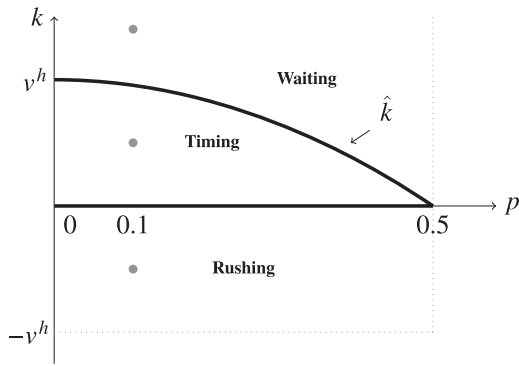
(a) Deterrence value and escape probabilities
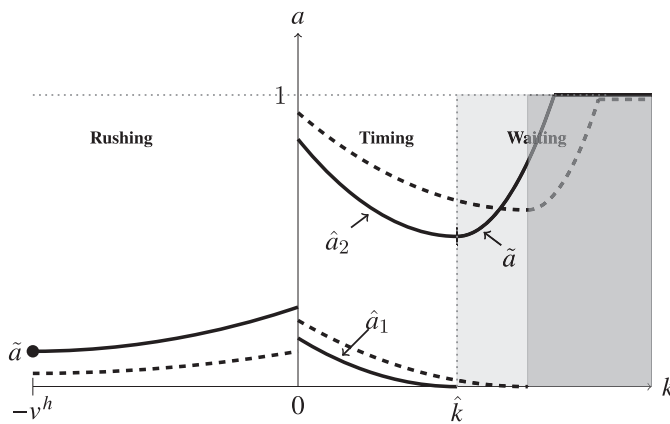
(b) Tresholds

(a) The solid dots indicate the experimentally implemented values (with $v^h = v^l = 10$, $p = 0.1$, and $k = \{-6, 6, 12\}$). (b) Rushing occurs to the left of the vertical axis, timing occurs between the vertical axis and the shaded area, waiting occurs in the shaded area. The dashed lines show a decrease in the escape probability (p) (for $p < \frac{1}{2}$). The dark shaded area shows the waiting equilibrium for the lower value of p.

In the dynamic game, the strongest types remain in the game until the last period, while some weaker types flee before any battle may take place. Moreover, a smaller fraction of types will fight; fighting becomes less attractive with fewer relatively weak players remaining.

Proposition 2 (Battles and Sorting). *Compared to a static (simultaneous-move) version of the game:*

(i) *The frequency of battles is reduced in case of a timing equilibrium and the same in case of a rushing or waiting equilibrium.*
(ii) *The rate at which the weaker player in a pair manages to escape is increased in case of a timing equilibrium and the same in case of a rushing or waiting equilibrium.*

C. Extensions

Heterogeneous risk aversion. A surprising feature of the analysis with a homogeneous population is that the set of de-

terrence values for which the timing equilibrium materializes does not depend on players' risk aversion. This result changes when the population is heterogeneous in the degree of risk aversion. Intuitively, players who are relatively averse to risks will want to flee earlier. Indeed, a population that is heterogeneous in the degree of risk aversion can sustain a timing equilibrium for a larger set of deterrence values. We show this in a simple framework with two levels of risk aversion and outline the two main strategic features of this model.

Suppose that a fraction $1 - q$ of the population has a risk aversion parameter λ_1, and a fraction q has $\lambda_2 > \lambda_1$. A player's value of λ is private information, but all players know the distribution. Consider the case where q is very small. In that case, the threshold levels derived assuming homogeneous risk aversion in section IIA are not much affected for the less risk-averse types. Fix an equilibrium in which $k > \hat{k}$, so that all types with λ_1 wait until period T.

If λ_2 is such that

$$\tilde{a}\frac{1}{2}(v^h + k) + (1 - \tilde{a})(1 - p)(-\lambda_2 v^l) < 0, \qquad (7)$$

then types with λ_2 and a fighting ability less than or equal to $\tilde{a}$ prefer to flee in period $T - 1$, while types with λ_1 prefer to wait until T. Thus, for the same level of k, we now have a timing equilibrium instead of a waiting equilibrium.

Another feature of this model is that the more risk-averse types will be the ones who flee more frequently before the end. To see this, note that for the ability level for which the less risk-averse type is indifferent between fleeing in period $T - 1$ and period T, the more risk-averse type still strictly prefers to flee in period $T - 1$. The reason is that the expected payoff of fleeing in period $T - 1$ is not affected by the degree of risk aversion (since there are no negative payoffs), while the expected payoff of fleeing in period T decreases in a player's risk aversion (since the negative payoff when a battle is lost weighs more heavily). In the experiment, we will test these two implications of the model with heterogeneous risk aversion.

Other extensions. We also considered some natural extensions of the model. Here we describe the main qualitative features of these extensions. In appendix B, we provide further details of these extensions, as well as discussion of the pros and cons of discrete versus continuous modeling in waiting games.

So far we simply assumed that the stronger player always wins a battle. A natural possibility is that stronger types are more likely to win but do not win with certainty. When relative strength correlates sufficiently strongly with winning a battle, the results are qualitatively the same. That is, with a positive deterrence value, all the action will be concentrated in the final two periods. The strongest types still want to fight in the final period, while no type wants to flee before the penultimate period. Likewise, with a negative deterrence value, all types will still act immediately. When the link between relative

strength and winning a battle becomes weak, other types of equilibria exist. In the extreme case, where each type has an almost equal chance of winning a battle against any other type, there can be equilibria where all types prefer to fight, possibly at different periods. There can also be an equilibrium in which all types prefer to flee in the last period.

Another natural extension is discounting of future payoffs. Conditional on a discount factor sufficiently close to 1, our main theoretical findings remain qualitatively similar. That is, we find a rushing equilibrium when $k < 0$, a timing equilibrium when $0 < k < \hat{k}$, and a waiting equilibrium when $k > \max(0, \hat{k})$. As in the timing equilibrium without discounting, all the action happens in the penultimate and last periods. The main difference with the model without discounting is that the thresholds now also depend on the discount factor. When discounting is important, the comparison between the static and dynamic case becomes less clear-cut in terms of welfare: the higher degree of sorting comes at the cost of waiting longer.

A final variant that yields somewhat different predictions is the one where players face a known cost for time. Here, it may happen that weak players decide to drop out earlier than the penultimate period. In a recent paper, Song and Houser (2021) study the interesting case of costly waiting in detail.

In the experiment, we focus on the variant where time is not costly for two reasons. First, it allows us to investigate in a meaningful way how the dynamic game helps players avoid costly battles compared to the static game where time plays no role. Second, we think that it is a stronger result if players use time as a sorting device when time is not costly.

III.   Experimental Design and Procedures

A.   Design

Subjects participated in a laboratory experiment in which they played the fight-or-flight game. In all treatments, we set the value of winning a battle to $v^h = 10$ and losing a battle to $-v^l = -10$. The probability of an escape when at the same time one player decided to fight and the other decided to flee was set to $p = 0.1$. Each subject played the game forty rounds, with random rematching after every round within a matching group of eight subjects. At the start of each round, the subjects were informed of their fighting ability for that round, which was an integer number from $\{0, 1, 2, \dots, 1000\}$. They knew that each number was equally likely, that each subject faced the same distribution, and that draws were independent across subjects and rounds. At the end of a round, each subject was informed of the outcome, the paired subject's fighting ability, and the resulting payoffs.

