# Costly Waiting in Dynamic Contests: Theory and Experiment

Daniel Houser and Jian Song

August 2021

Discussion Paper

Interdisciplinary Center for Economic Science
4400 University Drive, MSN 1B2, Fairfax, VA 22030
Tel: +1-703-993-4850    Fax: +1-703-993-4851
ICES Website: http://ices.gmu.edu
ICES RePEc Archive Online at: http://edirc.repec.org/data/icgmuus.html

# Costly Waiting in Dynamic Contests: Theory and Experiment*

Jian Song† Daniel Houser‡

August, 2021

**Abstract**

We extend the *war of attrition* by studying a three-period dynamic contest game. In our game, players can fight against their opponents at certain period of the contest and can flee at any time. Waiting is costly. We focus on the role of waiting costs and show that the value of waiting costs is a key factor in determining the type of equilibrium in such dynamic contests. Specifically, as waiting costs increase, contests end earlier, battles are less likely to occur, and the weaker player in a pair is more likely to flee. A lab experiment verifies most key features of our model. However, unlike theoretical predictions, we find that as waiting costs increase, the duration of contests and the frequency of battles fail to drop as significantly as theory predicted. Moreover, we find that in each treatment, individual players exit the contest significantly earlier than predicted.

**Keywords:** dynamic contest, waiting cost, frequency of battles, lab experiment

**JEL Codes:** D82, D90, C90

---

*We thank Jeroen van de Ven, Cesar Martinelli, Alexander Tabarrok and participants at the 2020 Utah Experimental Economics Conference, Contests: Theory and Evidence 6th Conference, 2020 ESA World Conference and SEA 90th Annual Meeting for helpful comments. This research was financed by the Interdisciplinary Center for Economic Science at George Mason University, and the Humane Studies Fellowship. Declarations of interest: none.

†Corresponding author: Department of Economics and Interdisciplinary Center for Economic Science, George Mason University and Economic Science Institute, Chapman University. Email: jsong25@gmu.edu

‡Department of Economics and Interdisciplinary Center for Economic Science, George Mason University. Email: dhouser@gmu.edu

# 1 Introduction

Dynamic contests are ubiquitous in social interactions. Examples include political campaigns, military wars, strikes and R&D races. Following Smith (1974), dynamic contests are often modeled using the *war of attrition*. In the original model, players compete in pairs and choose to exit the contest according to their strengths. The player who waits the longest to exit the contest wins. Players incur costs proportional to the time they wait to exit, i.e., their "wait time". The contest ends as soon as one player exits the game. As a result, the game always ends with "escape".

The real world differs from the models in that "escape" is not the only possible outcome in naturally occurring conflicts. Sometimes, players can actively fight with their opponents. In such cases, the contest ends in "battle". One example is the 2019 General Motors strike: after lengthy negotiations, the United Automobile Workers decided to strike for better working conditions.

To fill the gap between theoretical predictions and real-world conflicts, inspired by earlier closely related work by van Leeuwen *et al.* (2020), we extend the *war of attrition* by allowing players to choose to battle one another. In our three-period dynamic contest game, players must decide whether to fight, flee or wait, at different periods of the contest. Players can earn a prize, plus a positive "deterrence value" if their opponent flees before they do. In contrast to the model in van Leeuwen *et al.* (2020), we assume waiting is costly: the longer a player waits, the greater cost they incur.[1] We pay special attention to the role of waiting costs in dynamic contests. Moreover, our setting restricts players by allowing them to fight only in the final period. As in most natural environments, this allows opponents time to escape.[2] Another reason we only allow players to fight in the final period is to emphasize the effects of waiting costs (as battle cost rises with wait time).

Theoretically, we prove that the type of pure-strategy Bayesian Nash equilibrium is determined by the ratio of waiting costs to deterrence value. When waiting costs are relatively low in relation to deterrence value, all players will wait until the final period to decide whether to fight or flee. When waiting costs are relatively high in relation to deterrence value, players will flee gradually and are less likely to fight. Our model implies that greater waiting costs

[1]The role of waiting costs and deterrence values are different. Theoretically, when the waiting cost equals zero and the deterrence value is positive, sufficiently strong players would wait until the last period and fight against their opponents; while the waiting cost is positive and the deterrence value is zero, players would flee gradually and no battles occur.

[2]Moreover, fighting typically occurs only after players have first considered whether it is beneficial to concede. In some sense, our model simply separates the flee decision from the fight decision. It treats sequentially what is usually treated as a simultaneous decision, but it seems quite natural for the flee decision to precede the fight decision. One example captures such situations is, countries negotiate for multiple rounds about the development of nuclear weapons before one side actively declares a war or surrenders.

reduce the frequency of battles, shorten the duration of the dynamic contests, and enable the weaker player in a pair more likely to flee. Our model explains why countries and institutions find sanctions an effective tool for forcing their political opponents to surrender. The reason is that once sanctions are imposed, the waiting costs of conflicts rise significantly.[3]

Economic theory predicts that greater waiting costs should reduce the frequency of battles and accelerate the duration of contests. In practice, however, significant waiting costs may be counterproductive, particularly for lowering the battle rate. Previous studies suggest that due to the "sunk cost effect",[4] players tend to continue an endeavor once an investment has been made (e.g., Arkes & Hutzel (2000); Friedman *et al.* (2007)). Therefore, we would expect that players in our dynamic contests who have already paid a nonnegligible waiting cost might choose to stay for a longer period. Moreover, within each treatment, players may fail to make rational decisions based on their strength levels, making theoretical predictions less credible. This could be related to loss aversion (van Leeuwen *et al.* (2020)) or pro-social orientation (Abbink *et al.* (2012), Song & Houser (2021)). Taken together, practical results may be influenced by behavioral factors not easily predicted by the theory.

We test our theoretical predictions in an experiment with three treatments that differ according to waiting costs. Our results are broadly consistent with our theory. Our data demonstrate that, overall, as waiting costs increase, battles become less likely to occur. Likewise, contests end more quickly and the weaker player in a pair becomes more likely to flee. However, when waiting costs are initially low, an increase in waiting costs significantly reduces both the duration of contests and the frequency of battles. As waiting costs increase further, there is no effect on the duration of contests or the frequency of battles. And for each treatment, contests end earlier than prediction.

We scrutinize each player's behavior to investigate those deviations in greater details. One interesting finding contradicts the predictions: as waiting costs increase, players become less likely to flee gradually. Instead, their behavior tends to "bifurcate": they either flee immediately or fight in the last period, such behavior is consistent with the sunk cost effect. Additionally, similar to the behavioral patterns in van Leeuwen *et al.* (2020), we find that a proportion of players flee immediately in all treatments, even though they possess sufficient strength to continue. Consequently, contests end earlier than predicted in all treatments.

Our work contributes to the theory literature on dynamic contest games. An early contribution by Fudenberg *et al.* (1983) investigated the effect of players' strengths on results in the *war of attrition*. Bulow & Klemperer (1999) investigated the *war of attrition* with

[3]Another example demonstrating this fact occurred during WWII, when Japan surrendered quickly after being attacked with atomic bombs. The reason is that waiting longer would have raised the chance of suffering even more serious disasters.

[4]See Roth *et al.* (2015) as a review of the sunk cost effect in economic decision-making.

multiple players. Hörner & Sahuguet (2011) generalized the *war of attrition* by allowing players to signal. Rohner *et al.* (2013), Acemoglu & Wolitzky (2014), Baliga *et al.* (2011), and Gul & Pesendorfer (2012) modified the *war of attrition* to model inter-group conflicts. Our model, which is based on van Leeuwen *et al.* (2020), allows players to actively fight against their opponent in the last period of *war of attrition*. As a result, we can investigate how waiting costs affect equilibrium types.

Our work also contributes to the experimental contest literature. It is difficult to test dynamic contest models with naturally occurring data due to "self-selection, unobservable variables and unavoidable endogeneity in dynamic settings" (Kimbrough *et al.* (2017)). As a result, researchers have conducted experiments to verify a variety of implications of theoretical models.[5] Our experimental results, on the one hand, verify that waiting costs play an important role in altering the frequency of battles, the duration of contests, and the fraction of weaker players who choose to flee in dynamic contests. On the other hand, consistent with previous findings, our data suggest that there are behavioral deviations from theoretical predictions.

This paper proceeds as follows: Section 2 describes the model. Section 3 details the experiment design and predictions. Section 4 reports experimental results. Section 5 offers concluding remarks.

# 2 Model

## 2.1 Basic set up

Following van Leeuwen *et al.* (2020), we assume time is discrete, with finite periods $t = 0, 1, ...T$. For simplicity, in our model, let $T = 2$. Two players independently make their decision in each period. In $t = 0$ and 1, they players can only choose whether to "flee" (R)[6] or "wait" (W). In the last period, where $t = 2$, they must choose whether to "flee" (R) or "fight" (F), if no one has conceded so far.[7]

[5]For instance, Mago & Sheremeta (2019) experimentally compared dynamic contest games with static games. Oprea *et al.* (2013) investigated the impact of heterogenous players on war of attrition. Deck & Sheremeta (2012) investigated the effect of timing on conflict intensity. Deck & Jahedi (2015) studied bidders' behavior in different period via dynamic all-pay auction games. All the above papers reported overbidding in relation to theoretical predictions. By contrast, Hörisch & Kirchkamp (2010) found systematic underbidding in dynamic contests. Another strand of dynamic contest experiments focuses on identifying individual player behavior. Gelder & Kovenock (2017) classified players of a dynamic contest into various types according to their realized strategies.

[6]Same as van Leeuwen *et al.* (2020), here we use "R" to represent "retreat".

[7]Technically, if we relax this assumption and allow players to fight against their opponents at any period, it can be shown that when the waiting cost is sufficiently large in relation to the deterrence value, strong

Before the contest starts, each player privately knows their strength, $a_i$. It is common knowledge that $a_i \sim U(0,1)$. A player's strategy can be described as $S(a_i) = (t, A)$, where $A \in \{R, W\}$ for $t = 0, 1$ and $A \in \{R, F\}$ for $t = 2$. This represents player $i$ with strength $a_i$, who chooses their strategy A at period t, before the contest ends.

The contest ends as soon as (at least) one player flees in $t = 0$ and 1, or right after both players make their decision in $t = 2$. The outcome of the contest can be "battle" or "escape". A battle occurs if both players wait until the last period and choose to "fight". An escape occurs if one player chooses to flee at $t = 0$ or $t = 1$, while the other chooses to wait, or if both choose to flee in the same period. If one player chooses to fight and the other chooses to flee at $t = 2$, an escape occurs with probability $\frac{1}{2}$, and a battle occurs with probability $\frac{1}{2}$.

If battle occurs, payoffs are decided according to players' strengths: the player with greater strength wins a prize, $v_H > 0$, and the player with lesser strength loses $v_L > 0$. If escape occurs, the player who chooses to flee earns 0 and the player who chooses not to flee earns $v_H + k$, where $k$ represents the deterrence value. We restrict our analysis to $k > 0$, to represent the case where a player prefers to win the contest without costly battles. Waiting is costly, and a player incurs an identical waiting cost $c > 0$ per period.

As for the tie-breaking rules, if there is a battle between equally strong players, it is randomly determined which player wins $v_H$ and which player loses $v_L$. If both players flee in the same period, they share the prize for winning without a battle, the payoffs for each player are $\frac{v_H+k}{2}$.

We first assume that the degree of loss aversion ($\lambda$) to all players equals 1,[8] which means their utility function is defined as:

$$U(x) = x \tag{1}$$

Players' goal is to maximize their expected utility.

## 2.2 Equilibria

We look for pure-strategy Bayesian Nash equilibria. As in van Leeuwen *et al.* (2020), we derive equilibria under the assumption that players utilize threshold strategy: types[9] below a certain threshold flee and types above that threshold fight. Intuitively, given that the deterrence value, $k$, is positive, the sufficiently strong types should wait until the last period

types have the incentive to fight earlier and the gradual equilibrium of Proposition 3 does not exist. See Appendix F for details.

[8]As stated in Dutcher *et al.* (2015), considering loss-averse players implies that the experiment cannot be calibrated ex ante. We therefore simplify the theoretical model by assuming the degree of loss aversion to all players equals 1 and control for loss aversion econometrically. In Appendix D, we demonstrate how loss aversion parameters affect equilibrium predictions.

[9]The type of a player is characterized by their strength, $a_i$.

and fight, giving their opponent a chance to flee while gaining deterrence benefits. Given the strong types' strategy, it is not optimal for weak types to act later than strong types, as the weak types not only need to pay extra waiting costs, but also need to pay the costs of losing the battles.

We first consider the case where the waiting cost, $c$, is relatively small in relation to the deterrence value, $k$. Intuitively, when $k$ is sufficiently greater than $c$, to obtain the benefit of winning without costly battle, all types would wait until the last period ($t = 2$), while weaker types would "flee" to avoid further losses of battles, and stronger types would "fight" to win the prize. That result is consistent with the "waiting equilibrium" as stated in van Leeuwen *et al.* (2020).

The "waiting equilibrium" in our model can be characterized by one threshold value $\hat{a}$: in equilibrium, types below $\hat{a}$ flee at $t = 2$ and types above $\hat{a}$ fight at $t = 2$, type $\hat{a}$ is indifferent between fleeing at $t = 2$ and fighting at $t = 2$.

