When to Confront: The Role of Patience*

Alvaro Sandroni† Can Urgun‡

May 2, 2017

Abstract

This paper examines the effects of patience on ordinary conflicts such as divorce, price wars and commercial litigation. Players optimally decide when, if ever, to start a destructive confrontation. In the unique equilibrium, there is a tight connection between patience, aggressiveness and strength. In particular patience may lead to immediate confrontation (the most inefficient outcome). This inefficiency is caused by preemptive moves that deny option values to the opponent.

*We thank Bruno Strulovici, Sandeep Baliga, Jin Li, Ehud Kalai, Larbi Alaoui, Weifeng Zhong and Yeliz Kacamak for useful comments and suggestions. We also thank two anonymous referees and the editor Andrew Postlewaite for several comments that improved our proofs and the intuition behind our results and the overall quality of the exposition.

†Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA.

‡Department of Managerial Economics and Decision Sciences, Kellogg School of Management, Northwestern University, Evanston, IL 60208, USA.

"Warfare is the greatest affair of state, the basis of life and death, the way to survival or extinction. It must be thoroughly pondered and analyzed." Sun Tzu, Art of War.

1 Introduction

Conflict is a regular occurrence for almost everyone. People divorce, engage in commercial litigation, political infights, and professional disputes. The same applies to firms. They engage in price wars, hostile take-overs and aggressive marketing campaigns. While omnipresent, direct confrontations do not occur exogenously. Instead, someone must start it and the timing of the confrontation influences the resolution of the conflict. For example, consider two firms facing random stock valuations and deciding when, if ever, to engage in an aggressive take-over. The timing of the take-over is critical for its resolution. Hence, a basic question is when to start a confrontation.

We examine a two-player infinite horizon model where, at every period, each player decides whether to engage the other side. If one player attacks, the game ends with one side as the victor. Otherwise, they both face each other the next day. The odds of victory changes every period. Confrontations are destructive: the disutility of the loser outweighs the gains of the victor. So, this is a model of adversaries facing changing opportunities and assessing whether the time has come to attack. A decisive factor in the resolution of the conflict is the timing of the confrontation.

In this model, there is a unique equilibrium. Under suitable conditions, as players become more patient they become unambiguously more aggressive (i.e., they attack sooner). Patient players do not wait: when completely patient they become so aggressive that the confrontation starts *immediately*. This is the most inefficient outcome in this game. Thus, patience does not allow for efficiency. It precludes it and ensures inefficiency.

The players in this game start in a potential conflict and may decide whether to engage in an actual conflict. This is unlike repeated games where players are from the start in a conflict they cannot escape from (i.e., they cannot modify the game they are in). The results are in stark contrast: equilibrium payoffs are uniquely identified in our model whereas as several folk theorem have shown, there are multiple equilibrium payoffs in repeated games. Patience is detrimental to efficiency in this model whereas it is essential to achieve efficiency, at least potentially, in repeated games.

The basic trade-off in repeated games is whether to take a costly action that may induce the adversary to behave more favorably in the future. Hence, patience is necessary to induce mutually beneficial behavior. In this model, the basic trade-off is that waiting delivers an option value (i.e., the opportunity to start the actual conflict in more favorable terms). However, waiting also delivers the opponent an option value. So, the effects of patience in this game are, a priory, unclear. The option of waiting for a better opportunity to attack is more valuable to a more patient player. On the other hand, more patient players suffer greater disutility from delivering the opponent this option value. The formal analysis reveals that the results of patience are, in fact, unambiguous. The benefit of the former option value is dominated by the loss associated with the latter.

The results mentioned above assumed parity of strength (i.e., the odds of victory are generated, ex-ante, in a way that does not favor either player) and this is commonly known. Now consider the case where there is uncertainty about relative strength. This uncertainty is a potential deterrent for confrontations. So, one side is intrinsically stronger, but, initially, it is unknown which one. After several opportunities arise, players can learn about their relative strength. Here, the basic trade-off continues at a higher level. If a player does not stop the learning process (i.e., attacks) then this player obtains valuable knowledge, but this also grants the adversary the same knowledge.

This process is closely related to a game-theoretic bandit problem where players learn by experimentation. In this model, the bandits are negatively correlated, good news for one player is bad news for the other. However, the main novelty, compared to the existing game-theoretic literature on social experimentation (see, for example, Bolton and Harris (1999) and Rosenberg, Solan, and Vieille (2007)), is that one player stopping or continuing the learning process has more than an informational effect on the other player: it is also directly payoff-relevant for the other player.

Our main results are as follows: Uncertainty about relative strength does not ensure peace. This inefficiency cannot be resolved by patience, no matter how great. Moreover, if a player is revealed to be intrinsically stronger, then, for this player, greater patience leads to less aggressiveness. A weaker, more patient player attacks sooner. With perceived superior strength, it pays to wait before striking. In more vulnerable conditions, preemptive strikes are required. These results identify the key conditions under which patience is conducive to aggressiveness: relative strength is the key mediator on the relationship between patience and aggressiveness.