We implemented two treatment variations. The first treatment variable was the deterrence variable k, which was either -6, 6, or 12. The second concerned the dynamic or static nature of the flight-or-fight game. This gives a 3×2 design. Every subject participated in only one of the treatments. In to-

TABLE 1.—OVERVIEW OF TREATMENTS

Treatment Version	Deterrence Value (k)	N Subjects	N Matching Groups
Dynamic	−6	64	8
Dynamic	6	56	7
Dynamic	12	64	8
Static	−6	56	7
Static	6	64	8
Static	12	56	7

tal, 360 subjects participated, with seven or eight independent matching groups per treatment. Table 1 presents an overview.

In the *dynamic* fight-or-flight game, a 5 second countdown started after all subjects in the laboratory had indicated that they were ready to start. This ensured that subjects knew exactly when the game would start. During the game itself, a clock started counting down from 10 seconds to 0. The program divided the 10 seconds in 50 periods of 200 milliseconds each. Subjects implemented their strategies in real time. For instance, a subject could decide to wait for 5 seconds (i.e., for the first 25 periods) and then choose to fight, which would determine the outcome of the game (unless the other subject had already terminated the game earlier). This way she would implement the strategy $(25, F)$. If subjects let the time run down to 0, they entered the *endgame*, in which they simultaneously decided between fight and flee (with no time constraints, as they decided simultaneously anyway).

Our dynamic game has 50 periods, more than the minimum required to test the theoretical predictions of the model, and short time intervals of 200 ms. Our goal was to have a design that is closer to the examples that motivated our research. A disadvantage is that rational subjects might find it hard to exactly implement equilibrium strategies in our setup. A follow-up experiment with longer time intervals addresses this concern (see section IVD).

The *static* version of the game abstracted from the time element and only consisted of the endgame of the dynamic version. That is, in this version of the game, subjects were immediately put in the same position as the players of the dynamic game who had both decided to wait until the end of the game. So in the static game, both subjects simultaneously chose between fight and flee.

After the main part, we obtained additional measurements. We assessed subjects' risk aversion using the method of Gächter, Johnson, and Herrmann (2007). A subject chooses whether to accept or reject six different lotteries. In a lottery, the winning amount is 6 euros. The losing amount varies across lotteries, from two to seven. In each lottery, the winning and the losing amounts are equally likely. If a subject rejects a lottery, she surely receives 0 euro. At the end of the experiment, one of the six lotteries is selected at random and played out for actual payment. The number of rejected lotteries is our measure of a subject's degree of risk aversion.

We also measured physical strength. We asked subjects to press a hand dynamometer as hard as they could, following

the procedure of Sell et al. (2009). This measurement was obtained twice, and the best attempt was rewarded with 5 eurocents per kilo pushed. Finally, we obtained some self-reported measurements on social dominance and prestige (from Cheng, Tracy, & Henrich, 2010), perceived masculinity, sex, and age.[4]

This design allows us to investigate the predictions summarized in propositions 1 and 2. In addition, it makes it possible to test the predictions from the behavioral model of heterogeneous risk aversion.

B. Procedures

The experiment was computerized and run at CREED (University of Amsterdam). The instructions are in appendix E. Subjects read the instructions at their own pace. They could continue only after correctly answering test questions at the end of the instructions. To ease understanding, we used non-neutral labels such as "fight" and "escape." Subjects were informed that there would be two parts, receiving new instructions at the start of each part.

During the experiment, subjects earned points, where 1 point = €0.70 (≈$0.84). To avoid a net loss at the end of the experiment, they received a starting capital of 21 points, and any profits or losses would be added to or subtracted from this. At the end of the experiment, one round of the main part was randomly selected for payment. Total earnings averaged €19.09, ranging from €5.30 to €38.20.[5] A session took approximately 65 to 75 minutes.[6]

IV. Results

In sections IVA and IVB, we consider the testable predictions following from propositions 1 and 2, respectively. Then, in section IVC, we turn to decisions at the individual level. All statistical tests comparing treatment differences use matching group averages as the independent unit of observation, unless indicated otherwise.

A. Timing of Actions

Following proposition 1, we address the comparative static prediction that the timing of actions is influenced by the deterrence value. Specifically, we expect very quick decisions if the deterrence value is negative and decisions in the final periods if the deterrence value is positive. Figure 2 shows the average elapsed time before subjects made a decision
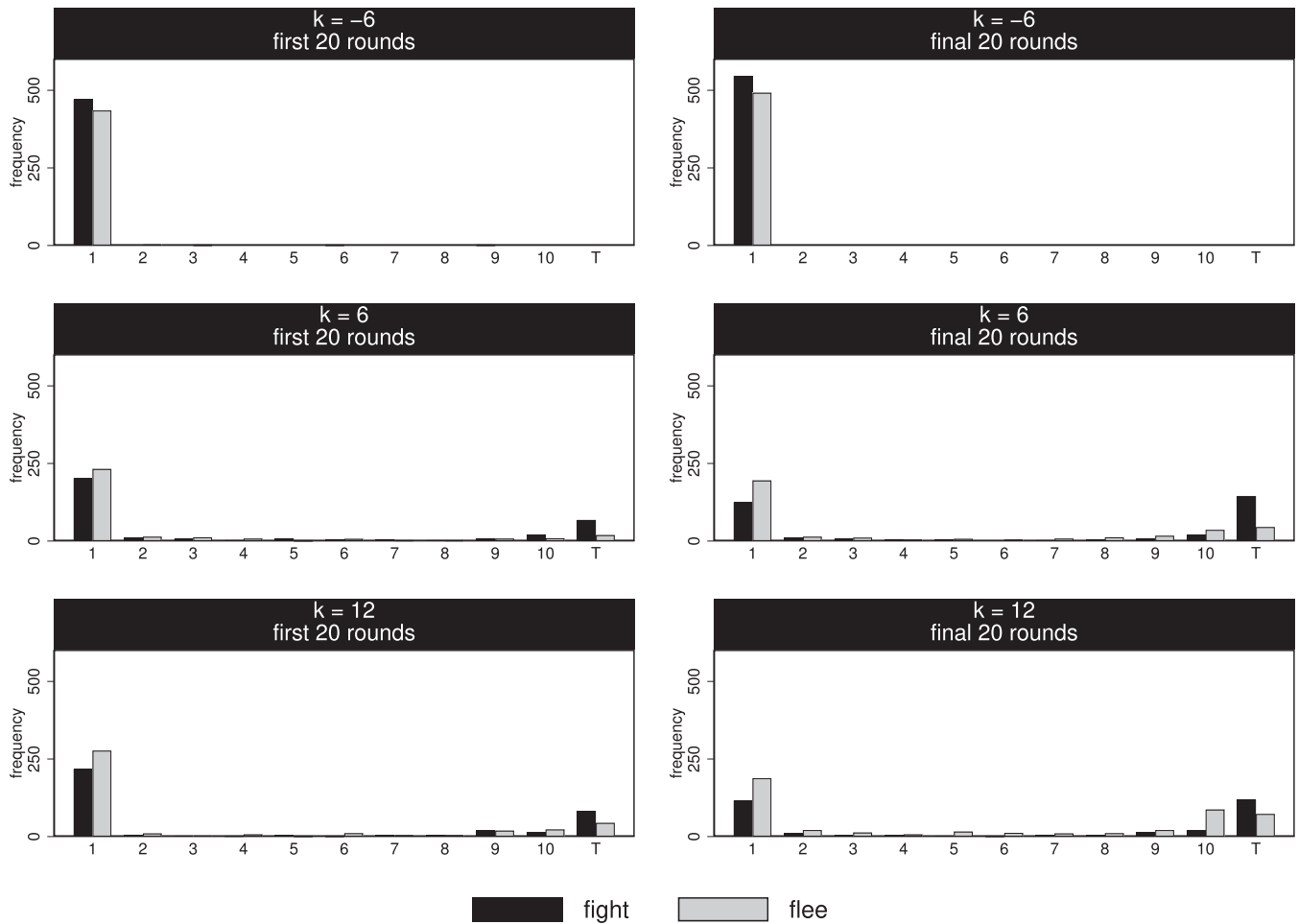
[4]Perceived masculinity is measured by the answer to the question: "On a scale from 1 (very feminine) to 7 (very masculine), how would you describe yourself?"