To obtain the value of $\hat{a}$, first consider that if type $\hat{a}$ chooses to flee at $t = 2$, they must pay the waiting cost, which equals $2c$. The probability for them to meet an opponent who also flees at $t = 2$ is $\hat{a}$. The contest then ends with an escape and the payoffs for type $\hat{a}$ equal to $\frac{v_H+k}{2}$. The probability for type $\hat{a}$ to meet an opponent who fights at $t = 2$ is $(1 - \hat{a})$. The contest then has a 50% chance of ending with a battle and another 50% chance of ending with an escape. The corresponding payoffs for type $\hat{a}$ are $-\frac{v_L}{2}$. Thus, the expected utility for type $\hat{a}$ fleeing at $t = 2$ equals to:

$$\hat{a}(\frac{v_H + k}{2}) + (1 - \hat{a})(-\frac{v_L}{2}) - 2c \tag{2}$$

Similarly, if $\hat{a}$ chooses to fight at $t = 2$, they have a probability of $\hat{a}$ of meeting an opponent who flees at $t = 2$. In that case, the expected payoffs for type $\hat{a}$ are $\frac{1}{2}(v_H+k)+\frac{1}{2}(v_H) = v_H+\frac{k}{2}$. Type $\hat{a}$ has a chance of $(1 - \hat{a})$ of meeting an opponent who also fights at $t = 2$, and the payoffs for $\hat{a}$ now become $-v_L$. Thus, the expected utility for type $\hat{a}$ fighting at $t = 2$ equals to:

$$\hat{a}(v_H + \frac{k}{2}) + (1 - \hat{a})(-v_L) - 2c \tag{3}$$

Since type $\hat{a}$ is indifferent between fleeing and fighting at $t = 2$, the expected utility would be the same. Let (2) = (3). We obtain:

$$\hat{a} = \frac{v_L}{v_H + v_L} \tag{4}$$

In Appendix A we prove that when $c \leq \frac{kv_L}{4(v_H+v_L)}$, there exists a waiting equilibrium. That gives our first proposition:

**Proposition 1 (Waiting equilibrium)** *If $0 \leq c \leq \frac{kv_L}{4(v_H+v_L)}$, there exists a waiting equilibrium, in which types below $\hat{a}$ flee at t=2 and types above $\hat{a}$ fight at t=2. Type $\hat{a}$ is indifferent between fleeing and fighting at t=2.*

**Proof.** See Appendix A for the proofs. ■

If the waiting cost becomes relatively greater than the deterrence value, intuitively, since the waiting cost is not negligible, the deterrence value would not be attractive for weaker types. However, stronger types would still prefer to wait to gain deterrence benefits. Weaker types would flee immediately to avoid costly waiting, while stronger types would wait to gain deterrence benefits. We refer this kind of equilibrium as "jump equilibrium": in equilibrium, a fraction of types would flee at $t = 0$ and a fraction of types would flee at $t = 2$. The strongest types would fight at $t = 2$.

The jump equilibrium can be characterized by two threshold values: $0 \leq \overline{a_1} \leq \overline{a_2} \leq 1$. As described above, we assume that in equilibrium, a fraction of $\overline{a_1}$ flee at t=0; a fraction of $\overline{a_2} - \overline{a_1}$ flee at $t = 2$; and types above $\overline{a_2}$ fight at $t = 2$. Type $\overline{a_1}$ is indifferent between fleeing at $t = 0$ and fleeing at $t = 2$; type $\overline{a_2}$ is indifferent between fleeing at $t = 2$ and fighting at $t = 2$. We then derive the threshold values, using the method mentioned above, equal to:

$$\overline{a_1} = \frac{4c(v_H + v_L) - kv_L}{4c(v_H + v_L) + v_H(k + v_H + v_L)} \tag{5}$$

$$\overline{a_2} = \frac{4c(v_H + v_L) + v_H v_L}{4c(v_H + v_L) + v_H(k + v_H + v_L)} \tag{6}$$

When $\frac{kv_L}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2(v_H+v_L)}$, there exists a jump equilibrium. Unlike waiting equilibrium, the threshold values in jump equilibrium depend on both waiting cost and deterrence value. We pay special attention to how waiting costs affect threshold values. The results are stated as follows:

**Proposition 2 (Jump equilibrium)** *(1) If $\frac{kv_L}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2(v_H+v_L)}$, there exists a "jump equilibrium," in which a fraction of $\overline{a_1}$ flee at $t = 0$, a fraction of $\overline{a_2} - \overline{a_1}$ flee at $t = 2$, and types above $\overline{a_2}$ fight at $t = 2$.*
*(2) In jump equilibrium, as waiting cost increases, more types would flee at $t = 0$, less types would flee at $t = 2$ and fight at $t = 2$.*

**Proof.** See Appendix B for the proofs. ■

What if the waiting cost becomes even greater in relation to the deterrence value? Intuitively, a fraction of weaker types would flee immediately, or flee at $t = 1$ to avoid further losses on waiting. In contrast, stronger types would wait until the final period to gain deterrence benefits. Consequently, there is a fraction of types who would flee at the beginning ($t = 0$), a fraction of types who would flee at $t = 1$, and a fraction of types who would flee at $t = 2$. The strongest types would fight at $t = 2$. We refer to this type of equilibrium as "gradual equilibrium." The gradual equilibrium is characterized by three threshold values: $0 \leq a_1 \leq a_2 \leq a_3 \leq 1$. Without loss of generality, we suppose a fraction of $a_1$ flee at $t = 0$, a fraction of $a_2 - a_1$ flee at $t = 1$, a fraction of $a_3 - a_2$ flee at $t = 2$, and types above $a_3$ fight at $t = 2$. To solve those threshold values, we further assume that type $a_1$ is indifferent between fleeing at $t = 0$ and fleeing at $t = 1$; type $a_2$ is indifferent between fleeing at $t = 1$ and fleeing at $t = 2$; and type $a_3$ is indifferent between fleeing at $t = 2$ and fighting at $t = 2$. We obtain threshold values:

$$a_1 = \frac{4c^2(v_H + v_L) + kv_L(-2c + k + v_H)}{2cv_H(k + v_H + v_L) + (4c^2 + (k + v_H)^2)(v_H + v_L)} \tag{7}$$

$$a_2 = \frac{4c^2(v_H + v_L) + 2cv_H(k + v_H + v_L)}{2cv_H(k + v_H + v_L) + (4c^2 + (k + v_H)^2)(v_H + v_L)} \tag{8}$$

$$a_3 = \frac{4c^2(v_H + v_L) + 2cv_H(k + v_H + v_L) + v_L(k + v_H)^2}{2cv_H(k + v_H + v_L) + (4c^2 + (k + v_H)^2)(v_H + v_L)} \tag{9}$$

When $c \geq \frac{kv_L}{2(v_H+v_L)}$, a gradual equilibrium exists. In gradual equilibrium, the threshold values depend on both deterrence value and waiting costs. We obtain the following proposition:

**Proposition 3 (Gradual equilibrium)** *(1) If $c \geq \frac{kv_L}{2(v_H+v_L)}$, there exists a gradual equilibrium, in which a fraction of $a_1$ flee at $t = 0$ and a fraction of $a_2 - a_1$ flee at $t = 1$, a fraction of $a_3 - a_2$ flee at $t = 2$,and types above $a_3$ fight at $t = 2$.*
*(2) Suppose $v_H = v_L = v > 0$. In gradual equilibrium, as waiting costs increase, more types would flee at $t = 0$ and $t = 1$ and less types would flee at $t = 2$ and fight at $t = 2$.*

**Proof.** See Appendix C for the proofs. ■

From Propositions 1 to 3, we can infer that increasing waiting costs can effectively reduce the fraction of battles and shorten the duration of a contest. Moreover, as waiting costs increase, the weaker player in a pair would be more likely to choose to flee.

| Treatment | Waiting cost ($c$) | Deterrence value ($k$) | $v_H$ | $v_L$ |
|---|---|---|---|---|
| C5 | 5 | 100 | 100 | 100 |
| C20 | 20 | 100 | 100 | 100 |
| C40 | 40 | 100 | 100 | 100 |

Table 1: Treatment parameters

# 3 Experimental Design and Predictions

## 3.1 Experimental Design

The primary goal of our experiment is to investigate whether increasing waiting costs can effectively reduce the frequency of battles, shorten the duration of contests, and raise the fraction of weaker players who flee first. Therefore, our experiment has three treatments that vary according to waiting costs (referred as "C5","C20" and "C40" respectively). The parameters in each treatment are shown in Table 1.

The experiment proceeds as follows: At the beginning of each session, players are given the instructions of the experiment,[10] which are also read aloud by the experimenter. After players finish reading the instructions and completing the comprehensive quiz successfully, they proceed to a fight-or-flee game, which is the main part of our experiment.

The fight-or-flee game consists of 40 rounds. Players compete in a group of two. In each round, the groups are randomly re-matched. At the start of each round, each player is privately informed of their strength in this round. Strength is a random whole number between 0 and 1000. Every number has the same chance of being chosen. There are three periods in each round that represent the cases for $t = 0, 1, 2$. Each period lasts for fifteen seconds and players must make their decisions in each period, based on their strengths. Specifically, in $t = 0$ and $t = 1$, players must decide whether to "wait" or "flee," without knowing their opponents' choices. If time runs out and they have not made a decision, their choices are automatically counted as "wait"; In $t = 2$, they must decide whether to "fight" or "flee." Following van Leeuwen *et al.* (2020), if time runs out before they make their decisions then the players enter an "endgame" where they simultaneously choose whether to fight or flee. There is no time constraint in the endgame.

If both players flee in the same period, or one player flees while the other chooses to wait, then this round ends with "escape". If both players chose to fight at $t = 2$, then this round ends with "battle". If one player chooses to fight, while the other chooses to flee at $t = 2$, this round has a 50% chance of ending with "escape" and a 50% chance of ending with "battle".

[10]See Appendix E for the instructions.

The payoffs for players are defined as follows: If the round ends with an "escape", the player who chooses to flee earns 0 and the player who chooses not to flee earns $v_H + k$; if both players choose to flee in the same period, then each of them earn $\frac{v_H+k}{2}$. If the round ends with a "battle", the stronger player wins the battle and earns $v_H$, while the weaker player loses the battle and earns $-v_L$. Waiting is costly. If both players choose to wait at $t = 0$, then each of them must pay c for waiting; then, if both still choose to wait at $t = 1$, it costs each of them $2c$ in total for waiting.

At the end of each round, each player is informed of their payoffs, the outcome (escape/battle) of this round and their opponent's strength. All 40 rounds have the same procedure mentioned above.

After all players complete the fight-or-flee game, they are asked to complete a loss-aversion task (Gächter *et al.* (2007)), followed by a short demographic questionnaire. At the end of the experiment, players are paid in cash privately.

## 3.2 Predictions and Hypotheses

We use pure-strategy Bayesian Nash equilibrium as the prediction benchmark. Let the degree of loss aversion parameter $\lambda = 1$, and $v_H = v_L = k = 100$. Using the results stated in Section 2, when $0 \leq c \leq 12.5$, there exists a waiting equilibrium; when $12.5 \leq c \leq 25$, there exists a jump equilibrium; when $c \geq 25$, there exists a gradual equilibrium. The predictions are shown in Figure 1.

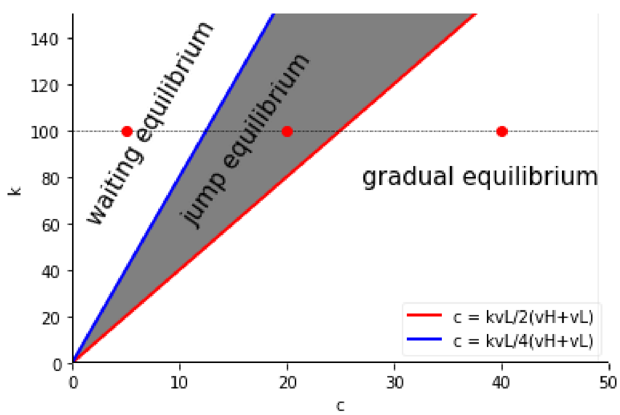
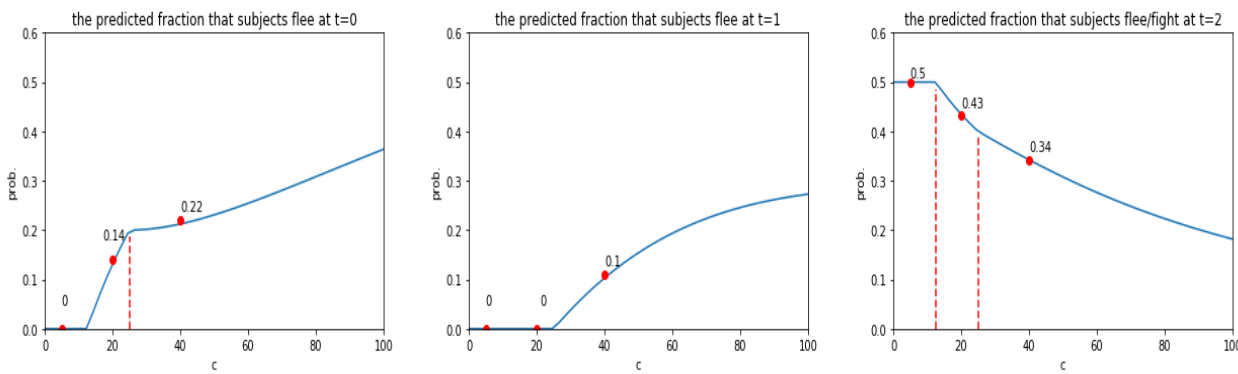



Figure 1: Equilibrium predictions, red dots indicate the locations of our experimental treatments.

From Figure 1, we predict that in the C5 treatment, all players will wait until the last period before deciding whether to fight or flee; in the C20 treatment, some players will flee immediately, while others will wait until the last period; and in the C40 treatment, players will flee gradually.

We next investigate how the fraction of players who flee (fight) at each period vary according to the value of the waiting cost.[11] The results are shown in Figure 2.




(a) t=0 (b) t=1 (c) t=2

Figure 2: Predicted fractions of players that flee (fight) at each period, red dots indicate the treatment values.