Our basic model is a stopping game. Stopping games are a natural framework for a wide variety of economic phenomena. While related to bargaining models and to repeated games, they are not nearly as used as these models. This is, in part, due to the fact that beyond existence of equilibria and, sometimes their value, little is known about stopping games. We characterize equilibria and more importantly obtain comparative statics results even when social experimentation is added to the model. We hope that our novel methods are a first step towards a widespread use of stopping games.

After a brief literature review, Section 2 introduces the game of conflict. The case of known strength is in Section 3. The learning model is in Section 4. Section 5 concludes. Proofs are in the appendix.

1.1 Relation with the Existing Literature

The inefficiency of conflict has been widely studied in the political economy literature. Garfinkel and Skaperdas (2007) is an extensive survey, but see also Fearon (1995), Powell (1999), Powell (2004), Powell (2006), Dal Bó, Hernández, and Mazzuca (2015). The political economy literature focuses on the understanding of the existence of inefficient inter-state wars, and how institutions can affect and prepare for them. We focus on everyday conflicts, the question of when to start a confrontation and the relation between patience, strength and aggressiveness. In addition, most of the political economy literature uses static models. There are notable exceptions. For example, Powell (1993) and Acemoglu and Robinson (2001) consider dynamic elements in a model fundamentally different from ours.

Our setting eventually divides a constant sum and so, can be related to bargaining models. We do not survey this large literature. Serrano (2007) provides an extensive survey, but see also Osborne and Rubinstein (1990) and Roth (1985). In Abreu and Gul (2000) and Compte and Jehiel (2002), players decide when to take action (concede). This gives higher payoffs to the other side. In our model, acting is to fight which never yields higher expected payoffs to the other side. Bargaining typically has a war of attrition structure. Our model does not have a war of attrition structure because the status-quo is efficient (peace) and the decisive action leads to an inefficient confrontation.

Single player stopping problems is also a large literature that we do not survey here. The current state of the art on the effects of discount rates in stopping problems is the work by Quah and Strulovici (2013). Under suitable conditions, our methods extend their decision theoretic results to a game theoretic framework.

The literature on stochastic games typically deals with more general problems such as existence of equilibria and sometimes, their value (see, for example, Shapley (1953), Mertens and Parthasarathy (1991), Solan and Vieille (2002)) with a main focus in irreducible games. A related strand in stochastic games considers games with absorbing states, where the game is *reducible* (when players' actions may change the game permanently). In particular Wiseman (2017) studies a game of repeated oligopoly, where firms might enter a price war and be forced to exit the market. Unlike our game, in Wiseman (2017) players have a strong incentive to induce the other players to exit the game. As long as players haven't exited, the oligopoly setting allows for intertemporal trade of continuation payoffs to avoid a price war. Even though there are no issues of dimensionality in the stage game, an anti-folk theorem is also obtained, in the sense that collusion is not sustainable. Our results are obtained under different conditions. In our setting when one player stops the game the other player typically receives low payoffs and cannot continue the game. Despite the differences in settings, the anti-folk theorems highlights the importance of reducibility in stochastic games. In fact our anti-folk result which arises due to reducibility is also in line with the sharp observation in Sorin (1986) where a stochastic discounted game with absorbing states (hence reducible) has a unique equilibrium payoff regardless of the discount factor.

Within stochastic games we build upon the work in stopping games (see Dynkin (1969), Neveu (1975), Yasuda (1985), Rosenberg, Solan, and Vieille (2001), Szajowski (1993), Shmaya and Solan (2004), Ekstrom and Villeneuve (2006), Ohtsubo (1987)). Stopping games are reducible. In many reducible games, it is difficult to obtain much more than existence of equilibria. We build upon Ohtsubo (1987) for characterizing equilibria and we utilize either direct calculations or the single player techniques of Quah and Strulovici (2013) on an extension of the Snell (1952) Envelope to achieve our comparative statics. The Snell envelope is a general, but mostly intractable, solution to stopping problems. We side step intractability by utilizing martingale techniques, focusing on ordering payoffs.

As mentioned in the introduction, our learning model resembles a social bandit problem where players learn by experimentation (see Bolton and Harris (1999), Bergemann and Välimäki (1996) and Rosenberg, Solan, and Vieille

(2007)). Klein and Rady (2011) consider negatively correlated bandits, where good news for one player is bad news for the other. However, in our model, when one player stops so must the other. Hence, in our model, the actions of one player are not only informationally relevant to the other player and, in this sense, indirectly payoff-relevant and but also directly payoff-relevant as they may change the game itself. This direct payoff relevance significantly changes the strategic interaction and required us to develop a novel approach. Here, we build upon techniques developed by Ohtsubo (1987) for characterizing the equilibria.