[5]The payment subjects received consisted of the starting capital and their earnings in the fight-or-flight game, the lottery task, and the physical strength task.

[6]In addition to the forty decision rounds (which lasted around 20 minutes), subjects spent time on the instructions and test questions (25 minutes), the lottery task, questionnaire and physical strength task (15 minutes), and payment of subjects (10 minutes).

FIGURE 2.—AVERAGE WAITING TIME (IN MS) BEFORE SUBJECTS MAKE A DECISION IN THE DYNAMIC GAME, BY TREATMENT AND ROUND
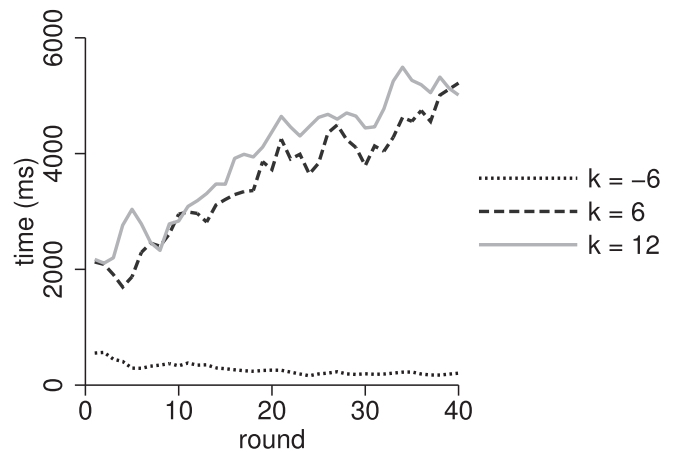
Lines are moving averages of three rounds.

in the dynamic games. As predicted, we observe a clear effect of the deterrence value on the timing of actions. With a negative deterrence value, subjects tend to fight or flee almost immediately. On average, subjects make a decision after 273 ms. When the deterrence value is positive, subjects tend to wait much longer. For $k = 6$, the average elapsed time before making a decision is 3,545 ms, and for $k = 12$, this is 3,973 ms. For both treatments with a positive deterrence value, the average waiting time is significantly longer than for $k = -6$ (Mann-Whitney tests, $p = 0.001$, $N = 15$ for $k = -6$ vs $k = 6$ and $p < 0.001$, $N = 16$ for $k = -6$ versus $k = 12$). While subjects wait slightly longer when $k = 12$ than with $k = 6$, the difference is not statistically significant (Mann-Whitney test, $p = 0.908$, $N = 15$ for $k = 6$ vs $k = 12$). For all three treatments, we observe learning effects. When the deterrence value is positive, subjects learn to wait, reflected by the strong, positive time trend over the rounds. The reverse holds for the negative deterrence value. In this case, subjects decide increasingly faster. The average elapsed time is 402 ms in the first ten rounds and 200 ms in the final ten rounds. When comparing the average waiting times in the first ten rounds and final ten rounds, all time trends are statistically significant (Wilcoxon signed-rank tests, $p = 0.017$, $N = 8$ for $k = -6$, $p = 0.018$, $N = 7$ for $k = 6$ and $p = 0.017$, $N = 8$ for $k = 12$).

Figure 3 gives a more detailed picture of the timing of decisions. The figure plots the distribution of actions for each of the 10 seconds plus the endgame (T). The left panels show this for the first twenty rounds and the right panels for the final twenty rounds. Several patterns emerge. First, with a negative deterrence value, we clearly observe rushing: subjects decide almost immediately. None of the matches make it to the endgame, and 99.5% of all matches end in the first second. In fact, 90% of all matches end within the very first 200 ms, that is, in the first period.[7] With a positive deterrence

[7]Figure A1 in appendix C shows the distribution of actions by 200 ms periods.

FIGURE 3.—DISTRIBUTION OF DECISIONS OVER TIME (SECONDS) BY DETERRENCE VALUE IN THE DYNAMIC GAME

Period T indicates the endgame. Left panels are for the first twenty rounds and right panels for the final twenty rounds. Only observations where a player made a decision to fight or flee are included in the graph, that is, observations where a player was waiting when the other moved are omitted.

value, most action is at the very beginning and the very end: subjects tend to decide either relatively quickly or wait until the final periods. In the final twenty rounds, a larger fraction of subjects waits until the end. This fraction might be underestimated because a subject who is willing to wait until the end will reach the end of the game only if the paired player is also willing to wait until then. Among those waiting, there are some subjects who flee right before the endgame.

Result 1. When the deterrence value is negative, players act immediately. When the deterrence value is positive, players are more likely to wait until the end of the game and they learn to wait longer.

In contrast to the theoretical predictions, some subjects move at the very beginning of the game when the deterrence value is positive. This fraction decreases over time, but even in the final twenty rounds (the right-hand panels of figure 3), we do observe such behavior. This behavior is not in line with the timing equilibrium or waiting equilibrium. We return to this anomaly when we discuss individual behavior (section

IVC). The comparative static results of increasing k are in line with the theoretical predictions.

B. Frequency of Battles and Sorting

The second main testable prediction, following from proposition 2, is that endogenous timing helps to avoid costly battles. Specifically, we expect fewer battles in the dynamic games in case of a timing equilibrium, but not in case of a rushing or waiting equilibrium. The left panel of figure 4 shows the frequency of battles for each treatment (we discuss the results for experiment 2 in section IVD). We do indeed observe fewer battles in the dynamic treatments compared to the static treatments. The difference varies between 15 and 26 percentage points depending on the deterrence value and is always highly significant ($p < 0.003$ in each case, two-sided Mann-Whitney tests). A regression analysis (table A1 in appendix C, column 1) confirms that there are fewer battles in the dynamic treatments, and this effect is slightly stronger when the deterrence value is positive.

FIGURE 4.—FRACTION OF BATTLES (LEFT PANEL) AND FRACTION OF TIMES THAT THE WEAKER PLAYER IN A PAIR ESCAPES (RIGHT PANEL)
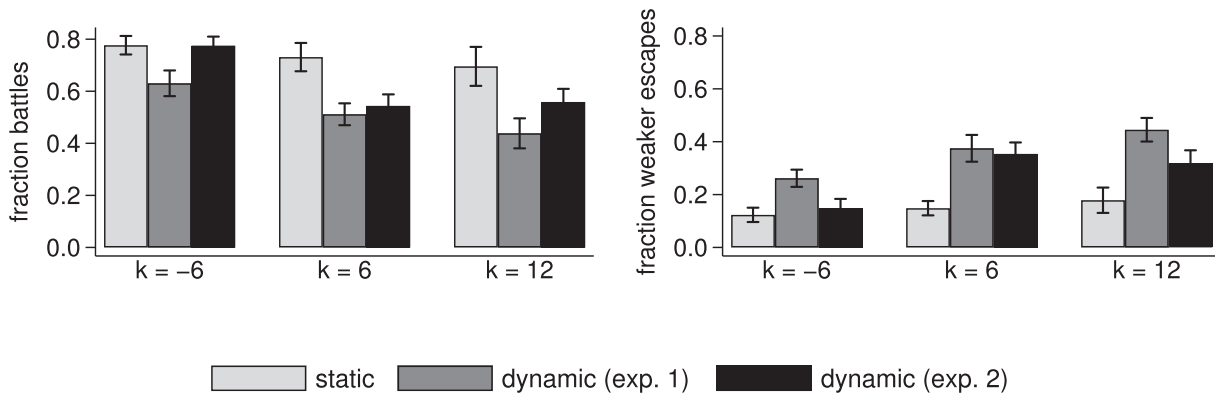
Error bars indicate 95% confidence intervals, based on matching groups as the independent unit of observation.