From Figure 2, we see that as waiting costs increase, more players will flee immediately and less players will choose to fight. This implies that waiting costs can be an effective tool for reducing battle rates, and accelerating the duration of contests. Moreover, similar to van Leeuwen *et al.* (2020), when waiting costs increase, the rate at which the weaker player in a pair escapes increases. Therefore, waiting costs also serve as a useful method for sorting. The qualitative hypotheses in our experiment are stated as:

**Qualitative Predictions**

**Hypothesis 1a (duration of contests)** *As waiting cost increases, contests will end in less time.*

**Hypothesis 1b (frequency of battles)** *As waiting cost increases, less battles will occur.*

**Hypothesis 1c (sorting)** *As waiting cost increases, the rate at which the weaker player in a pair escapes increases.*

[11] We use the gradual equilibrium stated in Proposition 3 as an example. The predicted fraction who flee at $t = 0$ is $a_1$, the predicted fraction who flee at $t = 1$ is $a_2 - a_1$, the predicted fraction who flee at $t = 2$ is $a_3 - a_2$ and the predicted fraction who fight at $t = 2$ is $1 - a_3$. When $v_H = v_L$, we can directly find that $a_3 - a_2 = 1 - a_3$, which means the fraction who flee at $t = 2$ should be the same as the fraction who fight at $t = 2$, that is demonstrated in Figure 2(c).

| Treatment | Equilibrium Type | Pr(t=0) | Pr(t=1) | Pr(t=2, w/escape) | Pr(t=2, w/battle) |
|---|---|---|---|---|---|
| C5 | Waiting | 0 | 0 | 0.5 | 0.5 |
| C20 | Jump | 0.24 | 0 | 0.38 | 0.38 |
| C40 | Gradual | 0.38 | 0.16 | 0.23 | 0.23 |

Table 2: Quantitative predictions of probabilities that contests end in each period.

Due to the fact that the experimental data fails to reflect players' strategies accurately (e.g., if players decide to fight at $t = 2$, while their opponent flees at $t = 0$, then the contest ends at $t = 0$, and their choice is counted as "wait"), we therefore use our prediction benchmarks as the probability that the contest will end at each period. The predicted probabilities are described in Table 2, which constitute our quantitative predictions.

**Quantitative Predictions**

**Hypothesis 2a (gradual equilibrium)** *There will be a "gradual equilibrium" in the C40 treatment, that is, the contest has a 38% chance of ending at $t = 0$, a 16% chance of ending at $t = 1$, and a 23% chance of ending at $t = 2$, with battle (escape).*

**Hypothesis 2b (jump equilibrium)** *There will be a "jump equilibrium" in the C20 treatment, that is, the contest has a 24% chance of ending at $t = 0$, and 38% chance of ending at $t = 2$, with battle (escape).*

**Hypothesis 2c (waiting equilibrium)** *There will be a "waiting equilibrium" in the C5 treatment, that is, the contest has a 50% chance of ending at $t = 2$, with battle (escape).*

# 4 Results

## 4.1 Overview of the experiments

The experiment was programmed in oTree (Chen *et al.* (2016)). We conducted all sessions at George Mason University, from September 2019 to November 2019. 204 (72 in C5, 64 in C20 and 68 in C40) undergraduate students participated in our experiment. The experiment lasted for about 60 minutes. Participants earned $20.30 (including the $5 show-up fees) on average.

As in van Leeuwen *et al.* (2020), all statistical tests comparing treatment differences use the session average as the independent unit of observation, unless indicated otherwise.

## 4.2 Treatment effects

To investigate whether increasing waiting costs can effectively shorten the duration of contests, we calculate the average length of time (in periods) in each treatment. The results are shown in Figure 3.

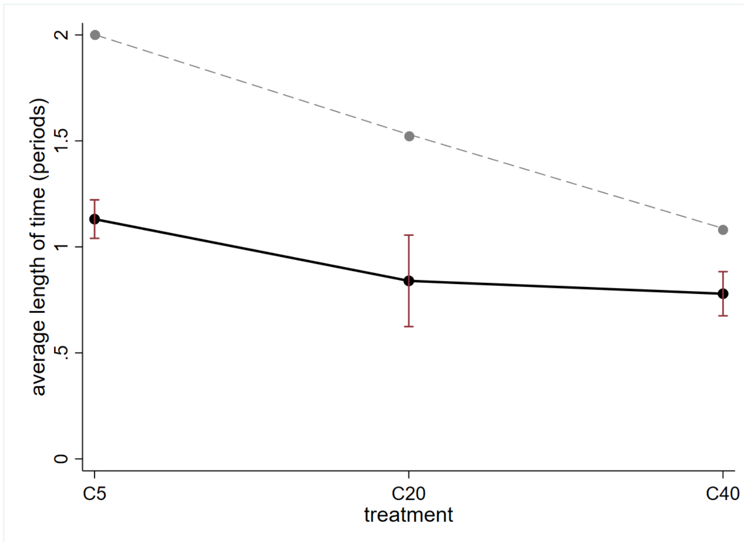



Figure 3: Average length of time (periods) in each treatment. Note: The solid line represents experimental results and the dash line represents theoretical predictions. Error bar indicates 95% confidence interval.

It is obvious that as waiting costs increase, contests end in less time (Jonckheere-Terpstra Test, $p < 0.001$ on descending ordered alternative, $N = 22$). However, if we make pair-wise comparisons, we find that contests in the C5 treatment take significantly more time to end than those in the C20 treatment (Mann-Whitney test, $p = 0.004$, $m = n = 7$); there is no significant difference in length of time between the C20 and C40 treatments (Mann-Whitney test, $p = 0.417$, $m = 7$, $n = 8$). Below, we derive our first result.

**Result 1** *When waiting costs are initially low, an increase in waiting costs significantly reduces the duration of contests. As waiting costs increase further, there is no effect on the duration of contests.*

Result 1 demonstrates that, within a certain range, waiting costs are an effective tool to adjust the duration of conflicts. For some conflicts, like military wars and strikes, it is better to end them sooner; however, for other conflicts, it is better to prolong their duration. For example, if two candidates compete for a promotion opportunity, the manager would like both to work hard for a long time and does not want either of them to quit early. Sports competitions are another example: it would be less enjoyable to watch a game where one side gives up quickly. From our experimental results, policymakers could adjust the length of conflicts by changing the waiting costs (through penalty, subsidy...), within a certain range, to improve the social efficiency.

Another qualitative hypothesis from our theory is that, as waiting costs increase, less battles will occur. To test this hypothesis, we calculate the frequency of battles in each treatment. The results are shown in Figure 4.

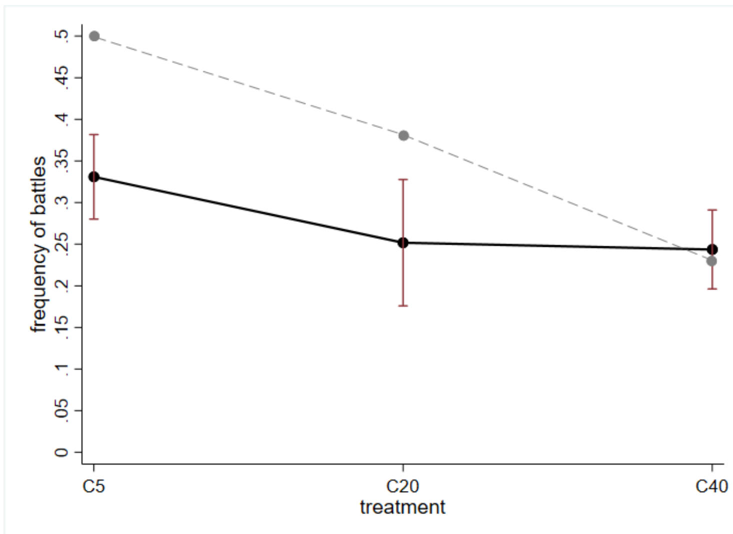



Figure 4: The frequency of battles in each treatment. Note: The solid line represents experimental results and the dash line represents theoretical predictions. Error bar shows 95% confidence interval.

The experimental results verify our qualitative prediction. Overall, as waiting costs become greater, battles are less likely to occur (Jonckheere test, $p = 0.018$ on descending ordered alternative, $N = 22$). In particular, when comparing the C5 with the C20 treatment, as waiting costs increase, battles are slightly less likely to occur (Mann-Whitney test, $p = 0.063$, $m = n = 7$). When comparing the C20 with the C40 treatment, rising waiting costs fail to significantly reduce the frequency of battles (Mann-Whitney test, $p = 0.908$, $m = 7$, $n = 8$). That gives our second result:

**Result 2** *When waiting costs are initially low, an increase in waiting costs significantly reduces the frequency of battles. As waiting costs increase further, there is no effect on the frequency of battles.*

One question that has been extensively investigated by researchers is how to effectively reduce the frequency of battles in conflicts. Our experimental results show that, when holding other factors constant, the frequency of battles has an "L" shaped relationship with the value of waiting costs. Specifically, when waiting costs are not too high, we could effectively reduce battles by raising waiting costs. However, when waiting costs become sufficiently great, the frequency of battles fails to decrease significantly as waiting costs increase.

We next investigate whether raising waiting costs will increase the rate that the weaker player in a pair chooses to flee. From theoretical predictions, as waiting costs increase, to avoid further losses on waiting, the weaker player in a pair should be more likely to choose to flee. Therefore, as in van Leeuwen *et al.* (2020), players will be more likely to sort themselves on their strengths when waiting costs become greater. We calculate the fraction of weaker players fleeing in each treatment. The results are shown in Figure 5.

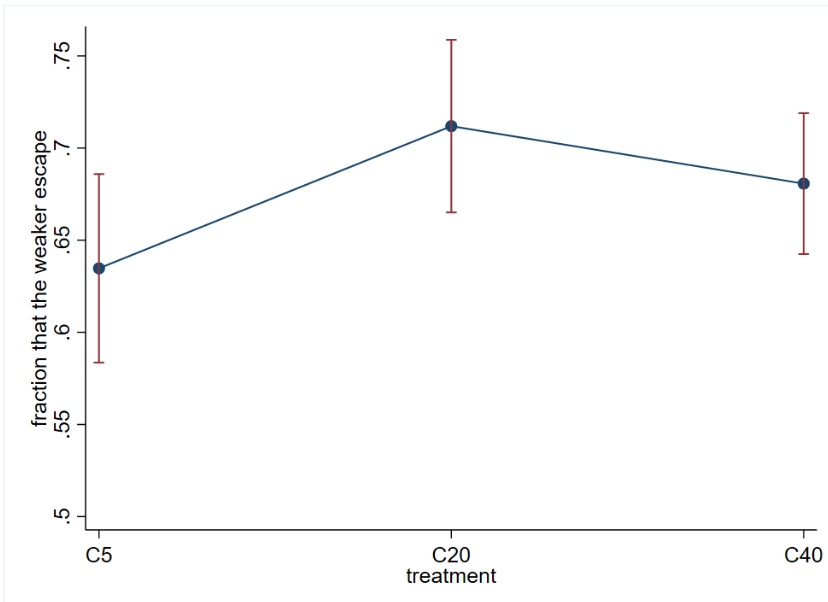



Figure 5: The fraction of weaker players flee in each treatment. Error bar indicates 95% significance level.

We find that as waiting costs increase, the weaker player in a pair chooses to flee becomes slightly more frequently (Jonckheere-Terpstra test, $p = 0.096$ on ascending ordered alternative, $N = 22$). Specifically, when comparing the C5 with the C20 treatment, as waiting costs increase, the weaker player in a pair become more likely to flee (Mann-Whitney test, $p = 0.025$, $m = n = 7$); by contrast, when comparing the C20 with the C40 treatment, there is no significant difference on sorting (Mann-Whitney test, $p = 0.224$, $m = 7$, $n = 8$). Thus, similar to previous results, our third result is stated as:

**Result 3** *When waiting costs are initially low, as waiting costs increase, players are significantly more likely to sort themselves by their strengths. As waiting costs increase further, there is no effect on sorting.*

We also conduct the diff-in-diff analysis. The regression results in Table 3 further confirm our previous findings. That is, when the waiting costs are initially not too great, increasing them can significantly shorten the duration of contests and reduce the frequency of battles. Likewise, players are more likely to sort themselves by their strengths. However, when waiting costs becomes sufficiently great, those treatment differences are not significant.

| | (1) Average length of time | (2) Frequency of battles | (3) Fraction that weaker escapes |
|---|---|---|---|
| C5 | $0.291^{***}$ | $0.079^{**}$ | $-0.077^{***}$ |
| | (0.086) | (0.035) | (0.027) |
| C40 | $-0.061$ | $-0.008$ | $-0.031$ |
| | (0.083) | (0.034) | (0.026) |
| Constant | $0.840^{***}$ | $0.252^{***}$ | $0.712^{***}$ |
| | (0.061) | (0.025) | (0.019) |
| Observations | 22 | 22 | 22 |

Table 3: OLS regressions on treatments comparisons. Note: We use the session average as the observational unit. The dependent variables are (1) the average length of time; (2) the frequency of battles; and (3) the fraction of the weaker player in a pair who escapes, in each session. All independent variables are dummy variables. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

We next test our quantitative hypotheses. The comparisons between the actual probabilities for contests end at each period, and the theoretical predictions are shown in Figure 6. Compared with theoretical predictions, in all treatments, contests are more likely to end at $t = 0$ (Mann-Whitney test, $p < 0.001$ for all treatments) and $t = 1$ (Mann-Whitney test, $p < 0.001$ for C5 and C20; $p = 0.072$ for C40); less likely to end at $t = 2$, with escape (Mann-Whitney test, $p < 0.001$ for all treatments). In the C5 and C20 treatments, contests are also less likely to end at $t = 2$, with battle (Mann-Whitney test, $p < 0.001$ for C5; $p = 0.017$ for C20). However, in the C40 treatment, contests are slightly more likely to end at $t = 2$, with battle (Mann-Whitney test, $p = 0.185$).[12] Therefore, our fourth results are stated as:

**Result 4** *Compared with theoretical predictions, contests are more likely to end earlier in all three treatments; In C5 and C20, contests are also less likely to end with battles. However, in C40, the actual probability that contests end with battles is slightly higher than the theoretical prediction.*

[12] If we use the loss aversion model predictions as the benchmarks, we find that $\lambda = 1.5$ is the optimal value to explain the patterns in our data; however, there are still significant discrepancies. Also, from the regression results in Table 4 & 5, we find that the degree of loss aversion levels plays only a small role in explaining why players flee earlier and fight less than theory predicts.