2 Basic Model and Notation

2.1 The game

There are two players 1 and 2. At each period either player either starts a confrontation (fight) or not. If neither player engages in a fight, they get 0 payoffs that period and the game continues. If one player starts a confrontation, the opponent cannot avoid it and the game ends with one player defeated. The winner gets utility $v > 0$, and the loser utility $-l$, $l > 0$, $l > v$. So, confrontations are destructive: the payoff of victory (v) is less than the disutility of defeat (l).[1] In case of a confrontation at period t, player 1 is the victor with probability p_t. At the beginning of each period, both players observe player 1's probability of winning (hence player 2's probability of winning which is, $1 - p_t$). Thus, players choose whether to start a confrontation after observing the odds of victory. The probability p_t of the player 1 winning a confrontation is produced by a random variable $\tilde{p}_t$ with continuous probability density and full support in $[0, 1]$. Both players discount future payoffs with a common, constant discount factor $\beta \in (0, 1)$. Both players have a common belief over the distribution of $\tilde{p}_t$.

Players 1 and 2 are in a potential conflict. The key question is whether and if so when one of them will strategically escalate the conflict in an irreversible payoff-relevant move. For example, consider two companies facing idiosyncratic shocks to costs and demands and deciding when to start a confrontation which can be in the form of a price war, a break off from an informal collusion or an aggressive marketing campaign. In the case of a troubled marriage, the open confrontation may occur when one of the par-

[1] Otherwise, the confrontation happens in the first period.

ties files for divorce and dispute, say, child-custody. A similar case occurs in commercial litigation, where the evidence/contractual claims can be changing before it goes to court. So, going to trial is confronting the opponent. In many, but not all, cases there is a natural point that can be construed as the start of the confrontation at its earnest. A significant limitation in our setting is that once the confrontation starts there are no more strategic interactions. While this limits the applicability of our model, it helps focus our results on the issue of when to start the conflict.

2.2 Strategies

The history at period t, $h_t = (p_0, p_1, \ldots, p_t)$, is the sequence of probabilities of winning for player 1 up to period t. Given that the game ends if a player starts a confrontation, we implicitly assume that no player has started a confrontation if a history is relevant for decision making. The set of all histories generate a growing sequence of σ-algebras, $\sigma(h_t)$ for the process $\{\tilde{p}_t\}$ or equivalently a filtration for $\{\tilde{p}_t\}$, denoted by $\{\mathcal{F}_t\}$. The probability triple (i.e., the filtered probability space) is given by $(\Omega, \{\mathcal{F}_t\}, \mathrm{P})$, where Ω is the set of all histories, i.e., $\Omega = \bigcup_{t=1}^{\infty} [0,1]^t$, and P is the probability measure over Ω. Let E be the expectation operator associated with P. Unless otherwise noted, we assume that $(\tilde{p}_{t+1} | \mathcal{F}_t)$ has a continuous density with full support on $[0,1]$ for all t and its realizations are denoted p_t. For consistency, with a slight abuse we assume time starts at the end of period -1 with a trivial algebra, right before p_0 is realized.

A pure strategy takes finite histories as input and returns, as output, the choice of whether to start a confrontation. We formalize pure strategies (in a way that is common in stopping games) as follows:

Definition 1. *A pure strategy is a stopping time τ for the filtration $\mathcal{F}_t$.*[2]

So, a pure strategy determines when to confront, depending on the current and (perhaps) past odds of victory. Given a history at period t, a player starts a confrontation at this history if and only if $\tau = t$. For example, consider the *hitting strategy with a fixed threshold* $\tau^1 = \inf\{\ t \geq 0 \mid p_t \geq \bar{p}\}$. In this strategy, player 1 starts the confrontation when the current odds of victory are greater than $\bar{p}$. Let τ^i be player i's pure strategy, $i = 1, 2$.

[2] Let $\mathbb{N}$ be the set of natural numbers. Given a triple $(\Omega, \{\mathcal{F}_t\}, \mathrm{P})$, a stopping time is a random variable $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ such that $\mathrm{P}(\tau = \mathrm{t})$ is $\mathcal{F}_t$ measurable.

Given a pure strategy profile $\tau = (\tau^1, \tau^2)$, the game ends at $\tau^1 \wedge \tau^2 = min(\tau^1, \tau^2)$. So, the overall payoff to players is given by

$$U^1(\tau) = \sum_{t=0}^{\infty} \beta^t \mathrm{P}(\tau^1 \wedge \tau^2 = t) E(\tilde{p}_t(v+l) - l | \tau^1 \wedge \tau^2 = t),$$
$$U^2(\tau) = \sum_{t=0}^{\infty} \beta^t \mathrm{P}(\tau^1 \wedge \tau^2 = t) E(-\tilde{p}_t(v+l) + v | \tau^1 \wedge \tau^2 = t).$$

Players' payoffs can also be written in a more compact form as follows:

$$U^1(\tau) = E(\beta^{\tau^1 \wedge \tau^2}(\tilde{p}_{\tau^1 \wedge \tau^2}(v+l) - l)),$$
$$U^2(\tau) = E(\beta^{\tau^1 \wedge \tau^2}(-\tilde{p}_{\tau^1 \wedge \tau^2}(v+l) + v)).$$