The reduction of battles for $k = 6$ is in line with the comparative static prediction following from proposition 2. For $k = 6$, the unique equilibrium outcome in the dynamic game is a timing equilibrium, resulting in fewer battles than in the equilibrium of the static game. Although we observe deviations from the timing equilibrium (in particular, some subjects move at the beginning of the game), we do find that the number of battles is reduced compared to the static case. The observed lower frequency of battles for $k = 12$ is not expected if players are homogeneous in their risk aversion, but it is consistent with the comparative static prediction of our version of the model in which players differ in their degree of risk aversion.[8] In contrast to the theoretical predictions, we also observe a decrease in battles when the deterrence value is negative. This result is, however, partly mechanical; even if all subjects wanted to act immediately, some subjects might be a fraction of a second slower than others, resulting in more escapes.[9]

It is also a possibility that random noise reduces the frequency of battles in the dynamic game. For instance, if players in the dynamic game choose fight, flee, and wait in each period with equal probabilities while players in the static game choose between flee and fight with equal probabilities, fewer battles will be observed in the static game.[10] However, as we will illustrate in section IVC, the behavior of our subjects is very remote from this random benchmark. Our subjects respond in a sensible way to their private strength parameters. Moreover, in agreement with theory but in contrast to the random benchmark, we find that the dynamic nature matters most for reducing the frequency of battles when $k > 0$.

Also following proposition 2, we expect that players sort themselves according to their fighting ability in case of a timing equilibrium. The strongest players should wait longer than weaker players, giving weaker players the opportunity to escape. Hence, weaker players should manage to escape more frequently in the dynamic games than the static games if the deterrence value is positive. Our results are in line with this prediction. The right panel of figure 4 shows how often the weaker subject in a pair escapes. Subjects sort on fighting ability more often in the dynamic than the static game and the increase is larger for dynamic games with a positive deterrence value. For $k = -6$, the weaker player escapes in 12% of the matches in the static game and 26% of the matches in the dynamic game. For $k = 6$ ($k = 12$), the weaker player escapes in 15% (18%) of the matches in the static game and 38% (45%) of the matches in the dynamic game. The difference-in-difference analysis reported in table A1 in appendix C shows that the larger increase for positive deterrence values is statistically significant.[11]

Result 2. There are fewer battles in the dynamic game than in the static game. The dynamic version of the game helps players to sort themselves according to their fighting ability, and this effect is stronger when the deterrence value is positive.

The reduced number of battles in the dynamic games also positively affects earnings. Figure 5 shows the mean earnings for each treatment and for different levels of fighting ability. As expected, stronger types attain higher earnings. Averaging across all fighting abilities, earnings are higher in the dynamic games than in the static games (Mann-Whitney tests, $p < 0.003$ for all three comparisons). Note that the difference for $k = -6$ is much smaller than the differences for the treatments with a positive deterrence value. Moreover, for

---

[8]As for $k = 6$, we also observe deviations from a timing equilibrium when $k = 12$ as a number of subjects move early in the game. We discuss these deviations in more detail in section IVC.

[9]Of the 15 percentage point difference in battles between static and dynamic games when $k = -6$, 6 percentage points can be attributed to escapes that occur just because the subject who wanted to fight is a fraction slower than the subject who wanted to flee. The remaining 9 percentage points can be attributed to more subjects fighting in the static games.

[10]We thank a referee for this insight.

[11]Figure A2 in appendix C shows decision times for weak and strong players separately. It confirms the comparative static prediction that stronger subjects wait longer than weaker subjects if the deterrence value is positive. Moreover, with experience, both weak and strong players learn to wait longer.

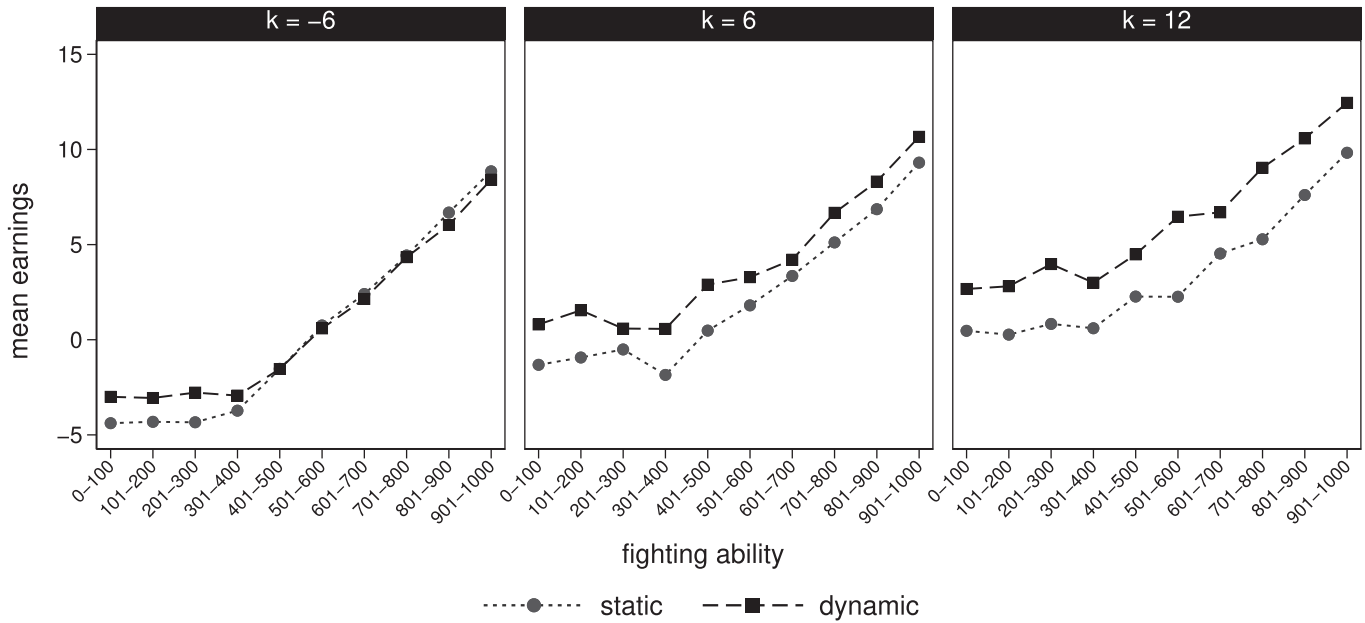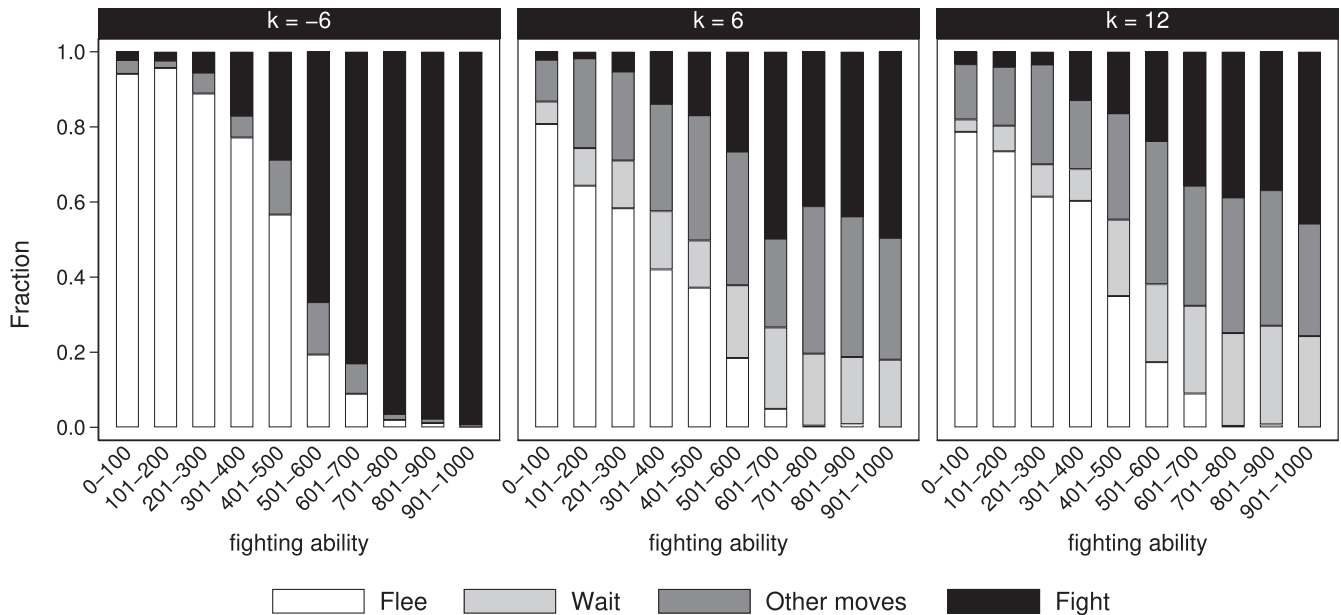FIGURE 5.—MEAN EARNINGS BY TREATMENT AND FIGHTING ABILITY