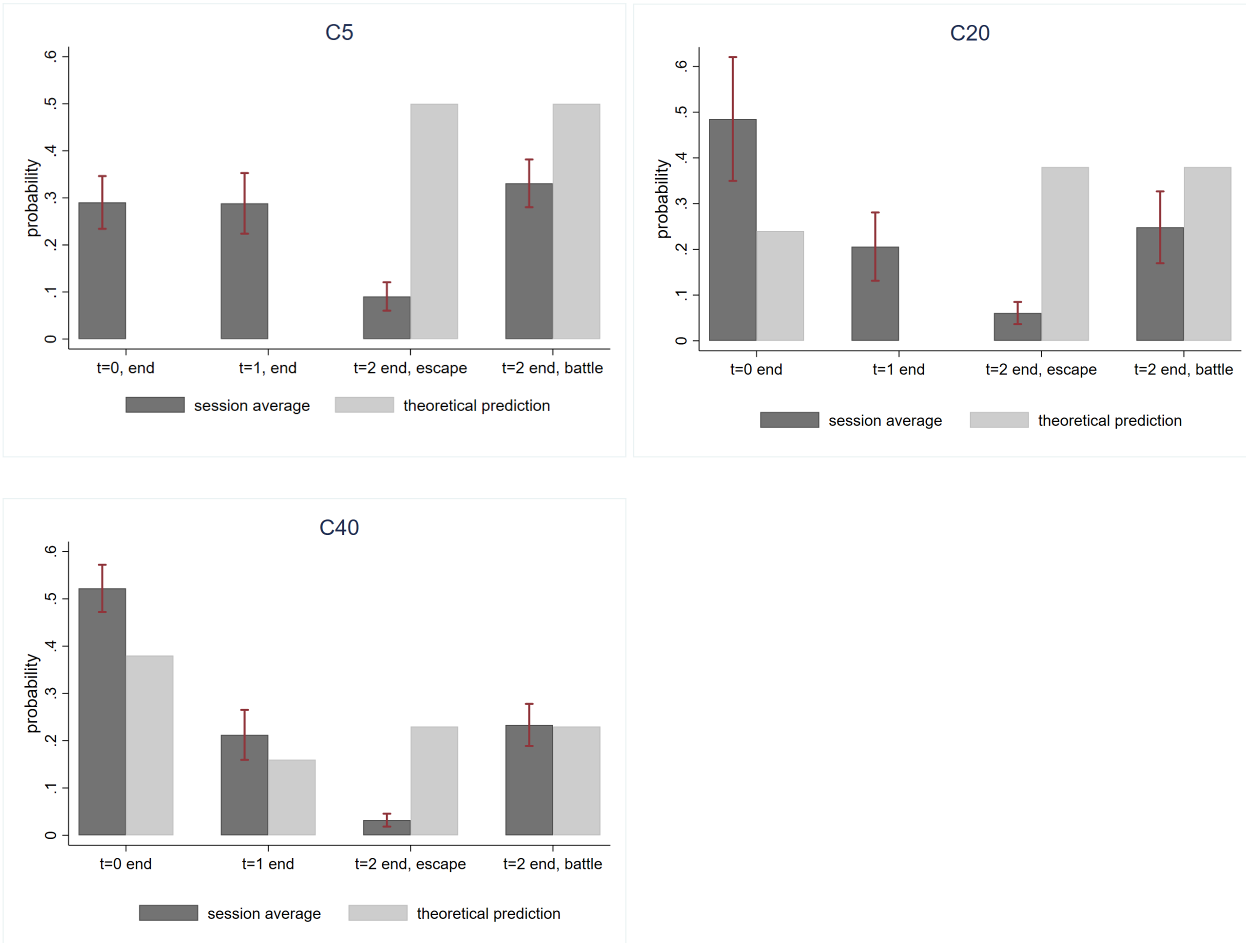



Figure 6: The comparisons of actual probabilities to the predicted values. The dark bars represent experimental results, while the light bars represent theoretical predictions. The error bar indicates 95% confidence interval.

## 4.3 Individual behavior

We now turn to investigating individual behavior; thus, the observational units mentioned in this subsection are at the individual level. First, we calculate the fraction of players who choose to flee, wait, or fight when the contest ends, in each treatment. The results are shown in Figure 7.

In all three treatments, in line with theoretical predictions, as a player's strength increases, the player is more likely to fight and less likely to flee. We next conduct regression analysis. The dependent variables indicate whether players flee immediately or fight at the last period. The results are set forth in Table 4 and Table 5, respectively.

From the estimation results, we find that players make decisions according to their strength; there are significant treatment differences on individual behavior, and both are

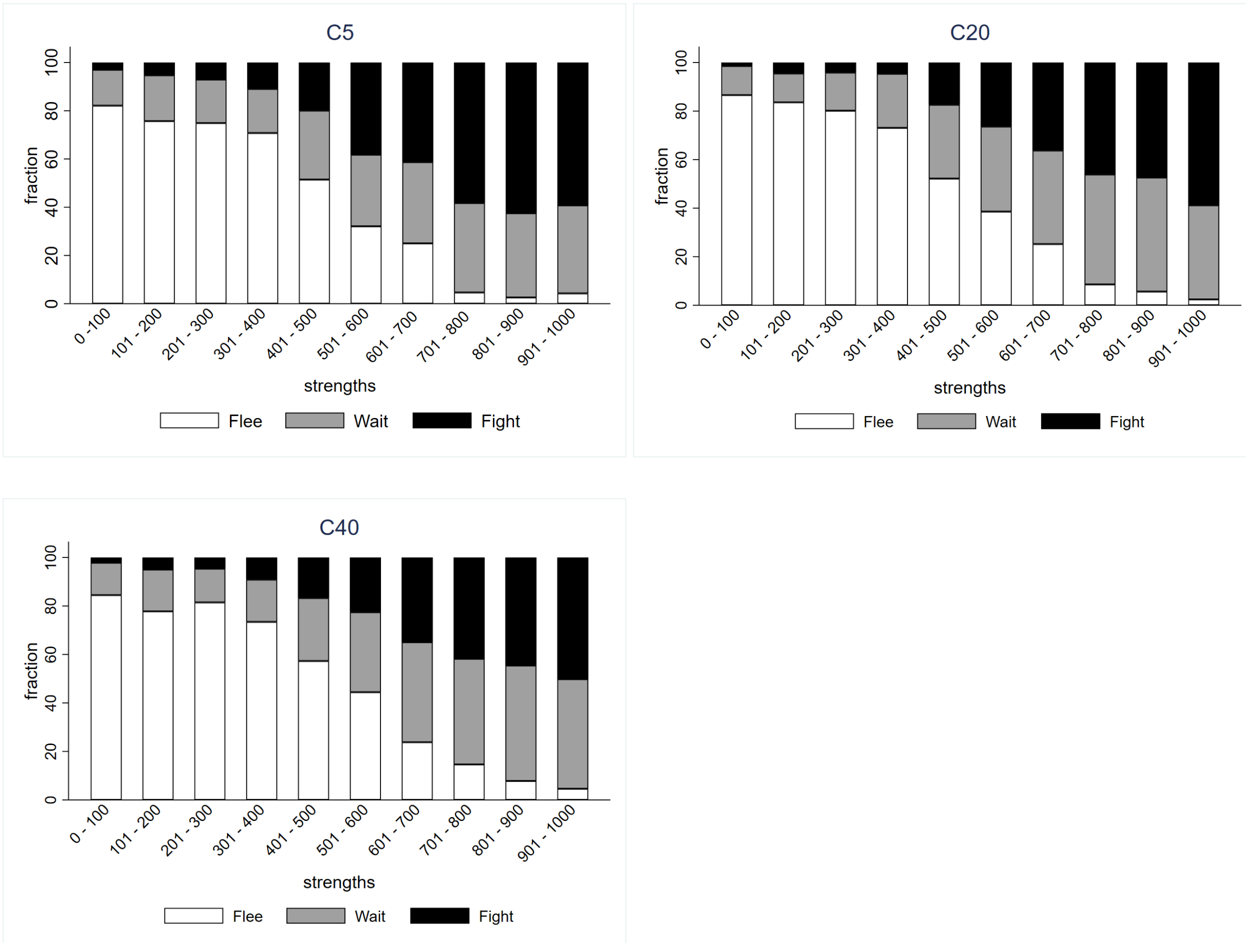



Figure 7: Individual behavior in each treatment.

consistent with our theoretical predictions. Aside from waiting costs and strength level, we observe a significant time trend: as time goes on, players are less likely to flee immediately, and slightly more likely to fight.

Following van Leeuwen *et al.* (2020), we use grid search to scrutinize each individual player's strategy. From Figure 7, we know that as strength increases, players are more likely to fight and less likely to flee immediately. Therefore, we assume players use the following cutoff strategy: each player's strategy can be characterized by two cutoff values, $e_1$,$e_2$, where $0 \leq e_1 \leq e_2 \leq 1000$. A player with strength below $e_1$ would flee immediately (at $t = 0$); a player with strength between $e_1$ and $e_2$ would flee during the middle of the contest (at $t = 1$ or $t = 2$); and a player with strength above $e_2$ would fight at $t = 2$. Figure 8 illustrates this cutoff strategy.

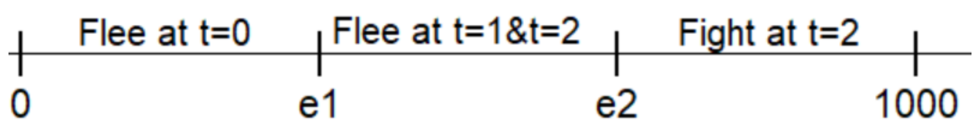



Figure 8: Individual cutoff strategy.

| | All periods | | | | Last 20 periods | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | C5 | C20 | C40 | All | C5 | C20 | C40 | All |
| Strengths | $-0.000^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ | $-0.000^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ | $-0.001^{***}$ |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Round | $-0.003^{***}$ | $-0.004^{***}$ | $-0.001$ | $-0.003^{***}$ | $-0.000$ | $-0.004^{***}$ | $-0.002$ | $-0.002^{**}$ |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.001) | (0.002) | (0.001) |
| Loss aversion | $-0.010$ | $-0.007$ | 0.007 | $-0.004$ | $-0.005$ | $-0.008$ | 0.014 | $-0.001$ |
| | (0.010) | (0.022) | (0.014) | (0.008) | (0.006) | (0.028) | (0.016) | (0.009) |
| Female | 0.057 | $-0.028$ | $-0.018$ | 0.014 | $0.065^{*}$ | $-0.012$ | 0.023 | 0.034 |
| | (0.048) | (0.036) | (0.085) | (0.034) | (0.039) | (0.034) | (0.087) | (0.031) |
| C5 | | | | $-0.113^{***}$ | | | | $-0.106^{***}$ |
| | | | | (0.033) | | | | (0.035) |
| C40 | | | | 0.042 | | | | $0.060^{**}$ |
| | | | | (0.029) | | | | (0.029) |
| Observations | 2880 | 2560 | 2720 | 8160 | 1440 | 1280 | 1360 | 4080 |

Table 4: Determinants on flee immediately. Note: Panel data probit regressions, allowing for random effects at the individual level. Coefficients are average marginal effects. The standard errors in parentheses are clustered at the session level. The dependent variables in each column indicate whether players choose to flee immediately in contests; Loss aversion is the number of lotteries players choose not to play in the loss aversion task; C5 equals 1 if the player is in the C5 treatment; C40 equals 1 if the player is in the C40 treatment. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

| | All periods | | | | Last 20 periods | | | |
|---|---|---|---|---|---|---|---|---|
| | (1) | (2) | (3) | (4) | (5) | (6) | (7) | (8) |
| | C5 | C20 | C40 | All | C5 | C20 | C40 | All |
| Strengths | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** | 0.001*** |
| | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) | (0.000) |
| Round | −0.001 | −0.001 | −0.001 | −0.001 | 0.001 | 0.003** | 0.004 | 0.003* |
| | (0.001) | (0.001) | (0.001) | (0.001) | (0.004) | (0.002) | (0.003) | (0.002) |
| Loss aversion | −0.020*** | 0.008 | −0.004 | −0.005 | −0.027*** | 0.003 | −0.006 | −0.010 |
| | (0.006) | (0.013) | (0.008) | (0.005) | (0.008) | (0.015) | (0.009) | (0.006) |
| Female | −0.052 | 0.006 | 0.036 | 0.001 | −0.065* | 0.004 | 0.025 | −0.008 |
| | (0.034) | (0.042) | (0.024) | (0.019) | (0.034) | (0.040) | (0.030) | (0.020) |
| C5 | | | | 0.062** | | | | 0.068* |
| | | | | (0.025) | | | | (0.037) |
| C40 | | | | −0.009 | | | | −0.002 |
| | | | | (0.032) | | | | (0.036) |
| Observations | 2880 | 2560 | 2720 | 8160 | 1440 | 1280 | 1360 | 4080 |

Table 5: Determinants on fight. Note: Panel data probit regressions, allowing for random effects at the individual level. Coefficients are average marginal effects. The standard errors in parentheses are clustered at the session level. The dependent variables in each column indicate whether players choose to fight in contests; Loss aversion is the number of lotteries players choose not to play in the loss aversion task; C5 equals 1 if the player is in the C5 treatment; C40 equals 1 if the player is in the C40 treatment. *** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

Specifically, if $e_2 > e_1 = 0$, it means the player would never flee at $t = 0$, no matter how small their strength; if $e_1 = e_2 > 0$, it means the player would either flee immediately or wait until the last period to fight; if $e_1 = e_2 = 0$, it means the player will always fight at the last period, no matter how small their strength; if $e_1 = e_2 = 1000$, that means the player will always flee immediately, no matter how great their strength. For each individual, we find their cutoffs $e_1$ and $e_2$, that maximize the number of accurately classified actions. The individual estimations are shown in Figure 9 and the average estimates on $e_1$ and $e_2$ are shown in Table 6. As in Table 6, on average, more than 85% of individuals' actions can be accurately classified by this cutoff estimation.