2.2.1 Equilibrium

Definition 2. *A pair of stopping times $\bar{\tau} = (\bar{\tau}^1, \bar{\tau}^2)$ is an equilibrium if*

$$U^1(\bar{\tau}) = \sup_{\tau^1} E(\beta^{\tau^1 \wedge \bar{\tau}^2}(\tilde{p}_{\tau_1 \wedge \bar{\tau}^2}(v+l) - l))$$
$$U^2(\bar{\tau}) = \sup_{\tau^2} E(\beta^{\bar{\tau}^1 \wedge \tau^2}(-\tilde{p}_{\bar{\tau}^1 \wedge \tau^2}(v+l) + v))$$

That is, a pair of stopping times $\bar{\tau} = (\bar{\tau}^1, \bar{\tau}^2)$ is an equilibrium if there are no profitable unilateral deviations from them. So, in equilibrium, each player best responds to the opponent's plans. In this definition only pure strategies are considered. In the appendix, we consider mixed strategies and show that even if mixed strategies are allowed, the equilibrium payoff set remains the same.

Remark 1 (Mutually assured destruction)**.** *The pair of strategies where players start a confrontation at every history is always an equilibrium.*

Once a player commits to fighting the other one can not back off. Hence, *mutually assured destruction* is always an equilibrium because if the opponent always attacks then anything is a best response. However these strategies are weakly dominated as players start confrontations even when they have almost no chance of winning.[3] For the remainder of the paper, we disregard this equilibrium and focus on non-dominated strategies.

Definition 3. *τ^i is a fighting strategy if $P(\tau^i < \infty) = 1$*

So, there is no lasting peace in fighting strategies. If fighting strategies are employed in equilibrium, we call it a *fighting equilibrium.* Due to the destructive nature of the confrontation, any fighting equilibrium is Pareto dominated by indefinite peace.

Remark 2 (Equilibria only in Fighting Strategies)**.** *Equilibria has to be in fighting strategies.*

When facing a destructive confrontation where span of the odds of victory is broad enough (i.e., the odds of victory can take values anywhere within bounds close to 0 and 1), fighting has to eventually occur. At one point, victory will be near certain for one side and that point, the player with favorable odds should start the confrontation. Both players know this and, hence, there is no hope for indefinite peace. This inevitability of fighting is also essential in characterizing the set of equilibrium payoffs as well. Once the confrontation happens, no matter the odds, the payoffs to the players is a constant sum $(v - l)$. The argument is slightly more involved but like a one-shot constant sum game, there is a unique pair of equilibrium values in our game. The uniqueness result can be found in Yasuda (1985) and makes use of a minmax like argument on a value operator, the value characterization is done by Ohtsubo (1987) by a backward induction like argument again using a value operator starting from the moment where the confrontation occurs. We include adapted versions of both results in the appendix for completeness.

[3] Consider the mutually assured destruction strategy of player 1, $\tau^{1,Mad} = \inf\{\ t \geq 0 \mid p_t \geq 0\}$, and consider the alternative strategy $\tau^{1,1/2}$, where player 1 confronts at period 0 if and only if $p_0 \geq 1/2$ and confronts in every other period at any history. For any strategy of player 2, $U^1(\tau^{1,Mad}, \tau^2) \leq U^1(\tau^{1,1/2}, \tau^2)$, with strictly inequality for some strategies of player 2. To see this observe that, in histories such that $p_0 < 1/2$, and $\tau^2 > 0$ the payoff associated with $\tau^{1,Mad}$ is strictly less than $\frac{v-l}{2} < 0$, whereas $\tau^{1,1/2}$ yields a payoff of $\beta\frac{v-l}{2} > \frac{v-l}{2}$. In all other histories, the two strategies produce the same payoffs. By a similar logic any strategy that prescibes certain fighting based on calendar time alone and disregards the current odds are also dominated and we disregard those as well.

2.3 Aggressiveness

Aggressiveness is measured by the tendency to confront early on. Consider player 1's hitting strategy with a fixed threshold of attacking when the odds of victory p_t are above $\bar{p}^1$. The lower the threshold $\bar{p}^1$ the less demanding the odds of victory player 1 requires to start a confrontation. Hence, *the lower the threshold* $\bar{p}^1$ *the more aggressive player 1 is.* Analogously, consider player 2's hitting strategy with a fixed threshold of attacking when player 1's odds of victory p_t are below $\bar{p}^2$ (and, hence, player 2's odds of victory $1-p_t$ are above $1-\bar{p}^2$). The higher the threshold $\bar{p}^2$ the less demanding the odds of victory player 2 requires to start a war. Hence, *the higher the threshold* $\bar{p}_2$ *the more aggressive player 2 is.* Therefore, in the special case of hitting strategies, aggressiveness can be defined by the thresholds $\bar{p}^1$ and $\bar{p}^2$. We also consider a definition of the impact of patience on aggressiveness that does not rely on hitting strategies. This broader definition of aggressiveness is particularly useful in the learning model. Letting $\tau^i(\beta)$ denote the equilibrium stopping time for player i associated with the discount factor β;

Definition 4. *For player i's pure strategies $\tau^i(\beta), \tau^i(\hat{\beta})$. A more patient player is more aggressive at time t if for any $s \geq t$*

$$P(\tau^i(\hat{\beta}) \leq s|\mathcal{F}_t) \geq P(\tau^i(\beta) \leq s|\mathcal{F}_t)$$

whenever $\beta \leq \hat{\beta}$. Conversely, at time t, a more patient player is less aggressive if this inequality holds with the roles of β and $\hat{\beta}$ reversed.