FIGURE 6.—BEHAVIOR BEFORE THE FINAL PERIOD IN THE DYNAMIC GAME, BY DETERRENCE VALUE k AND FIGHTING ABILITY a (IN TEN BINS OF EQUAL SIZE)

In the category "wait" refers to subjects who made it to the endgame, and in "other moves" refers to subjects who did not make a move before the endgame but the other subject did.

$k = -6$, the difference is driven by weaker subjects, whereas for the $k > 0$ treatments, all types on average benefit from endogenous timing.

C.   Individual Behavior

We start this section by considering how actions in the dynamic games depend on fighting ability. Figure 6 plots the fraction of subjects who flee or fight before the endgame, those who were waiting while the other moved, and those who wait until the endgame. We show this for the different deterrence values and for different fighting ability levels (in ten bins of equal size). In line with the results on decision times discussed in section IVA, no subject waits until the final period when the deterrence value is negative. Only a few subjects (6%) are still waiting when the other moves. When the deterrence value is positive, many subjects wait until the endgame or are waiting when the other moves. Combining

TABLE 2.—FLEEING BEFORE ENDGAME

	(1) $k = -6$	(2) $k = 6$	(3) $k = 12$	(4) All k
Risk aversion	0.037***	−0.001	0.041**	0.026***
	(0.007)	(0.018)	(0.018)	(0.009)
Female	−0.025	−0.012	0.003	−0.008
	(0.031)	(0.052)	(0.060)	(0.030)
Dominance	0.007	−0.028	−0.001	−0.006
	(0.011)	(0.023)	(0.015)	(0.010)
Physical strength	−0.015	−0.004	0.003	−0.002
	(0.019)	(0.020)	(0.020)	(0.011)
Fighting ability	−0.980***	−0.881***	−0.889***	−0.935***
	(0.003)	(0.034)	(0.016)	(0.013)
Round	0.000	−0.000	−0.000	−0.000
	(0.001)	(0.001)	(0.001)	(0.000)
$k = 6$				−0.127***
				(0.030)
$k = 12$				−0.112***
				(0.022)
Observations	2,520	2,080	2,520	7,120

Panel data probit regressions, with random effects at the subject level. Coefficients are average marginal effects. Dependent variable is a dummy indicating whether the player decided to flee before the endgame or not. Risk aversion is measured as the number of rejected lotteries. Dominance and physical strength are normalized (mean 0 and SD 1). Fighting ability takes on values between 0 and 1. Standard errors (clustered at the matching group level) in parentheses. Additional specifications with fewer or more controls are reported in table A4 in appendix C. *$p < 0.10$, **$p < 0.05$, and ***$p < 0.01$.

those groups, we find that 44% of subjects (intend to) wait for both $k = 6$ and $k = 12$. In line with theory, we find in all treatments that weaker players are much more likely to flee and stronger players are much more likely to wait or fight. This pattern clearly shows that the behavior of subjects is far from a random benchmark.

In the appendix, we provide further details on individual strategies. In appendix D, we estimate individual cutoff strategies. We find that most behavior is consistent with the use of cutoff strategies: around 90% of all decisions are captured by individual cutoff strategies. There is substantial heterogeneity in the type of cutoff strategies that individuals employ. Although the estimated cutoffs organize the data very well, for a substantial number of subjects, the estimated cutoffs are remote from the theoretical prediction.

In section IVB, we reported that sorting was observed not only for $k = 6$ but also for $k = 12$. Although behavior in both treatments does not exactly follow the predictions from the timing equilibrium (notably, some subjects move early on in the game), the finding that subjects sort in $k = 12$ is consistent with the idea that heterogeneous risk aversion enlarges the set of environments for which the timing equilibrium applies. A more direct implication of heterogeneous risk aversion is that the more risk-averse players should flee early more often. Table 2 presents panel data probit regressions of how the probability of choosing to flee before the endgame (T) depends on a subject's level of risk aversion, together with some controls. In agreement with the model of heterogeneous risk aversion, more risk-averse subjects are more likely to flee before the endgame when $k = -6$ and when $k = 12$, and the effect survives when we combine all three treatments.[12]

An anomaly is the fighting behavior early on in the game when there are benefits of letting the other escape, that is, when $k > 0$. In this case, fighting early is weakly dominated. Given the observed actions in the experiment, the losses of fighting early are substantial. Consider the strongest possible type who wins every fight. This type would earn 14% higher expected payoffs by waiting to fight in the endgame if $k = 6$ and 42% higher expected payoffs if $k = 12$. Note that fighting early is even more costly for weaker types. One possible reason for why we observe this anomalous behavior is that subjects may need some time to learn. As figure 3 shows, we do indeed observe less of this behavior in the final twenty rounds compared to the first twenty rounds. Another, more psychological, explanation for fighting early on in the game might be a preference for social dominance. The evidence does not support this. Table 3 shows that the survey measure of social dominance is not a predictor of fighting early. We also do not find an association with physical strength, but we do find that women are more likely to fight early than men.[13]

It may be that some of our subjects start playing the game with a misguided behavioral rule that in contests, it generally pays off to strike first. Myerson (1991) proposes that behavior that is apparently suboptimal behavior can sometimes be understood by assuming that observed behavior is optimal in a related but more familiar environment, which he calls a "salient perturbation" (see Myerson, 1991; Samuelson, 2001; Jehiel, 2005). Alternatively, it could be that intuition favors fighting behavior. According to the social heuristics hypothesis (Rand, Greene, & Nowak, 2012; Rand et al., 2014) applied to our setting, if fighting is typically advantageous, it could become the intuitive response. Note that subjects who fight

[12]When we regress the estimated cutoff fighting ability below which subjects flee before the endgame on risk aversion and other individual characteristics, we obtain qualitatively similar results. The regressions are reported in table A2 in appendix C.

[13]When we regress the estimated cutoff fighting ability above which subjects fight before the endgame on risk aversion and other individual characteristics, we obtain qualitatively similar results. The regressions are reported in table A3 in appendix C.