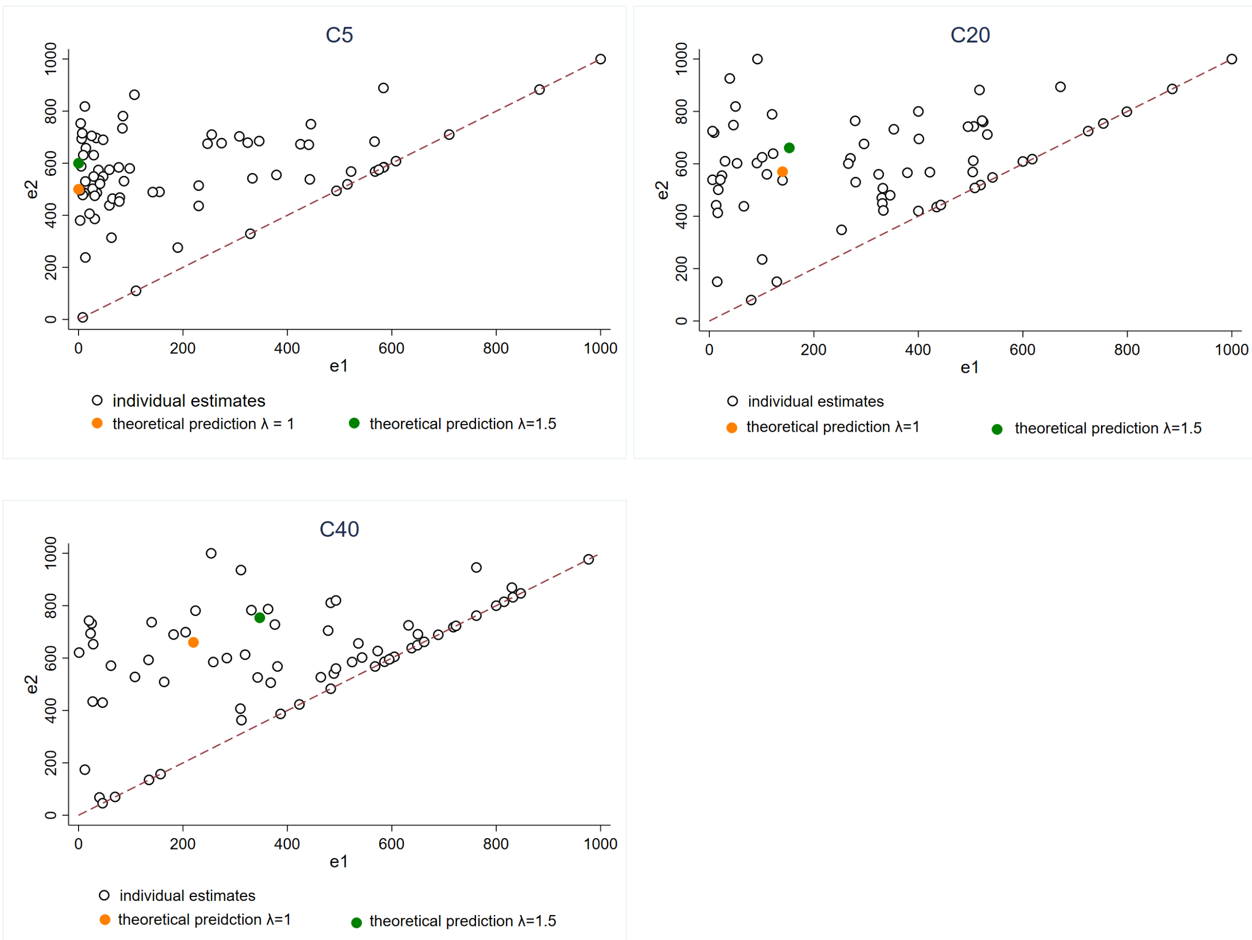



Figure 9: Individual estimated cutoff strategies. The orange dots show the theoretical predictions when the degree of loss aversion, $\lambda = 1$; the green dots show the theoretical predictions when the degree of loss aversion, $\lambda = 1.5$.

From individual cutoff estimations, as waiting costs increase, players are more likely to cluster on the 45-degree line where $e_1 = e_2$. That means players are less likely to flee during

| Treatment | Average cutoff estimated | | Correctly classified |
|---|---|---|---|
| | $e_1$ | $e_2$ | |
| C5 | 212.96 | 571.26 | 87.28% |
| | (0) | (500) | |
| C20 | 314.03 | 604.31 | 85.07% |
| | (140) | (570) | |
| C40 | 402.44 | 610.13 | 86.94% |
| | (220) | (660) | |

Table 6: Average cutoff estimates, theoretical predictions of $\lambda = 1$ are in parenthesis.

the middle of the contest (at $t = 1$ or $t = 2$).[13] Such behavior is in opposition to theoretical predictions, which predict that as waiting costs increase, players will flee gradually. One explanation consistent with those discrepancies is that players may fall into the "sunk cost fallacy." In our setting, the waiting cost can be treated as a sunk cost: once players have paid for it, they cannot get it back. Intuitively, the greater the amount of sunk cost players have invested, the less likely they will quit in the middle of the contest. Therefore, in the C40 treatment, individual behavior is more likely to "bifurcate": players either flee immediately or wait until the last period to fight against their opponents. As a result, the frequency of battles and the duration of contests in the C40 treatment fail to drop significantly when compared with the C20 treatment.

Also from Table 6, we find that individual cutoffs, $e_1$, are significantly greater than theoretical predictions (Mann-Whitney test, $p < 0.001$ in all three treatments). This means that a number of players prefer to flee immediately even though they have sufficient strength. Such behavior is consistent with the fact that contests end earlier than predicted in all three treatments.

As shown in Appendix D, theoretically, more loss averse players tend to flee immediately and fight less. Therefore, we first investigate whether loss aversion is related to the fact that players flee earlier than predicted. However, the regression results in Table 4 demonstrate that more loss averse players do not flee earlier in all three treatments.[14]

Another possible reason for why players flee earlier than predicted is that players may need time to learn. As Table 4 shows, players are less likely to flee immediately as time

[13] The regression results further support this finding: Using linear random effect regressions, allowing the random effects at the session level and the dependent variable is a dummy indicating whether the individual's estimated $e_1 = e_2$. We find that players in the C40 treatment are significant less likely to flee gradually ($p = 0.037$).

[14] van Leeuwen *et al.* (2020) observed that more loss averse players are more likely to flee before the last period. One possible explanation for the discrepancy between our finding and van Leeuwen *et al.* (2020) is that, in their work, players are allowed to fight at any period, and some players chose to fight earlier (which should not happen in equilibrium). As a result, it was risky to wait until the penultimate period before fleeing, which may have led to a correlation between early fleeing and loss aversion in their study.

goes on. Pro-social preferences may be related to early flight. To maximize payoffs, players have to end the contest immediately. If both players do this, they split the prize plus the deterrence value and the equitable outcome is realized.

# 5 Conclusion

Dynamic contests are ubiquitous and widely used in real-world social interactions; as a result, plenty of research has investigated how players behave in dynamic contests by using the *war of attrition*. However, in the original *war of attrition* model, players can only decide when to exit the contest. As a result, the contest has only one result: "escape." This single result cannot effectively capture real-world conflicts or model them accurately. In reality, there are many dynamic contests that end with players directly confronting their opponents (e.g., military wars and strikes).

Inspired by van Leeuwen *et al.* (2020), we extend the *war of attrition* by allowing players to actively battle their opponent at a certain point in time. In contrast to van Leeuwen *et al.* (2020), our three-period dynamic contest model allows players to choose whether to flee or wait in the first two periods. In the final period, they can choose whether to flee or fight. The reason we restrict players to fighting in the last period is that, similar to most natural environments, this allows their opponents time to escape. Thus, players can avoid costly battles. If players win the contest without battling, they can earn a prize, and a positive deterrence value.

Another, more important, feature of our model is that waiting is costly. That is, the longer players wait in the contest, the greater the waiting cost they incur. We show that the ratio of waiting cost to deterrence value is a key determinant to the type of equilibrium. When the waiting cost is relatively less than the deterrence value, all players wait until the last period, and decide whether to flee or fight; when the waiting cost is relatively greater than the deterrence value, players will flee gradually. The results in our model imply that contests will end earlier and result in less battles. Likewise, the weaker player in a pair will be more likely to flee as the waiting cost increases.

We pay special attention to the role of waiting costs in such dynamic contests. To verify our theoretical predictions, we conduct a laboratory experiment with three treatments that differ according to waiting costs. In our experiment, players compete within a group of two. They must decide whether to flee or wait (fight) at each period of the contest. The experimental results verify most key features of our model. Overall, as waiting cost increase, battles are less likely to occur, contests end in less time, and the weaker player in a pair is more likely to flee. However, as waiting costs become greater, the frequency of battles, the duration of contests, and the fraction that the weaker player in a pair manage to flee fail to decrease significantly. Moreover, if we make quantitative comparisons within each treatment, contests end earlier than theory predicts.

To investigate behavioral deviations from theoretical predictions in both between-treatment and within-treatment comparisons, we analyze each player's strategy. For between-treatment comparisons, unlike what theory predicts, as waiting cost increases, players are less likely to flee gradually. Instead, they either flee immediately or wait until the final period and fight against their opponents. Such behavioral patterns are consistent with the sunk cost effect. For within-treatment comparisons, there are some players who flee immediately even though they are sufficiently strong. As a result, contests end earlier than predicted in all three treatments.

Our model, combined with laboratory findings, has important policy implications. As is known theoretically, increasing the waiting cost can effectively reduce the frequency of battles and shorten the duration of the conflicts. Therefore, in real world conflicts, countries and institutions commonly use sanctions as a tool to raise waiting costs, thereby forcing their opponents to surrender quickly. However, our experimental results show that when waiting costs are sufficiently great, policymakers may be unable to alter the frequency of battles and the duration of conflicts simply by adjusting the waiting cost. Even worse, the sunk cost effect may lead battles becoming more prevalent when waiting costs are high. One interesting direction for follow-up studies is how to accurately identify the sunk cost effect and effectively lessen its influence in costly dynamic conflicts. Additionally, studying how to model and verify early flight behavior in dynamic contests is another direction for further investigation.

# References

Abbink, Klaus, Brandts, Jordi, Herrmann, Benedikt, & Orzen, Henrik. 2012. Parochial altruism in inter-group conflicts. *Economics Letters*, **117**(1), 45–48.

Acemoglu, Daron, & Wolitzky, Alexander. 2014. Cycles of conflict: An economic model. *American Economic Review*, **104**(4), 1350–67.

Arkes, Hal R, & Hutzel, Laura. 2000. The role of probability of success estimates in the sunk cost effect. *Journal of Behavioral Decision Making*, **13**(3), 295–306.

Baliga, Sandeep, Lucca, David O, & Sjöström, Tomas. 2011. Domestic political survival and international conflict: is democracy good for peace? *The Review of Economic Studies*, **78**(2), 458–486.

Bulow, Jeremy, & Klemperer, Paul. 1999. The generalized war of attrition. *American Economic Review*, **89**(1), 175–189.

Chen, Daniel L, Schonger, Martin, & Wickens, Chris. 2016. oTree—An open-source platform for laboratory, online, and field experiments. *Journal of Behavioral and Experimental Finance*, **9**, 88–97.

Deck, Cary, & Jahedi, Salar. 2015. An experimental investigation of time discounting in strategic settings. *Journal of Behavioral and Experimental Economics*, **54**, 95–104.

Deck, Cary, & Sheremeta, Roman M. 2012. Fight or flight? Defending against sequential attacks in the game of siege. *Journal of Conflict Resolution*, **56**(6), 1069–1088.

Dutcher, E Glenn, Balafoutas, Loukas, Lindner, Florian, Ryvkin, Dmitry, & Sutter, Matthias. 2015. Strive to be first or avoid being last: An experiment on relative performance incentives. *Games and Economic Behavior*, **94**, 39–56.

Friedman, Daniel, Pommerenke, Kai, Lukose, Rajan, Milam, Garrett, & Huberman, Bernardo A. 2007. Searching for the sunk cost fallacy. *Experimental Economics*, **10**(1), 79–104.

Fudenberg, Drew, Gilbert, Richard, Stiglitz, Joseph, & Tirole, Jean. 1983. Preemption, leapfrogging and competition in patent races. *European Economic Review*, **22**(1), 3–31.

Gächter, Simon, Johnson, Eric J, & Herrmann, Andreas. 2007. Individual-level loss aversion in riskless and risky choices.

Gelder, Alan, & Kovenock, Dan. 2017. Dynamic behavior and player types in majoritarian multi-battle contests. *Games and Economic Behavior*, **104**, 444–455.

Gul, Faruk, & Pesendorfer, Wolfgang. 2012. The war of information. *The Review of Economic Studies*, **79**(2), 707–734.

Hörisch, Hannah, & Kirchkamp, Oliver. 2010. Less fighting than expected. *Public choice*, **144**(1-2), 347–367.

Hörner, Johannes, & Sahuguet, Nicolas. 2011. A war of attrition with endogenous effort levels. *Economic Theory*, **47**(1), 1–27.

Kimbrough, Erik O, Laughren, Kevin, & Sheremeta, Roman. 2017. War and conflict in economics: Theories, applications, and recent trends. *Journal of Economic Behavior & Organization*.

Mago, Shakun D, & Sheremeta, Roman M. 2019. New Hampshire Effect: behavior in sequential and simultaneous multi-battle contests. *Experimental Economics*, **22**(2), 325–349.

Oprea, Ryan, Wilson, Bart J, & Zillante, Arthur. 2013. War of attrition: Evidence from a laboratory experiment on market exit. *Economic Inquiry*, **51**(4), 2018–2027.

Rohner, Dominic, Thoenig, Mathias, & Zilibotti, Fabrizio. 2013. War signals: A theory of trade, trust, and conflict. *Review of Economic Studies*, **80**(3), 1114–1147.

Roth, Stefan, Robbert, Thomas, & Straus, Lennart. 2015. On the sunk-cost effect in economic decision-making: a meta-analytic review. *Business research*, **8**(1), 99–138.