So, an increase in patience makes the player more aggressive if the strategies with higher discount factor increase the odds of starting the confrontation within any finite time frame.

2.4 Strength

In this model, the stronger party is not assured to win a fight, but strength delivers better odds of victory. We formalize strength as follows:

Definition 5 (Expectations of Imminent Confrontation)**.**

$$w_t^1 = vp_t - (1-p_t)l,$$
$$w_t^2 = -p_t l + (1-p_t)v$$

So, w_t^1 is the expected payoff of player 1 at period t, if a confrontation starts in that period. Analogously, w_t^2 is the expected payoff of player 2 at period t, if a confrontation starts in that period. These values are exogenously given and are not related to the strategies chosen by the players.

Definition 6 (Intrinsic Strength)**.** *Player i is intrinsically strong if for all t, $E(w_t^i) > 0$.*

So, player i is intrinsically strong if, at any period, ex-ante (i.e., before the odds of victory are realized) the expected payoff of a confrontation is greater than the payoff obtained in peace. It is not possible for both sides to be strong, but it is possible that neither is (e.g., when there is parity in power).

3 Known Strength

We first examine a special case where both players have equal intrinsic strength and they know it. This determines the extent to which parity in strength and the destructive nature of confrontations can ensure peace and hence, efficiency.

Theorem 1. *Assume that player 1's probability of victory p_t is uniform i.i.d., over $[0,1]$ and this is commonly known. There is a unique set of sustainable payoffs in equilibrium. These payoffs are supported by fighting strategies, given by hitting strategies with fixed thresholds $\bar{p}_1 < 1$ and $\bar{p}_2 > 0$:*

$$\bar{\tau}^1 = \inf\{t \geq 0 \mid p_t \geq \bar{p}^1\} \text{ and } \bar{\tau}^2 = \inf\{t \geq 0 \mid p_t \leq \bar{p}^2\} \text{ where}$$

$$\bar{p}^1 = \frac{1}{2} + \frac{1}{4\beta} - \frac{\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}{4\beta(l+v)}$$

$$\bar{p}^2 = \frac{1}{2} - \frac{1}{4\beta} + \frac{\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}{4\beta(l+v)}$$

So, parity of power does *not* ensure peace. Theorem 1 shows that the destructive nature of confrontations is not a sufficient deterrent either. Once the odds of victory tilts sufficiently towards one side, the confrontation starts. That is, the equilibrium is in fighting strategies and, hence, inefficient.

The closed-form solutions in Theorem 1 are obtained through a combination of straightforward calculations along with the existence of Markovian equilibria in Mertens (2002), and a key result in Yasuda (1985) to ascertain that there is a unique pair of equilibrium payoffs. We show that, in our model, any Markovian equilibria is also in hitting strategies with fixed thresholds. Hence, there exists an equilibria in hitting strategies with fixed thresholds. With some algebra, we obtain closed-form solutions for such an equilibrium.

The intuitive reason for a unique pair of sustainable values is reducibility. After a confrontation the game ends, thus there is no way to make subgame perfect threats that would enlarge the set of equilibrium payoffs. The fighting strategies identified here are not necessarily the only pair of strategies that sustain the equilibrium payoffs, but any other pair of equilibrium strategies (possibly mixed) sustain the same payoffs. Thus, we focus on fighting strategies given by hitting strategies with fixed thresholds because, with them, we can calculate an equilibrium in closed-form. This permits direct comparative statics results. Paradoxically, we now show that the destructive nature of confrontations makes patience conducive to hasty engagements.

Corollary 1. *In equilibrium, both players become more aggressive as the discount factor increases. A confrontation arises almost immediately as players become fully patient (i.e.,* $\bar{p}^1$ *and* $\bar{p}^2$ *converge to* 0.5 *as* β *goes to* 1*).*

The relationship between patience and aggressiveness is unambiguous. More patient players are more aggressive. As the players become completely patient, with near certainty, the confrontation starts *immediately.* This follows even if the odds of victory in the first period are such that the confrontation is ex-ante damaging *for both sides.*

Corollary 1 is seemingly in contrast with commonsensical ideas on the relationship between patience and aggressiveness. An appeal to patience is often understood as an appeal to peace and tranquility, but seldom, if ever, to aggressivity. Indeed forgoing an opportunity to attack delivers an option value. It gives the player a chance to confront the opponent in better terms in the future. More patient players obtain greater discounted utility from this option value. However, forgoing an opportunity to attack can be fatal because it also grants enemies the option value of starting the confrontation under better conditions for them. This discounted disutility is greater if the player is more patient. Overall, given that confrontations are destructive, the incremental disutility of granting an option value to the opponent surpasses the additional benefit from obtaining an option value. More patient players have a greater incentive to preempt the opponent and start the conflict earlier since the discounted expected disutility of future destructive confrontations are enhanced. Hence, they are more aggressive.