TABLE 3.—FIGHTING IN THE FIRST SECOND

	All Rounds			Final 20 Rounds		
	(1) $k = 6$	(2) $k = 12$	(3) all $k > 0$	(4) $k = 6$	(5) $k = 12$	(6) all $k > 0$
Risk aversion	-0.030^{**}	-0.013	-0.020^{**}	-0.020	-0.015	-0.018^{*}
	(0.015)	(0.018)	(0.010)	(0.017)	(0.013)	(0.011)
Female	0.130	0.029	0.066	0.161	0.046	0.097^{**}
	(0.100)	(0.058)	(0.050)	(0.111)	(0.049)	(0.048)
Dominance	0.002	0.018	0.015	-0.025	0.011	0.006
	(0.019)	(0.021)	(0.016)	(0.021)	(0.014)	(0.011)
Physical strength	0.009	0.001	-0.001	0.025	0.022	0.022
	(0.054)	(0.036)	(0.029)	(0.053)	(0.027)	(0.025)
Fighting ability	0.415^{***}	0.368^{***}	0.389^{***}	0.272^{***}	0.218^{**}	0.242^{***}
	(0.067)	(0.060)	(0.044)	(0.072)	(0.090)	(0.048)
Round	-0.004^{***}	-0.005^{***}	-0.005^{***}	-0.002	-0.003^{*}	-0.002^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
$k = 12$			-0.017			-0.026
			(0.042)			(0.036)
Observations	2,080	2,520	4,600	1,040	1,260	2,300

Panel data prohibit regression with random effects at the subject level. Coefficients are average marginal effects. Dependent variable is a dummy indicating whether the player decided to fight in the first second or not. Risk aversion is measured as the number of rejected lotteries. Dominance and physical strength are normalized (mean 0 and SD 1). Fighting ability takes on values between 0 and 1. Standard errors (clustered at the matching group level) in parentheses. Additional specifications with fewer or more controls are reported in table A5 in appendix C. * $p < 0.10$, ** $p < 0.05$, and *** $p < 0.01$.

early on have limited opportunities to learn, since they never experience the benefits of waiting. This could explain why they do not converge fully to waiting until the end of the game.

The fact that we observe an approximately equal frequency of early battles when $k = 6$ as when $k = 12$ suggests that this behavior is not due to a separate utility component reflecting (for instance) a desire to control the outcome or a joy of winning. If people have a preference to control the outcome, we would expect fewer early battles when it becomes more costly in $k = 12$.[14] Still, when play has not yet converged to equilibrium, we cannot exclude that early fighting is encouraged by players who experience a joy of winning when they beat the other in a battle. In our follow-up experiments reported in section IVD, we include some measures of joy of winning to get direct evidence for this possibility.[15]

Result 3. A sizable minority of players acts immediately when the deterrence value is positive. This behavior decreases with experience.

D.   Experiment 2

In the dynamic treatments, a period lasted 200 ms. Such short periods can make it hard for participants to precisely time their actions. This could potentially explain why even for $k = -6$, we observe fewer battles and more escapes in

the dynamic game compared to the static game. We address this in a follow-up experiment.[16]

Experimental design and procedures.   The design of experiment 2 closely follows that of the first experiment. We collected data for all dynamic treatments, using periods of 5 seconds instead of 200 ms and four periods per round (with forty rounds in total). This gives subjects more scope to time their actions. We also added two items to the survey, measuring subjects' joy of winning. The first (incentivized) measure is taken from Sheremeta (2010, 2018). In this task, subjects can bid to win a contest with a prize of 0 points. For the second (nonincentivized) measure, subjects indicated how strongly they agreed with the statement: "I enjoy winning an amount by competing against another person more than I enjoy receiving that same amount without having to compete for it" (rated on a 7-point Likert scale).

The experiment was run online. Participants were recruited from the same subject pool as for the first experiment (excluding subjects who already participated). As in the first experiment, we included test questions at the end of the instructions. We showed the correct answers after two failed attempts on a question. We did this to prevent that subjects would log out if they had to wait too long. We kept track of the mistakes they made so that we can control for this in the analysis.

In total, 168 subjects participated, with seven matching groups of eight subjects in each of the three dynamic treatments ($k = -6$, $k = 6$, $k = 12$).[17] Sessions lasted around 60 minutes in total, and earnings varied between €4.20 and €35.70 (€19.00 on average).

---

[14] The same argumentation would apply to a distaste for surprise or suspense.

[15] Sheremeta (2010), Price and Sheremeta (2011), and Cason, Masters, and Sheremeta (2018) all report evidence that joy of winning and risk aversion are important factors in driving subjects' behavior in contest games. In a second price auction with value uncertainty, Goeree and Offerman (2003b) find that bidders tend to submit bids below the expected value of the object, which suggests that risk aversion may be the stronger force. Sheremeta (2013) provides a survey.
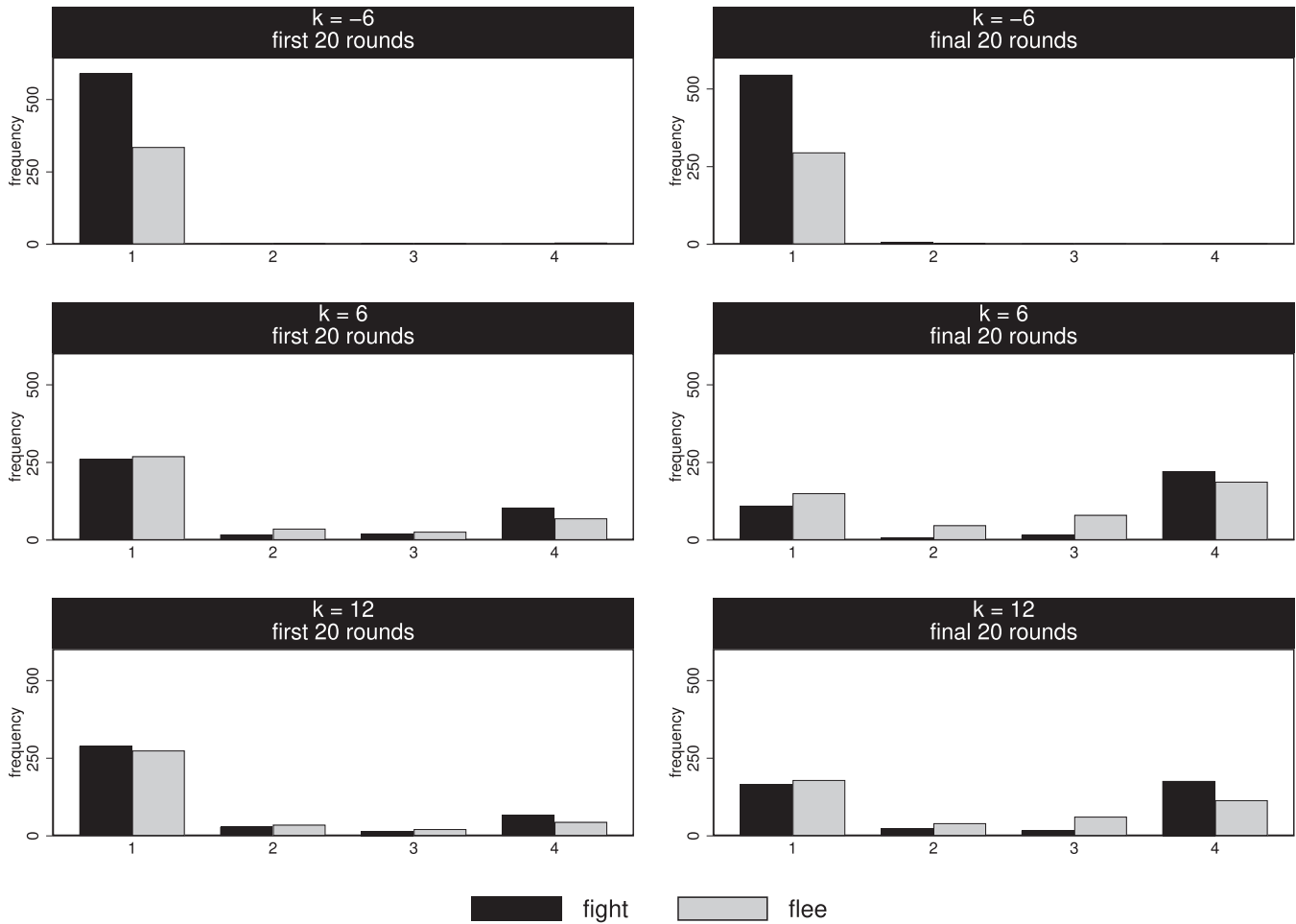
[16] We thank a referee for this suggestion.