Smith, J Maynard. 1974. The theory of games and the evolution of animal conflicts. *Journal of theoretical biology*, **47**(1), 209–221.

Song, Jian, & Houser, Daniel. 2021. Non-exclusive group contests: an experimental analysis. *Working paper*.

van Leeuwen, Boris, Offerman, Theo, & van de Ven, Jeroen. 2020. Fight or flight: Endogenous timing in conflicts. *Review of Economics and Statistics*, 1–45.

# Appendix A. Waiting equilibrium

**Lemma 1 (Monotony, adapted from van Leeuwen *et al.* (2020))** *(i) Suppose there is an interval of types $a_t = (a_1, a_2)$ that act at period t=2 and let $a_i$,$a_j \in a_t$. If there is a type $a_i$ that is indifferent between $(t = 2, F)$ and $(t = 2, R)$, then all types $a_j > a_i$ strictly prefer $(t = 2, F)$ to $(t = 2, R)$ and all types $a_j < a_i$ strictly prefer $(t = 2, R)$ to $(t = 2, F)$.*
*(ii) If there is an equilibrium in which there is a player with type $a_i$ strictly prefers $(t = 2, F)$ to $(t^{'}, R)$ for any $t^{'}$, then any players with type $a_j > a_i$ strictly prefers $(t = 2, F)$ to $(t^{'}, R)$ for any $t^{'}$.*

**Proof.** (i) Consider two types, $a^{'}, a^{''} > a^{'}$. Suppose $a^{'}$ prefers $(t = 2, F)$ to $(t = 2, R)$, which means:

$$\Delta(a^{'}) = V_i((t = 2, F), a^{'}, \cdot) - V_i((t = 2, R), a^{'}, \cdot) \geq 0 \tag{A.1}$$

For type $a \in (a^{'}, a^{''})$, fighting in $t = 2$ instead of fleeing in $t = 2$ will never decrease and may increase the likelihood of ending up in a battle. And $a^{'}$ will lose such battle, $a^{''}$ would win it. If opponent has the ability $a \notin (a^{'}, a^{''})$, then $\Delta(\cdot)$ is affected equally for $a^{'}$ and $a^{''}$. Thus, for any $a^{''} > a^{'}$, we have:

$$\Delta(a^{''}) - \Delta(a^{'}) \propto \int_{a \in (a', a'') | S(a) \in (F,R)} g(a)\, da \geq 0 \tag{A.2}$$

Which means, $\Delta(a^{''}) = V_i((t = 2, F), a'', \cdot) - V_i((t = 2, R), a'', \cdot) \geq 0$. (i) can be proved.

To show (ii), if $a_j > a_i$, there is strictly positive probability of meeting an opponent with ability $a \in (a_i, a_j)$, and $\Delta(a_j) > \Delta(a_i)$. Vice versa. ■

We need to prove that when $c \leq \frac{kv_L}{4(v_H+v_L)}$, there exists a waiting equilibrium:
(1) The monotonicity stated in the Lemma is also valid here. There exists a positive fraction of types that for some time $t^{'}$, they prefer $(t^{'}, R)$ to $(t = 2, F)$.
(2) We show that, given types $(\hat{a}, 1]$ fight at $t = 2$, the best response for types $[0, \hat{a}]$ is to flee at $t = 2$: For type $a_j \in [0, \hat{a}]$, if they flee at $t = 0$, then they earn 0; and if they flee at $t = 1$, they earn $-c$. So, if the expected payoff for $a_j \in [0, \hat{a}]$ to flee at $t = 2$ is greater than zero, then they should choose to flee at $t = 2$. The expected payoff for $a_j \in [0, \hat{a}]$ to flee at $t = 2$ is:

$$\hat{a}(\frac{v_H + k}{2}) + (1 - \hat{a})(-\frac{v_L}{2}) - 2c \tag{A.3}$$

Plug $\hat{a}$ into (A.3), and rearrange, it equals:

$$\frac{kv_L}{2(v_H + v_L)} - 2c \tag{A.4}$$

If $c \leq \frac{kv_L}{4(v_H+v_L)}$, then (A.4) is greater than zero, types $[0, \hat{a}]$ have no incentive to deviate. Then, from Lemma, type $\hat{a}$ is indifferent between $(t=2, R)$ and $(t=2, F)$ and all types above $\hat{a}$ would strictly prefer to fight at $t=2$.

# Appendix B. Jump equilibrium

First, to satisfy the constraints that $0 \leq \overline{a_1} \leq \overline{a_2} \leq 1$, we have $c \geq \frac{kv_L}{4(v_H+v_L)}$.

Second, when $\frac{kv_L}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2(v_H+v_L)}$, the jump equilibrium exists:
(1) The monotonicity stated in the Lemma is also valid here: there is a positive fraction of types that for some $t^{'}$, they prefer $(t^{'}, R)$ to $(t=2, F)$.
(2) For types below $\overline{a_1}$, they have no incentive to deviate: If they flee late, they need to pay the waiting cost. And for types between $\overline{a_1}$ and $\overline{a_2}$, they are indifferent between $(t=0, R)$ and $(t=2, R)$. We need to show that given other strategies, for any $a_j \in (\overline{a_1}, \overline{a_2}]$, fleeing at $t=2$ produces higher payoff than fleeing at $t=1$.

The expected payoff for $a_j \in (\overline{a_1}, \overline{a_2}]$ fleeing at $t=2$ is:

$$\overline{a_1}(v_H + k) + (\overline{a_2} - \overline{a_1})\frac{v_H + k}{2} - (1 - \overline{a_2})\frac{v_L}{2} - (1 - \overline{a_1})2c \tag{B.1}$$

The expected payoff for $a_j \in (\overline{a_1}, \overline{a_2}]$ fleeing at t=1 is:

$$\overline{a_1}(v_H + k) - (1 - \overline{a_1})c \tag{B.2}$$

Plug $\overline{a_1}$ and $\overline{a_2}$ into (B.1) and (B.2), and let $(B.1) - (B.2) \geq 0$, we can get:

$$c \leq \frac{kv_L}{2(v_H + v_L)} \tag{B.3}$$

(B.3) is always satisfied when $\frac{kv_L}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2(v_H+v_L)}$.

(3) We need to prove types above $\overline{a_2}$ have no incentive to deviate. From monotonicity, type $\overline{a_2}$ is indifferent between $(t=2, F)$ and $(t=2, R)$, then all types above $\overline{a_2}$ will strictly prefer to fight at $t=2$.

Finally, we can obtain:

$$\frac{\partial \overline{a_1}}{\partial c} = \frac{4(v_H + v_L)(kv_L + v_H(k + v_H + v_L))}{(4c(v_H + v_L) + v_H(k + v_H + v_L))^2} > 0 \tag{B.4}$$

And

$$\overline{a_2} - \overline{a_1} = \frac{v_H v_L + k v_L}{4c(v_H + v_L) + v_H(k + v_H + v_L)} \rightarrow \frac{\partial(\overline{a_2} - \overline{a_1})}{\partial c} < 0 \tag{B.5}$$

$$1 - \overline{a_2} = \frac{v_H(k + v_H)}{4c(v_H + v_L) + v_H(k + v_H + v_L)} \rightarrow \frac{\partial(1 - \overline{a_2})}{\partial c} < 0 \tag{B.6}$$

For jump equilibrium, as waiting costs increase, more types will flee at t=0, while less types flee (fight) at t=2.

# Appendix C. Gradual equilibrium

First, to satisfy the condition: $0 \leq a_1 \leq a_2 \leq a_3 \leq 1$, the deterrence value k and the waiting cost c should have the relation: $c \geq \frac{kv_L}{2(v_H+v_L)}$.
To satisfy the constraint $a_1 \geq 0$, we must have:

$$4(v_H + v_L)c^2 - 2kv_L c + kv_L(k + v_H) \geq 0 \tag{C.1}$$

We can prove that, if $v_H$,$v_L$ and k are all positive, then for any $c > 0$, (C.1) always greater than zero.
To satisfy the constraint $a_1 \leq a_2$, we must have:

$$-v_L k^2 + (2cv_H - v_H v_L + 2cv_L)k + 2c(v_H^2 + v_H v_L) \geq 0 \tag{C.2}$$

Combined with $k > 0$, we can get:

$$c \geq \frac{kv_L}{2(v_H + v_L)} \tag{C.3}$$

To satisfy $a_2 \leq a_3$, let $a_3 - a_2$, we can get:

$$a_3 - a_2 = \frac{(k + v_H)^2 v_L}{2cv_H(k + v_H + v_L) + [4c^2 + (k + v_H)^2](v_H + v_L)} \tag{C.4}$$

Which is obviously that $a_2 \leq a_3$.
To prove $a_3 \leq 1$, rearrange $a_3$, we can obtain:

$$a_3 = 1 - \frac{(k + v_H)^2 v_H}{2cv_H(k + v_H + v_L) + [4c^2 + (k + v_H)^2](v_H + v_L)} \tag{C.5}$$

Which is obvious that $a_3 \leq 1$.

Second, we prove that when $c \geq \frac{kv_L}{2(v_H+v_L)}$, the gradual equilibrium exists:
(1) There is a positive fraction of types that for some $t'$, they prefer $(t', R)$ to $(t=2, F)$. If this was not the case, and no positive fraction flees at some point, then for sufficiently small $\epsilon > 0$, all types on $[0, \epsilon)$ strictly gain by fleeing at $t = 0$: this would strictly increase the probability of an escape and they almost surely lose a battle for $\epsilon$ sufficiently small.
(2) The Lemma implies that if some types fight at $t = 2$, then all stronger types must fight at $t = 2$.
(3) Under threshold strategy, no types gain from deviating. All types $[0, a_3]$ have the same equilibrium payoff $a_1(\frac{v_H+k}{2}) > 0$, and are indifferent between (t=0, R), (t=1, R), (t=2, R) and (t=2, F). From the Lemma, if type $a_3$ is indifferent between $(t = 2, R)$ and $(t = 2, F)$, then all types above $a_3$ will strictly prefer to fight at $t = 2$.

Finally, we can obtain that:

$$\frac{\partial a_1}{\partial c} = \frac{2(g_1(v_H, v_L, c, k) - g_2(v_H, v_L, c, k))}{(2cv_H(k + v_H + v_L) + (4c^2 + (k + v_H)^2)(v_H + v_L))^2} \tag{C.6}$$

Where:

$$g_1(v_H, v_L, c, k) = [4c(v_H+v_L) - kv_L][(2cv_H(k+v_H+v_L) + (4c^2 + (k+v_H)^2)(v_H+v_L)] \tag{C.7}$$

$$g_2(v_H, v_L, c, k) = [4c(v_H+v_L) + v_H(k+v_H+v_L)][4c^2(v_H+v_L) + kv_L(-2c+k+v_H)] \tag{C.8}$$

To simplify, we now suppose $v_H = v_L = v > 0$, then $\frac{\partial a_1}{\partial c}$ can be written as:

$$\frac{\partial a_1}{\partial c} = \frac{(8c - k)[4c^2 + c(k + 2v) + (k + v)^2] - [8c^2 + k(-2c + k + v)](\frac{8c+k+2v}{2})}{[4c^2 + c(k + 2v) + (k + v)^2]^2} \tag{C.9}$$

And we can show that as $c \geq \frac{kv_L}{2(v_H+v_L)} = \frac{k}{4}$, $\frac{\partial a_1}{\partial c} > 0$. As waiting costs increase, more types would flee at $t = 0$.

We can also show that, as $v_H = v_L = v > 0$ and $c \geq \frac{kv_L}{2(v_H+v_L)} = \frac{k}{4}$

$$\frac{\partial(a_2 - a_1)}{\partial c} \propto \frac{-16c^2 + 8kc + k(k + 2v)^2 + 4(k + v)^2}{[4c^2 + c(k + 2v) + (k + v)^2]^2} > 0 \tag{C.10}$$

And it is obvious that:

$$a_3 - a_2 = \frac{v(k + v)^2}{2cv(k + 2v) + 2v[4c^2 + (k + v)^2]} \rightarrow \frac{\partial(a_3 - a_2)}{\partial c} < 0 \tag{C.11}$$

$$1 - a_3 = \frac{v(k+v)^2}{2cv(k+2v) + 2v[4c^2 + (k+v)^2]} \rightarrow \frac{\partial(1-a_3)}{\partial c} < 0 \tag{C.12}$$

Which demonstrates that as waiting costs increase, more types will flee at $t = 1$, less types will flee (fight) at $t = 2$.

# Appendix D. Homogenous loss aversion model

We extend our model by assuming that players are (homogenous) loss averse. As in van Leeuwen *et al.* (2020), we use the following utility function to characterize loss aversion situations:

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

Where $\lambda \geq 1$ measures the degree of loss aversion, higher $\lambda$ means players are more loss averse. Intuitively, the more loss averse players will flee more frequently and earlier, also fight less. Similar to the above, we prove that under certain conditions, there exists pure-strategy Bayesian equilibrium; the proportion of players who flee (fight) at each period depends on the degree of loss aversion.

**Proposition 4 (Waiting equilibrium under loss aversion)**

1. *Suppose* $\frac{v_H+k}{4} \geq \max\{c, \frac{kv_L}{2[v_H+(2-\lambda)v_L]}\}$ *and* $\lambda \leq \min\{\frac{v_H}{v_L}, \frac{2k}{v_H+k} - \frac{v_H}{v_L}\} + 2$.*If* $c \leq \frac{kv_L}{4(v_H+v_l)}$, *there exist a "waiting equilibrium": types below* $\hat{a_\lambda}$ *flee at* $t = 2$ *and types above* $\hat{a_\lambda}$ *fight at t=2. Type* $\hat{a_\lambda}$ *is indifferent between fleeing at* $t = 2$ *and fighting at* $t = 2$. *Where:*

$$\hat{a_\lambda} = \frac{\lambda vl}{vH + \lambda vl} \tag{D.2}$$

2. *If players are more loss averse, then more fractions flee at* $t = 2$ *and less fractions fight at* $t = 2$.