The setting of theorem 1, features an i.i.d. distribution and parity in strength. In this sense, this is the closest setting to a repeated game, permitted by our framework. Yet, the results are strikingly different from those in repeated games. Unlike folk theorems, there is a unique pair of equilibrium payoffs. More patient players fight sooner, diminishing expected payoffs to a minimum as they become more patient. So, patience does not enable players to approach Pareto efficient behavior (which would be indefinite peace). Instead, it leads to the most extreme inefficiency. Unlike repeated games, where patience is instrumental for welfare, patience leads to an immediate destructive confrontation in this game.

The effect of patience in parity of strength can be reinforced by the idea that, in equilibrium, the increased levels of aggressiveness hold even if we fix the discount of the opponent as shown below: Let $\bar{\tau}^i(\beta^1, \beta^2)$ be the equilibrium strategies associated with the different discount factors (under Theorem 1's setting):

Corollary 2. *In equilibrium, for any discount factor β with $\beta^1 = \beta^2 = \beta$ and for any period $t \geq 0$,*

$$\frac{\partial \mathrm{P}(\bar{\tau}^i(\beta^i, \beta^j) \leq t)}{\partial \beta^i} > 0$$

So, in equilibrium, a more patient player is also more aggressive, even when the discount factor of the other player remains the same.

In this section, we made assumptions that were meant to simplify enough the notoriously difficult stopping games so that it becomes sufficiently tractable to obtain close-form solutions and comparative statics results. This allows for insights that we now show are robust in more complex settings. In particular, we relax key assumptions and introduce social experimentation into this model.

4 Unknown Strength and Learning

In a conflict players may not know how comparatively powerful their adversaries are. This uncertainty is a potential deterrent for confrontation and is now captured in a learning model. Players learn about their relative strengths by observing the opportunities arising, while still deciding when to confront. This fits well with everyday conflicts because it allows for underlying trends that need to be learned. For example, in commercial litigation if one side has a more legitimate claim then the evidence that is slowly gathered will tend to favor that side. Similarly, in case of a potential dispute over child custody, there are trends in the parents careers that effect their ability to provide for the child.

As before, there are changing opportunities to engage in a confrontation. If neither side start a confrontation early on, both players observe these opportunities and infer from them the odds of victory in each period and also, in part, the process generating these odds. So, this is a stopping game with social experimentation. Now, waiting is an opportunity to learn the relative strengths and conversely either player attacking concludes the learning period for *both* players.

Each player now faces two different drawbacks and two different benefits in starting a confrontation. Fighting early on forgoes the opportunity to learn more about the relative strength before the confrontation starts. Hence, it forgoes critical information for a proper decision. Moreover, fighting also forgoes the opportunity of starting the confrontation in better terms. On the other hand, fighting early on precludes the opponent to obtain the same opportunities for learning and for determining the timing of the confrontation. So, the basic trade-off is whether the option value from not fighting exceeds the disutility of granting this option values to the opponent. But now this trade-off has the extra complexity coming from the uncertainty over relative strength and the value of experimentation.

So, assume that the i.i.d. process that generates the odds of victory is either coming from a distribution F^1 or F^2. Both distributions have continuous densities $f^1$and f^2 with full support in $[0,1]$. We further assume that $f^1(x) = f^2(1-x)$ and so if a distribution favors one of the players, the other distribution favors the other player in the same way. In the first distribution, F^1, player 1 is *intrinsically strong.* So, in distribution F^2, player 2 is *intrinsically strong.* The players share the same non-degenerate prior.[4] Let μ_t denote the posterior belief, at period t, that player 1 is *intrinsically strong.* In this setting, good news for one player is bad news for the other, like in a negatively correlated bandit setting.

The main novelty in this model is that there are two different ways in

[4]There is a common inital belief about the likelihood of F_1, strictly between 0 and 1.

which the experimentation of one player affects the other player. If a player decides to fight then, like in other game theoretic models of experimentation (e.g. Bolton and Harris (1999) and Rosenberg, Solan, and Vieille (2007)), this has an indirect effect on the payoffs on the other player because the opponent can no longer obtain critical information. However, in addition to the interactions through information, there is also a direct effect on the payoffs of the other player. The payoffs of the opponent changes when the confrontation starts.