[17] We have some missing data for five subjects who lost the connection. If a subject could not be paired in a round because of this, he or she received the maximal payoff.

FIGURE 7.—DISTRIBUTION OF DECISIONS OVER PERIODS BY DETERRENCE VALUE IN THE DYNAMIC GAME WITH FOUR PERIODS (EXPERIMENT 2)

Left: First twenty rounds. Right: Final twenty rounds. Only observations where a player made a decision to fight or flee are included (omitting observations where a player was waiting when the other moved).

Results. Figure 7 shows the timing of actions, which strongly resembles the results of experiment 1. With a negative deterrence value, virtually all action happens in the first period. With a positive deterrence value, many subjects wait until later periods. Compared to the first twenty rounds (left panels), more subjects wait in the final twenty rounds (right panels). The mean waiting time does not increase with experience for $k = -6$ and does increase for positive deterrence values (figure A3, appendix C). Moreover, we again observe that some participants act in the first period when $k > 0$. In line with the theoretical predictions, some subjects flee just before the endgame, although some do so in period 2 rather than period 3, and very few fight just before the endgame.

Figure 4 plots the frequency of battles and escapes by the weaker player in both experiments. The results are very comparable to those of experiment 1. In particular, for positive values of k, the dynamic game leads to a reduction in battles and an increase in escapes by the weaker player compared to the static version. The difference between the static and dynamic game is significant in all those cases (Mann-Whitney test, $p < 0.005$ in all cases). The main difference with ex-

periment 1 is that for a negative k, there is no reduction in battles or increase in escapes compared to the static games ($p = 0.898$ for battles, $p = 0.368$ for escapes). This supports the idea that in experiment 1, the decrease in battles and increase in escapes are driven by coordination failures: subjects may have attempted to immediately fight but were not always able to precisely time their action.[18]

In experiment 2, we again observe anomalous early fighting if $k > 0$. The two measures of joy of winning do not explain this early fighting, while the number of mistakes in the test questions and the social dominance score do explain (some) of the anomalous behavior (see table A7 in appendix C). In experiment 2, we do not replicate the finding that risk aversion correlates with fleeing before the endgame (see table A8 in appendix C).[19]

---

[18] A regression analysis confirms these results. The interaction effects between dynamic timing and positive deterrence values are statistically significant, indicating that the effect of dynamic timing on battles and escapes matters more for $k > 0$ (table A6 in appendix C).

[19] If we combine the data of both experiments, risk aversion is significantly correlated with fleeing before the endgame and dominance with early fighting. See tables A9 and A10 in appendix C.

V.   Conclusion

In this paper, we present a dynamic fight-or-flight game that makes sense of a large range of conflicts observed in practice. We highlight the crucial role that the deterrence value plays that players receive when the other player successfully escapes. If it is negative, players will act in a split-second. When it is positive, players will be patient and try to make the other player flee. An interesting feature of the analysis is that if the deterrence value is positive but not too large, sorting will occur. That is, the weakest players will flee just before the end, and thereby avoid costly battles. Thus, this paper clarifies how time can help people reach better outcomes in dynamic games, even when time is not costly. The important role of the deterrence value is confirmed in our experiments. Compared to a static version of the game, players are better able to avoid costly battles.

In the experiment, we find support for a behavioral version of the model that allows for heterogeneous risk aversion. In agreement with this model, sorting occurs for a wider range of situations than predicted by the model with standard preferences. In addition, subjects who appear to be more risk averse in an independent task tend to be the ones who more frequently flee early, although we do not replicate this in the follow-up experiment. We also observe an interesting anomaly. A fraction of the players choose to fight early even in situations where the strategic incentive is to be patient. Our conjecture is that some subjects come to the interaction with a homegrown notion that it generally pays off to strike early in contests. Over time, this costly behavior diminishes but does not disappear.

We think that our setup provides a lower limit of the amount of sorting that can be expected in practice. In our game, players manage to sort even though they do not receive any sensory input about the ability of the opponent. In particular when there is a strategic incentive to wait, sensory cues before or during the contest may help players to avoid costly fights. In an actual display, body odor or a high-pitched voice may reveal fear and help identify the weaker player (Mujica-Parodi et al., 2009; Sobin & Alpert, 1999). A dominant performance in a television show by a candidate running for presidential office may convince a weaker opponent that it is better to flee early. In the future, artificial intelligence may further help players to agree on how they are ranked in terms of ability before they engage in a costly battle. Relevant information about the opponent's ability will also affect players' decisions when the deterrence value is negative. However, in such situations, a positive frequency of battles cannot be avoided. Even when information about the opponent helps players to perfectly forecast who will win the fight, the stronger player will still want to catch the weaker player in a battle. We think that extending the analysis in this direction is an interesting avenue for future research.

REFERENCES

Agranov, Marina, and Matt Elliott, "Commitment and (In)efficiency: A Bargaining Experiment," Cambridge-INET working paper 2017/20 (2017).

Bigoni, Maria, Marco Casari, Andrzej Skrzypacz, and Giancarlo Spagnolo, "Time Horizon and Cooperation in Continuous Time," *Econometrica* 83:2 (2015), 587–616. 10.3982/ECTA11380

Bliss, Christopher, and Barry Nalebuff, "Dragon-Slaying and Ballroom Dancing: The Private Supply of a Public Good," *Journal of Public Economics* 25:1–2 (1984), 1–12. 10.1016/0047-2727(84)90041-0

Calford, Evan, and Ryan Oprea, "Continuity, Inertia, and Strategic Uncertainty: A Test of the Theory of Continuous Time Games," *Econometrica* 85:3 (2017), 915–935. 10.3982/ECTA14346

Carrillo, Juan D., and Thomas R. Palfrey, "The Compromise Game: Two-Sided Adverse Selection in the Laboratory," *American Economic Journal: Microeconomics* 1:1 (2009), 151–181. 10.1257/mic.1.1.151

Cason, Timothy N., William A. Masters, and Roman M. Sheremeta, "Winner-Take-All and Proportional-Prize Contests: Theory and Experimental Results," *Journal of Economic Behavior and Organization* 175 (2018).