**Proof.**

1. First, to obtain the threshold value, $\hat{a_\lambda}$, we use the same method as stated in the case that $\lambda = 1$. Define:

$$\Phi(x) = \begin{cases} x, & \text{for } x \geq 0 \\ -\lambda x, & \text{for } x < 0 \end{cases} \tag{D.3}$$

We assume that type $\hat{a_\lambda}$ is indifferent between $(t = 2, R)$ and $(t = 2, F)$. Then, from the Lemma, types below $\hat{a_\lambda}$ prefer $(t = 2, R)$ to $(t = 2, F)$ and types above $\hat{a_\lambda}$ prefer

$(t=2,R)$ to $(t=2,F)$. The expected payoff for type $\hat{a_\lambda}$ that flees at $t=2$ is:

$$\Phi(\frac{v_H+k}{2}-2c)+(1-\hat{a_\lambda})[\frac{1}{2}\Phi(0-2c)+\frac{1}{2}\Phi(-v_L-2c)] \tag{D.4}$$

Under the assumption that $c \leq \frac{v_H+k}{4}$, (D.4) can be simplified as:

$$\hat{a_\lambda}(\frac{v_H+k}{2}-2c)-\lambda(1-\hat{a_\lambda})(2c+\frac{v_L}{2}) \tag{D.5}$$

The expected payoff for $\hat{a_\lambda}$ that fights at $t=2$ is:

$$\hat{a_\lambda}(v_H+\frac{k}{2}-2c)-\lambda(1-\hat{a_\lambda})(v_L+2c) \tag{D.6}$$

Since $\hat{a_\lambda}$ is indifferent between $(t=2,R)$ and $(t=2,F)$. Let (D.5)=(D.6), we can get

$$\hat{a_\lambda}=\frac{\lambda vl}{vH+\lambda vl} \tag{D.7}$$

Next, we need to prove that when $c \leq \frac{kv_L}{4(v_H+v_L)}$, such waiting equilibrium exists:

(1) The monotonicity which stated in the Lemma is also valid here. There exists a positive fraction of types that for some time $t'$, they prefer $(t',R)$ to $(t=2,F)$.

(2) Given types $(\hat{a_\lambda},1]$ that fight at $t=2$, the best response for types $[0,\hat{a_\lambda}]$ is flee at t=2: For type $a_{\lambda j} \in [0,\hat{a_\lambda}]$, if they flee at $t=0$, they would earn 0; and if they flee at $t=1$, they would earn $-\lambda c$. So, if the expected payoff for $a_{\lambda j} \in [0,\hat{a_\lambda}]$ to flee at $t=2$ is greater than zero, then they would choose to flee at $t=2$.

Given others' strategies, the expected payoff for $a_{\lambda j}$ fleeing at $t=2$ is:

$$\hat{a_\lambda}(\frac{v_H+k}{2}-2c)-\lambda(1-\hat{a_\lambda})(\frac{v_L}{2}+2c) \tag{D.8}$$

Let $(D.8) \geq 0$, we obtain:

$$4c \leq k+v_H-\lambda\frac{\hat{a_\lambda}}{1-\hat{a_\lambda}}(4c+v_L) \rightarrow c \leq \frac{kv_L}{4(v_H+v_l)} \tag{D.9}$$

Which demonstrates that when $c \leq \frac{kv_L}{4(v_H+v_l)}$, such waiting equilibrium exists.

2. It is obvious that:

$$\hat{a_\lambda}=\frac{\lambda v_L}{v_H+\lambda v_L}=1-\frac{v_H}{v_H+\lambda v_L} \tag{D.10}$$

Thus,

$$\frac{\partial \hat{a_\lambda}}{\partial \lambda} > 0 \tag{D.11}$$

$$\frac{\partial (1-\hat{a_\lambda})}{\partial \lambda} < 0 \tag{D.12}$$

As players become more loss averse, they are more likely to flee at $t = 2$ and less likely to fight at $t = 2$.

■

**Proposition 5 (Jump equilibrium under loss aversion)**

1. *Suppose* $\frac{v_H+k}{4} \geq max\{c, \frac{kv_L}{2[v_H+(2-\lambda)v_L]}\}$ *and* $\lambda \leq min\{\frac{v_H}{v_L}, \frac{2k}{v_H+k} - \frac{v_H}{v_L}\} + 2$. *If* $\frac{kvl}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2[v_H+(2-\lambda)v_L]}$, *there exist a "jump equilibrium": a fraction of* $\overline{a_{1\lambda}}$ *flee at* $t = 0$, *a fraction of* $\overline{a_{2\lambda}} - \overline{a_{1\lambda}}$ *flee at* $t = 2$, *types above* $\overline{a_{2\lambda}}$ *fight at* $t = 2$. *where:*

$$\overline{a_{1\lambda}} = \frac{\lambda[4c(v_H+v_L) - kv_L]}{4c\lambda(v_H+v_L) + v_H(k+v_H+\lambda v_L)} \tag{D.13}$$

$$\overline{a_{2\lambda}} = \frac{\lambda[v_Hv_L + 4c(v_H+v_L)]}{4c\lambda(v_H+v_L) + v_H(k+v_H+\lambda v_L)} \tag{D.14}$$

2. *If players are more loss averse, more fractions flee at t=0, more fractions flee at t=2 and less fractions fight at t=2.*

**Proof.**

1. To satisfy the constraint: $0 \leq \overline{a_{1\lambda}} \leq \overline{a_{2\lambda}} \leq 1$, the following condition should be satisfied:

$$c \geq \frac{kv_L}{4(v_H+v_L)} \tag{D.15}$$

Next we need to prove that when $\frac{kv_L}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2[v_H+(2-\lambda)v_L]}$, such jump equilibrium exists:

(1) The monotonicity that stated in the Lemma is also valid here: there is a positive fraction of types that for some $t'$, they prefer $(t', R)$ to $(t = 2, F)$.

(2) For types below $\overline{a_{1\lambda}}$, they have no incentive to deviate: If they flee late, they need to pay waiting costs. And for types between $\overline{a_{1\lambda}}$ and $\overline{a_{2\lambda}}$, they are indifferent between $(t = 0, R)$ and $(t = 2, R)$. We need to show given other strategies, for any $a_{j\lambda} \in (\overline{a_{1\lambda}}, \overline{a_{2\lambda}}]$, fleeing at $t = 2$ produces higher payoff than fleeing at $t = 1$:
The expected payoff for $a_{j\lambda} \in (\overline{a_{1\lambda}}, \overline{a_{2\lambda}}]$ that flees at $t = 2$ is:

$$\overline{a_{1\lambda}}(v_H+k) + (\overline{a_{2\lambda}} - \overline{a_{1\lambda}})(\frac{v_H+k}{2} - 2c) - \lambda(1-\overline{a_{2\lambda}})(\frac{v_L}{2} + 2c) \tag{D.16}$$

The expected payoff for $a_{j\lambda} \in (\overline{a_{1\lambda}}, \overline{a_{2\lambda}}]$ that flees at $t = 1$ is:

$$\overline{a_{1\lambda}}(v_H + k) - \lambda(1 - \overline{a_{1\lambda}})c \tag{D.17}$$

Plug $\overline{a_{1\lambda}}$ and $\overline{a_{2\lambda}}$ into (D.16) and (D.17), and let (D.16)-(D.17)$\geq 0$, we can get:

$$c \leq \frac{kv_L}{2[v_H + (2-\lambda)v_L]} \tag{D.18}$$

(D.18) is always satisfied when $\frac{kv_L}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2[v_H+(2-\lambda)v_L]}$.

(3) We then need to prove that types above $\overline{a_{2\lambda}}$ have no incentive to deviate. From monotonicity, type $\overline{a_{2\lambda}}$ is indifferent between $(t = 2, F)$ and $(t = 2, R)$, then all types above $\overline{a_{2\lambda}}$ will strictly prefer to fight at $t = 2$.

2. We can obtain:

$$\frac{\partial \overline{a_{1\lambda}}}{\partial \lambda} = \frac{[4c(v_H + v_L) - kv_L]v_H(k + v_H)}{[4c\lambda(v_H + v_L) + v_H(k + v_H + v_L)]^2} \tag{D.19}$$

Since $c \geq \frac{kv_L}{4(v_H+v_L)}$, so $\frac{\partial \overline{a_{1\lambda}}}{\partial \lambda} \geq 0$.
Also, we can get:

$$\frac{\partial(\overline{a_{2\lambda}} - \overline{a_{1\lambda}})}{\partial \lambda} = \frac{v_H v_L(k + v_H)^2}{[4c\lambda(v_H + v_L) + v_H(k + v_H + v_L)]^2} \rightarrow \frac{\partial(\overline{a_{2\lambda}} - \overline{a_{1\lambda}})}{\partial \lambda} > 0 \tag{D.20}$$

$$\frac{\partial(1 - \overline{a_{2\lambda}})}{\partial \lambda} = \frac{[4c(v_H + v_L) + v_H v_L][-v_H(k + v_H)]}{[4c\lambda(v_H + v_L) + v_H(k + v_H + v_L)]^2} \rightarrow \frac{\partial(1 - \overline{a_{2\lambda}})}{\partial \lambda} < 0 \tag{D.21}$$

Which demonstrate that, as players become more loss averse, more fractions will flee at $t = 0$, more fractions will flee at $t = 2$ and less fractions will flee at $t = 2$.

■

**Proposition 6 (Gradual equilibrium under loss aversion)**
*Suppose* $\frac{v_H+k}{4} \geq \max\{c, \frac{kv_L}{2[v_H+(2-\lambda)v_L]}\}$ *and* $\lambda \leq \min\{\frac{v_H}{v_L}, \frac{2k}{v_H+k} - \frac{v_H}{v_L}\} + 2$. *If* $\frac{kv_L}{2[v_H+(2-\lambda)v_L]} \leq c \leq \frac{vH+k}{4}$, *there exist a "gradual equilibrium": a fraction of* $a_{1\lambda}$ *flee at* $t = 0$, *a fraction of* $a_{2\lambda} - a_{1\lambda}$ *flee at* $t = 1$, *a fraction of* $a_{3\lambda} - a_{2\lambda}$ *flee at* $t = 2$, *types above* $a_{3\lambda}$ *fight at* $t = 2$. *Where:*

$$a_{1\lambda} = \frac{\lambda[4c^2(v_H+(2-\lambda)v_L)+kv_L((4\lambda-6)c+k+v_H)+4cv_Hv_L(\lambda-1)]}{\lambda[2cv_H(k+v_H+\lambda v_L)+4c^2v_H+v_L(4c^2(2-\lambda)+(\lambda-1)2ck)]+(v_H+v_L)(k+v_H)^2} \tag{D.22}$$

$$a_{2\lambda} = \frac{\lambda[4c^2(v_H+(2-\lambda)v_L)+2cv_H(k+v_H+\lambda v_L)+2ckv_L(\lambda-1)]}{\lambda[2cv_H(k+v_H+\lambda v_L)+4c^2v_H+v_L(4c^2(2-\lambda)+(\lambda-1)2ck)]+(v_H+v_L)(k+v_H)^2} \tag{D.23}$$

$$a_{3\lambda}=\frac{\lambda[4c^2(v_H+(2-\lambda)v_L)+2cv_H(k+v_H+v_L)+2ckv_L(\lambda-1)+v_L((k+v_H)^2+2cv_H(\lambda-1))]}{\lambda[2cv_H(k+v_H+\lambda v_L)+4c^2v_H+v_L(4c^2(2-\lambda)+(\lambda-1)2ck)]+(v_H+v_L)(k+v_H)^2} \tag{D.24}$$

**Proof.** Using the same method stated above, we obtain threshold values, $a_{1\lambda}$, $a_{2\lambda}$ and $a_{3\lambda}$. To satisfy the condition that: $0 \le a_{1\lambda} \le a_{2\lambda} \le a_{3\lambda} \le 1$, we have the following constraints:

$$c \ge \frac{kv_L}{2[v_H + (2-\lambda)v_L]} \tag{D.25}$$

Combined with $c \le \frac{v_H+k}{4}$, we derive that:

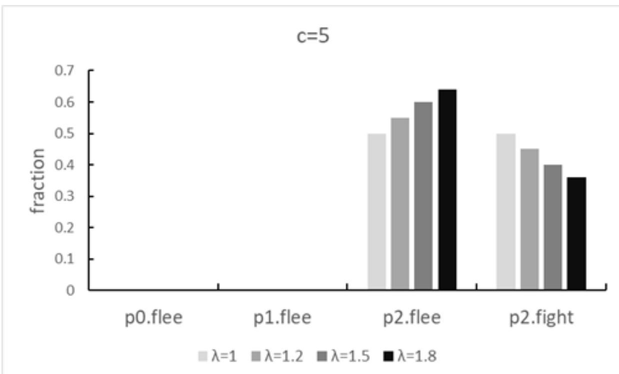
$$\lambda \le \min\{\frac{v_H}{v_L}, \frac{2k}{v_H+k} - \frac{v_H}{v_L}\} + 2 \tag{D.26}$$

Therefore, using it the same way as stated in Appendix C, we can show that when $\frac{kv_L}{2[v_H+(2-\lambda)v_L]} \le c \le \frac{v_H+k}{4}$ and $\lambda \le \min\{\frac{v_H}{v_L}, \frac{2k}{v_H+k} - \frac{v_H}{v_L}\} + 2$, such gradual equilibrium exists. ■

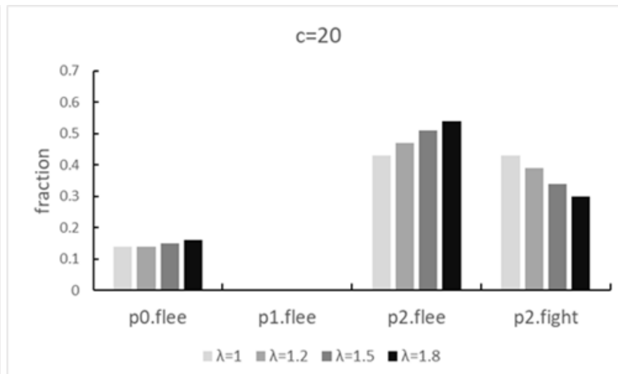
Specifically, suppose $v_H = v_L = k = 100$ and $c = 40$. If general equilibrium exists, then as players become more loss averse, more fractions will flee at $t = 0$, less fractions will flee (fight) at $t = 2$, and the fractions of players fleeing at $t = 1$ will first increase and then decrease.[15]

Use the results stated in above, we calculate the predicted fractions of players who flee (fight) at each period, when $v_H = v_L = k = 100$; the degree of loss aversion ($\lambda$), equals $1, 1.2, 1.5$ and $1.8$; and the waiting cost, $c$ equals 5, 20 and 40 respectively. Figure D.1 demonstrates the results. It is obvious that as players become more loss averse, they are more likely to flee early and less likely to fight. If we compare across different treatments, we find that raising waiting costs can effectively reduce the frequency of battles, shorten the duration of contests, and raise the fraction of weaker players who flee first.