Uncertainty about relative strength may make players hesitant to confront their adversaries. Hence, this uncertainty is a potential deterrent for conflict, but, as we show, it is not an effective deterrent. In our model the uncertainty diminishes, but it is never fully resolved in finite time. In spite of this uncertainty the players stop the process in finite time and attack. This follows because players are once again reluctant to give option values to the opponent.

Our main results show that there is a unique sustainable payoff, which is supported by fighting strategies. Hence, the equilibrium is never efficient. Uncertainty about relative strength is not enough to ensure indefinite peace, no matter how patient players might be. Moreover, the model is still tractable enough to deliver clear comparative static results.

Our results are formalized as follows: for player 1, the expected payoff of a confrontation next period is $\mathrm{E}\left(w^1_{t+1}|\mathcal{F}_t\right) = \mathrm{E}\left(w^1_{t+1}|\mu_t\right)$. For player 2, it is $\mathrm{E}\left(w^2_{t+1}|\mathcal{F}_t\right) = \mathrm{E}\left(w^2_{t+1}|\mu_t\right)$. Assume that μ_t is sufficiently high at period t so that $\mathrm{E}\left(w^1_{t+1}|\mu_t\right) > 0$. Then, the belief that player 1 is strong is sufficiently firm so that player 1 has, ex-ante, a positive expected valuation for a confrontation next period (even though the confrontation is destructive). In this case, we say that players believe that player 1 is *intrinsically strong at period t*. The same terminology applies to player 2, with w^2_{t+1} replacing w^1_{t+1}.

Theorem 2. *There is a unique set of sustainable payoffs in equilibrium. These payoffs are supported by fighting strategies. Moreover, if player i is believed to be intrinsically strong at period t, then a more patient player i is less aggressive at period t. If player i is not believed to be intrinsically strong at period t, then a more patient player i is more aggressive.*

So, for a (believed to be) strong player, greater patience leads to less aggressiveness. On the other hand, if a player's strength is not believed to be sufficiently great then higher patience leads to more aggressiveness.

Thus, perception of strength is critical in conflicts. Even when the relative strength must be learned, the perception of strength is a key mediator in the relationship between patience and aggressiveness.

The trade-off between the option value of waiting for both player versus denying both players this option value delivers the critical insight in Theorem 2. A strong player can afford to wait for better future opportunities because the opponent is less likely to benefit from them. This holds even if strength is not known and can only be partially inferred from the histories. It follows that a more patient, and strong, player has less incentives to conduce a preemptive attack. The opposite occurs for a more patient, but weaker, player who has greater incentives to preempt the opponent. Since it is impossible for both players to be relatively strong at the same time our results show the difficulties with peaceful arrangements and the additional role of perceived strength in our game of warfare.

A remarkable quality of Theorem 2 is that rational players account for the entire stream of future payoffs and not just one-period ahead payoffs (as myopic players do). Yet, the signs of one-step look ahead expected values i.e., $\mathrm{E}\left(w^1_{t+1}|\mathcal{F}_t\right)$ and $\mathrm{E}\left(w^2_{t+1}|\mathcal{F}_t\right)$, suffice to determine the connection between patience and aggressiveness. This relies on an important observation: the belief about intrinsic strength, μ, is a martingale. Intrinsic strength at period t is also a martingale. Therefore, the one-step look ahead values suffice to characterize any finite step look-ahead values, i.e. $\mathrm{E}\left(w^i_{t+1}|\mathcal{F}_t\right) = \mathrm{E}\left(w^i_{t+k}|\mathcal{F}_t\right)$ for any $k > 1$.

On the methodological side, the key innovation in this section is to use martingale techniques to connect the single-player work of Quah and Strulovici (2013) and the existence of equilibrium results in the two-player work of Ohtsubo (1987). Similar to the fair case, Yasuda (1985) is utilized to ascertain that there is a unique set of sustainable values (even if mixed strategies are employed). We restrict attention to the strategies identified in Ohtsubo (1987). These are not necessarily the only strategies that support these payoffs, but any other equilibrium strategy also achieves the same payoffs. These characterized values and strategies are not in closed form, but they are closely related to Snell (1952) envelopes used in single player stopping problems. By utilizing martingale techniques we extend the single player results of Quah and Strulovici (2013) to order the values without directly computing them. This permits comparative statics results.

5 Conclusions and Future Work

5.1 Conclusions

This paper develops novel methods for stopping games with and without the addition of social experimentation. The results reveal inefficiencies produced by preemptive moves leading to early destructive confrontations. Unlike repeated games, this inefficiency cannot be alleviated by patience. The results also show that strength is the key mediator on the relation between patience and aggressiveness. Destructive confrontations are triggered sooner by preemptive moves of patient players. The exception to this strategic aggressiveness takes place at the side (if it exists) that is sufficiently strong and so, may optimally wait long for a better opportunity to strike.

The reducibility of the process (the game ends when one player starts a confrontation) makes the underlying logic in this game completely different from those in repeated games. Unlike folk theorems there is no possible threat that the players can make in order to sustain different equilibrium payoffs, no matter how patient players might be. This is the fundamental reason why there is no folk theorem in this stopping game with destructive confrontations.