Cheng, Joey T., Jessica L. Tracy, and Joseph Henrich, "Pride, Personality, and the Evolutionary Foundations of Human Social Status," *Evolution and Human Behavior* 31:5 (2010), 334–347. 10.1016/j.evolhumbehav.2010.02.004

De Dreu, Carsten K. W., Jörg Gross, Zsombor Méder, Michael Giffin, Eliska Prochazkova, Jonathan Krikeb, and Simon Columbus, "In-Group Defense, Out-Group Aggression, and Coordination Failures in Intergroup Conflict," *Proceedings of the National Academy of Sciences* 113:38 (2016), 10524–10529. 10.1073/pnas.1605115113

Dechenaux, Emmanuel Dan Kovenock, and Roman M. Sheremeta, "A Survey of Experimental Research on Contests, All-Pay Auctions and Tournaments," *Experimental Economics* 18:4 (2015), 609–669. 10.1007/s10683-014-9421-0

Friedman, Daniel, and Ryan Oprea, "A Continuous Dilemma," *American Economic Review* 102:1 (2012), 337–363. 10.1257/aer.102.1.337

Fudenberg, Drew, and Jean Tirole, "A Theory of Exit in Duopoly," *Econometrica* 54 (1986), 943–960. 10.2307/1912845

Gächter, Simon, Eric J. Johnson, and Andreas Herrmann, "Individual-Level Loss Aversion in Riskless and Risky Choices," CEDEX discussion paper 1749–3293 (2007).

Goeree, Jacob K., and Theo Offerman, "Competitive Bidding in Auctions with Private and Common Values," *Economic Journal* 113:489 (2003a), 598–613. 10.1111/1468-0297.t01-1-00142

——— "Winner's Curse without Overbidding," *European Economic Review* 47:4 (2003b), 625–644. 10.1016/S0014-2921(02)00290-8

Hörisch, Hannah, and Oliver Kirchkamp, "Less Fighting Than Expected," *Public Choice* 144:1–2 (2010), 347–367.

Ivanov, Asen, Dan Levin, and James Peck, "Hindsight, Foresight, and Insight: An Experimental Study of a Small-Market Investment Game with Common and Private Values," *American Economic Review* 99:4 (2009), 1484–1507. 10.1257/aer.99.4.1484

Jehiel, Philippe, "Analogy-Based Expectation Equilibrium," *Journal of Economic Theory* 123:2 (2005), 81–104. 10.1016/j.jet.2003.12.003

Kirchkamp, Oliver, and Benny Moldovanu, "An Experimental Analysis of Auctions with Interdependent Valuations," *Games and Economic Behavior* 1:48 (2004), 54–854. 10.1016/j.geb.2002.02.001

Kolb, Aaron M., "Optimal Entry Timing," *Journal of Economic Theory* 157 (2015), 973–1000. 10.1016/j.jet.2015.03.002

Levin, Dan, and James Peck, "Investment Dynamics with Common and Private Values," *Journal of Economic Theory* 143:1 (2008), 114–139. 10.1016/j.jet.2008.02.001

Maynard Smith, John, "The Theory of Games and the Evolution of Animal Conflicts," *Journal of Theoretical Biology* 47:1 (1974), 209–221. 10.1016/0022-5193(74)90110-6, PubMed: 4459582

Mujica-Parodi, Lilianne R., Helmut H. Strey, Blaise Frederick, Robert Savoy, David Cox, Yevgeny Botanov, Denis Tolkunov, Denis Rubin, and Jochen Weber, "Chemosensory Cues to Conspecific Emotional Stress Activate Amygdala in Humans," *PLOS One* 4:7 (2009), e6415. 10.1371/journal.pone.0006415, PubMed: 19641623

Myerson, Roger B., *Game Theory: Analysis of Conflict* Cambridge, MA: Harvard University Press 1991.

Ockenfels, Axel, and Alvin E. Roth, "Late and Multiple Bidding in Second Price Internet Auctions: Theory and Evidence Concerning Different Rules for Ending an Auction," *Games and Economic Behavior* 55:2 (2006), 297–320. 10.1016/j.geb.2005.02.010

Oprea, Ryan, Gary Charness, and Daniel Friedman, "Continuous Time and Communication in a Public-Goods Experiment," *Journal of*

Economic Behavior and Organization 108 (2014), 212–223. 10.1016/j.jebo.2014.09.012

Oprea, Ryan, Keith Henwood, and Daniel Friedman, "Separating the Hawks from the Doves: Evidence from Continuous Time Laboratory Games," *Journal of Economic Theory* 146:6 (2011), 2206–2225. 10.1016/j.jet.2011.10.014

Oprea, Ryan, Bart J. Wilson, and Arthur Zillante, "War of Attrition: Evidence from a Laboratory Experiment on Market Exit," *Economic Inquiry* 51:4 (2013), 2018–2027. 10.1111/ecin.12014

Ponsati, Clara, and József Sákovics, "The War of Attrition with Incomplete Information," *Mathematical Social Sciences* 29:3 (1995), 239–254. 10.1016/0165-4896(94)00771-Y

Potters, Jan, Martin Sefton, and Lise Vesterlund, "After You— Endogenous Sequencing in Voluntary Contribution Games," *Journal of Public Economics* 89:8 (2005), 1399–1419. 10.1016/j.jpubeco.2004.02.008

Price, Curtis R., and Roman M. Sheremeta, "Endowment Effects in Contests," *Economics Letters* 111:3 (2011), 217–219. 10.1016/j.econlet.2011.02.003

Rand, David G., Joshua D. Greene, and Martin A. Nowak, "Spontaneous Giving and Calculated Greed" *Nature* 489:7416 (2012), 427. 10.1038/nature11467, PubMed: 22996558

Rand, David G., Alexander Peysakhovich, Gordon T. Kraft-Todd, George E. Newman, Owen Wurzbacher, Martin A. Nowak, and Joshua D. Greene, "Social Heuristics Shape Intuitive Co-Operation," *Nature Communications* 5:3677 (2014). 10.1038/ncomms4677

Roth, Alvin E., J. Keith Murnighan, and Françoise Schoumaker, "The Deadline Effect in Bargaining: Some Experimental Evidence," *American Economic Review* 78:4 (1988), 806–823.

Roth, Alvin E., and Axel Ockenfels, "Last-Minute Bidding and the Rules for Ending Second-Price Auctions: Evidence from eBay and Amazon Auctions on the Internet," *American Economic Review* 92:4 (2002), 1093–1103. 10.1257/00028280260344632

Samuelson, Larry, "Analogies, Adaptation, and Anomalies," *Journal of Economic Theory* 97:2 (2001), 320–366. 10.1006/jeth.2000.2754

Sell, Aaron, Leda Cosmides, John Tooby, Daniel Sznycer, Christopher Von Rueden, and Michael Gurven, "Human Adaptations for the Visual Assessment of Strength and Fighting Ability from the Body and Face," *Proceedings of the Royal Society B: Biological Sciences* 276:1656 (2009), 575–584. 10.1098/rspb.2008.1177

Sheremeta, Roman M., "Experimental Comparison of Multi-Stage and One-Stage Contests," *Games and Economic Behavior* 68:2 (2010), 731–747. 10.1016/j.geb.2009.08.001

——— "Overbidding and Heterogeneous Behavior in Contest Experiments," *Journal of Economic Surveys* 27:3 (2013), 491–514. 10.1111/joes.12022

——— "Impulsive Behavior in Competition: Testing Theories of Overbidding in Rent-Seeking Contests, " SSRN 2676419 (2018).

Sobin, Christina, and Murray Alpert, "Emotion in Speech: The Acoustic Attributes of Fear, Anger, Sadness, and Joy," *Journal of Psycholinguistic Research* 28:4 (1999), 347–365. 10.1023/A:1023237014909, PubMed: 10380660

Song, Jian, and Daniel Houser, "Costly Waiting in Dynamic Contests: Theory and Experiment," Working Paper 1082, George Mason University, Interdisciplinary Center for Economic Science (2021).

Weesie, Jeroen, "Asymmetry and Timing in the Volunteer's Dilemma," *Journal of Conflict Resolution* 37:3 (1993), 569–590. 10.1177/0022002793037003008