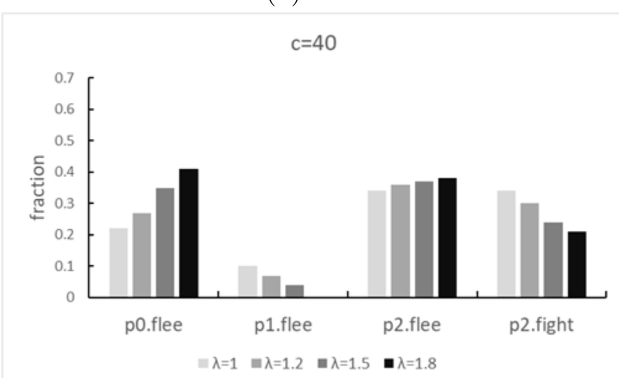
[15]Due to the complicity in deriving the comparative statistics results in the general case, here we focus only on the case where the values of parameters equal our experiment's setting.




(a) c=5




(b) c=20




(c) c=40

Figure D.1: Predicted fractions of players who flee (fight) at different periods, under different degrees of loss aversion and different waiting costs.

# Appendix E. Experimental instructions (C40 treatment)

Welcome to this experiment on decision-making. You've already earned a $5 show-up bonus. We thank you for your participation!

Please read the following instructions carefully. If you have any questions at any time, please raise your hand. An experimenter will assist you privately.

Today's experiment consists of 2 parts. At the beginning of each part, you will receive new instructions. You will spend most time on first part. Your decisions in one part have no influence on the proceedings or earnings of the other part.

Your decisions and those of other participants will determine your earnings. Your earnings will be paid to you privately at the end of today's session. Your earnings in Part 1 will be denoted in points. At the end of the experiment, each point that you earned will be converted into 6 US cents (1 point = 0.06 US dollar).

**Part 1: Decisions and Payoffs**

This part consists of 40 rounds. In each round you will be anonymously and randomly paired with another participant (your opponent) in the laboratory. Therefore, in each round you will (most likely) be paired with a different participant from the previous round. You will never learn with whom you are paired. At the end of the experiment, one of the 40 rounds of Part 1 will be randomly selected for payment. Your earnings for Part 1 will be completely determined by what happened in this round.

**Your choices**

There will be 3 periods in each round. In each period, there will be a clock that counts down from 10 seconds to 0. In Period 1 and 2, you and your opponent have the option to choose to flee at any point during the countdown, you have also the option to choose wait and thereby postpone your decision. If time runs out and you haven't made any decisions, your choice will be automatically counted as wait. In Period 3, you and your opponent have the option to choose fight or flee at any point during the countdown. If time runs out and you haven't made any decisions in this period, you are forced to make a decision to fight or flee, without knowing what your opponent chooses. So please pay attention to the screen.

**Possible Outcomes**

If both of you decide to flee in the same period, or one of you decide to flee while the other is still waiting to make a decision, an escape occurs.

If both of you decide to wait in Period 1 (2), then it proceeds to Period 2 (3).

If both of you decide to wait in the first two periods, and fight in Period 3, a battle occurs.

If one of you decide to flee and the other decides to fight in Period 3, a battle occurs 50% of time and an escape occurs the other 50% of time.

The possible scenarios are illustrated in the figure D.2 below:

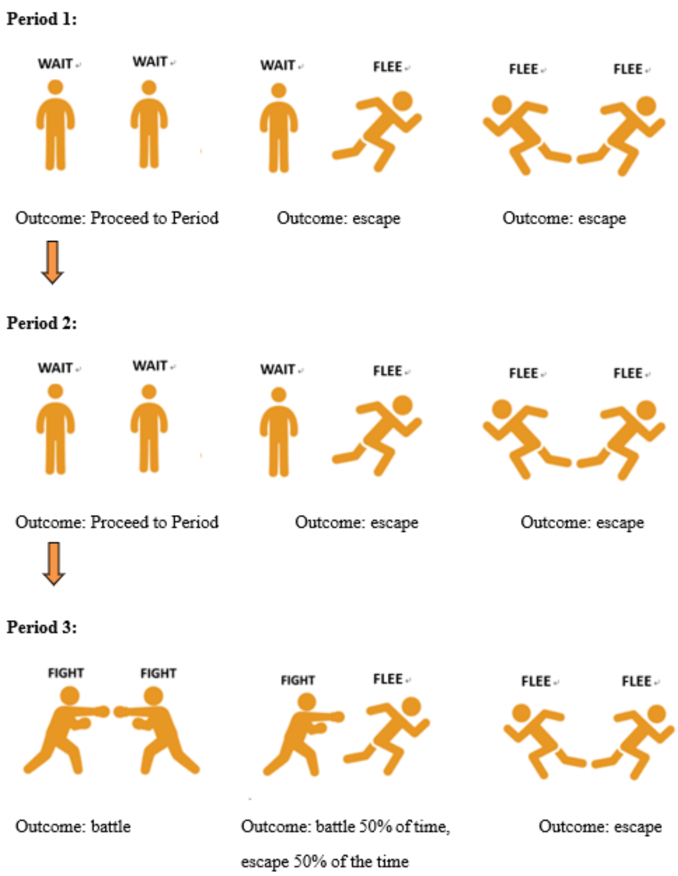



Figure D.2: Possible outcomes.

### Possible earnings

In each round, you will be given a starting capital of 190 points. Any earnings or losses you make in this round will be added to or subtracted from this starting capital.

If a battle occurs between you and your opponent, the participant with higher strength will earn the prize of 100 points and the participant with lower strength will lose 100 points.

If an escape occurs, the participant who flees guarantees her-/himself 0 point (instead of winning or losing 100 points depending on her/his strength). The participant that chooses not to flee then automatically receives the prize of 100 points plus an additional 100 points (thus earning 200 points in total). If both of you flee in the same period, you will split the earnings of the escape. That is, both of you will earn 100 points.

Waiting is costly: If both of you choose to wait in Period 1, then it will cost each of you 40 points for waiting; Then if both of you still choose to wait in Period 2, it will cost each of you 40 more points for waiting (that's 80 points in total).

**Your strength**

At the start of each round, each participant will be informed of her/his strength in that round.

- A participant's strength will be a random whole number between 0 and 1000 (0 and 1000 are also possible). Each of these numbers is equally likely to be chosen.
- In each round, every participant is assigned a new (and independent) strength. Therefore, different participants (most likely) have different strengths in a round, and the same participant (most likely) has different strengths across rounds.
- At the start of a round, each participant is only informed about her/his own strength.
- It is very unlikely that you and your opponent have the same strength, but if this happens, it will be randomly determined who is the stronger player.

**Information at the end of a round**

At the end of a round, each participant will be informed of the outcome (escape/battle), the opponent's strength and the resulting payoffs.

**Examples**

Let's look at several examples: Consider a round in which your strength is 889 and your opponent has a strength of 181.

- If you choose to wait in Period 1, and your opponent chooses to flee in Period 1. Then, this round ends as an escape in Period 1 and the cost of waiting is 0 point:
  - You choose not to flee, your earnings in this round are 200 points.
  - Your opponent chooses to flee, your opponent's earnings in this round are 0 point.
  - Plus the 190 points initial capital, your payoffs in this round are 390 points, and your opponent's payoffs are 190 points.
- If in Period 1 both you and your opponent choose to wait. And in Period 2, you choose to flee and your opponent chooses to wait. Then, this round ends as an escape in Period 2 and the cost of waiting is 40 points.

- You choose to flee, your earnings in this round are 0-40 = -40 points.
- Your opponent chooses not to flee, your opponent's earnings in this round are 200-40 = 160 points.
- Plus the 190 points initial capital, your payoffs in this round are 150 points, and your opponent's payoffs are 350 points.

- If in Period 1 and 2 both you and your opponent choose to wait. And in Period 3, you choose to fight and your opponent chooses to flee. Then, this round has 50% chance to end with an escape, and 50% chance to end with a battle in Period 3. The cost of waiting is 80 points.

  **If the round ends with an escape:**

  - You choose not to flee, your earnings in this round are 200-80 = 120 points.
  - Your opponent chooses to flee, your opponent's earnings in this round are 0-80 = -80 points.
  - Plus the 190 points initial capital, your payoffs in this round are 310 points, your opponent's payoffs are 110 points.

  **If the round ends with a battle:**

  - You have higher strength, you win the battle, and your earnings in this round are 100-80= 20 points.
  - Your opponent has lower strength, she/he loses the battle, and her/his earnings in this round are -100-80 = -180 points.
  - Plus the 190 points initial capital, your payoffs in this round are 210 points, your opponent's payoff is 10 points.

This is the end of the instructions. You will be given a short quiz to ensure that you understand the instructions. Once you complete the quiz successfully, you will proceed to the experiment.

# Appendix F. Equilibria when allowing players to fight at any period

**Waiting equilibrium**
We first prove that when $0 \leq c \leq \frac{kv_L}{4(v_H+v_L)}$, given others' strategy, any type $a_j \in (\hat{a}, 1]$ has no incentive to fight earlier.

If $a_j$ fights at $t = 2$, the expected payoff is:

$$\hat{a} \times \frac{k}{2} + a_j(v_H + v_L) - v_L - 2c \tag{F.1}$$

If $a_j$ fights at $t = 1$, the expected payoff is:

$$\hat{a}(v_H - c) + (a_j - \hat{a})(v_H - c) + (1 - a_j)(-v_L - c) \tag{F.2}$$

If $a_j$ fights at $t = 0$, the expected payoff is:

$$\hat{a}v_H + (a_j - \hat{a})v_H - (1 - a_j)v_L \tag{F.3}$$

Since $(F.2) < (F.3)$, we just need to compare the payoff between (F.1) and (F.3). It is obvious that when $0 \leq c \leq \frac{kv_L}{4(v_H+v_L)}$, $(F.3) \leq (F.1)$. Therefore, types above $\hat{a}$ have no incentive to fight earlier. Given their strategy, using the results in Appendix A, we can derive that types below $\hat{a}$ have no incentive to flee earlier. Which means the waiting equilibrium stated in Proposition 1 still holds when allowing players to fight at any period.

**Jump equilibrium**
Next, we prove that when $\frac{kv_L}{4(v_H+v_L)} \leq c \leq \frac{kv_L}{2(v_H+v_L)}$, given others' strategy, any type $a_j \in (\overline{a_2}, 1]$ has no incentive to fight earlier.

If $a_j$ fights at $t = 2$, the expected payoff is:

$$(\overline{a_1} + \overline{a_2}) \times \frac{k}{2} + a_j(v_H + v_L) - v_L - 2c(1 - \overline{a_1}) \tag{F.4}$$

If $a_j$ fights at $t = 1$, the expected payoff is:

$$(\overline{a_2} - \overline{a_1})(v_H + \frac{k}{2} - 2c) + (a_j - \overline{a_2})(v_H - 2c) + (1 - a_j)(-v_L - 2c) \tag{F.5}$$

If $a_j$ fights at $t = 0$, the expected payoff is:

$$\overline{a_1}(v_H + \frac{k}{2}) + (\overline{a_2} - \overline{a_1})v_H + (a_j - \overline{a_2})v_H - (1 - a_j)v_L \tag{F.6}$$

When $c \leq \frac{kv_L}{2(v_H+v_L)}$, $(F.5) \leq (F.4)$; and when $c \leq \frac{kv_L}{(v_H+v_L)}$, $(F.6) \leq (F.4)$. Therefore, types above $\overline{a_2}$ have no incentive to fight earlier. Given their strategy, using the results in Appendix B, we can derive that types below $\overline{a_2}$ have no incentive to deviate. Which means the jump equilibrium stated in Proposition 2 still holds when allowing players to fight at any period.

**Gradual equilibrium**

Finally, we prove that when $c > \frac{kv_L}{2(v_H+v_L)} \times \frac{k+v_H}{v_H}$, given others' strategy, any type $a_j \in (a_3, 1]$ has the incentive to fight earlier.

If $a_j$ fights at $t = 2$, the expected payoff is:

$$a_1(v_H+k)+(a_2-a_1)(v_H+k-c)+(a_3-a_2)(v_H+\tfrac{k}{2}-2c)+(a_j-a_3)(v_H-2c)-(1-a_j)(v_L+2c) \tag{F.7}$$

If $a_j$ fights at $t = 1$, the expected payoff is:

$$a_1(v_H+k)+(a_2-a_1)(v_H+\tfrac{k}{2}-c)+(a_3-a_2)(v_H-c)+(a_j-a_3)(v_H-c)-(1-a_j)(v_L+c) \tag{F.8}$$

We can find that $(F.7) = (F.8) + (a_3 - a_1)\frac{k}{2} - (1 - a_2)c$. When $c > \frac{kv_L}{2(v_H+v_L)} \times \frac{k+v_H}{v_H}$, $(a_3 - a_1)\frac{k}{2} - (1 - a_2)c < 0$. $a_j$ will be strictly better off if fleeing earlier. Therefore, the gradual equilibrium stated in Proposition 3 will not hold if the waiting cost is sufficiently large in relation to the deterrence value.