5.2 Future work

A natural follow-up to this model would combine our stopping game and a repeated game. Like in this model, there would be a pre-stage whether each player decides, in each period, whether or not to engage the opponent. However, if so, the game would not end, but instead the players would be irrevocably engaged in a repeated game. We conjecture that this will reveal new tensions in the welfare implications of patience. This is left for future work.

Another natural follow-up to is to explore the case whether initiating a fight can give a direct advantage or disadvantage. This is an important question, and existence and uniqueness of values have already been explored in some settings (usually continuous-time models) within the stopping games literature (see Dynkin (1969), Neveu (1975), Rosenberg, Solan, and Vieille (2001), Szajowski (1993), Shmaya and Solan (2004), Ekstrom and Villeneuve (2006)). However most of these results are focused on existence with involved characterizations of equilibria. Thus, obtaining comparative statics results in this setting is a daunting task. This is left for future work.

A quite different follow-up would be in social experimentation where players' experimentation are directly payoff relevant to other players, and players are not necessarily adversaries. This can fit well situations like adopting new technologies or forming social norms. Our techniques could be utilized to shed light on these phenomena. This is also left for future work.

6 Appendix

6.1 Uniqueness of Values and Random Stopping Times

6.1.1 Mixed Strategies

In the literature of stopping games, there are three concepts of random stopping times. Proving the equivalence of these concepts is challenging, but can be found in Solan, Tsirelson, and Vieille (2012) in the context of stopping games, in Mertens, Sorin, and Zamir (2015) for a broader class of extensive form games. For the convenience of the reader, we give a brief overview of the three definitions. Let $\{\mathcal{F}_t\}$ be a filtration of a given probability triple $(\Omega, \{\mathcal{F}_t\}, \mathrm{P})$. Recall that a pure stopping time (strategy) is a function $\tau : \Omega \to \mathbb{N} \cup \{\infty\}$ such that $\{\omega \in \Omega : \tau(\omega) = n\} \in \mathcal{F}_n$ for all n.

The concepts of mixtures are as follows:

Definition 7 (Randomized Stopping Time)**.** *A randomized stopping time is a nonnegative adapted real valued process* $\rho = (\rho_n)_{n \in \mathbb{N} \cup \{\infty\}}$ *that satisfies* $\sum_{n \in \mathbb{N} \cup \{\infty\}} \rho_n(\omega) = 1$ *for every* $\omega \in \Omega$.

The interpretation is that when the true state of the world is ω the probability that the player stops at time n is $\rho_n(\omega)$.

Definition 8 (Behavior Stopping Time)**.** *A behavior stopping time is an adapted* $[0, 1]$ *valued process* $\pi = (\pi_n)_{n \in \mathbb{N}}$.

The interpretation is that when the true state of the world is ω the probability that the player stops at time n is $\pi_n(\omega)$ *conditional* on stopping occurring after time $n - 1$. For brevity we refer to a behavior stopping time as a behavior strategy.

Definition 9 (Mixed Stopping time)**.** *Let* I *denote the unit interval and* $\mathcal{B}$ *denote the borel algebra with respect to the Lebesgue measure. A mixed* *stopping time is a* $(\mathcal{B}\times\mathcal{F})$ *measurable function* $\mu : I \times \Omega \to \mathbb{N} \cup \{\infty\}$ *such that for every* $r \in I$*, the function* $\mu(r, \cdot)$ *is a stopping time.*

The interpretation is that a pure stopping time is chosen according to the uniform distribution at the beginning and used throughout the stopping problem.

6.1.2 Uniqueness of Values

We adopt the notion of a randomized strategy from Yasuda (1985) , which corresponds to behavior stopping times. We are only going to use behavioral stopping times, but by Solan, Tsirelson, and Vieille (2012) and Mertens, Sorin, and Zamir (2015) we know that other forms of mixtures are covered. For completeness, here we reproduce Yasuda (1985)'s result and proof with a simplified notation to match our setting. The result shows that there is a single sustainable pair of values in equilibrium. Thus, after this section we restrict attention to the equilibrium strategies we identify. Any other equilibrium attains the same values. Interested readers can consult Yasuda (1985) for broader results.

Definition 10. *Let* $(\pi^i_n)_{n\in\mathbb{N}} = \pi^i$ *be a behavior strategy for player* i*. That is* π^i_n *is a process adapted to* $\mathcal{F}_n$*, with* $0 \leq \pi^i_n \leq 1$ *for all* n*. And let* x^i_n *be an i.i.d. uniform variable independent of* $\mathcal{F}_n$*.* [5] *Then a* random stopping time *associated with a behavior strategy* π^i *is defined by*

$$\tau(\pi^i) = \inf\{n \geq 0 : x^i_n \leq \pi^i_n\}$$

The interpretation is that player i's behavior strategy describes the odds of stopping, π^i_n, as a function of the past and current history. The random variable x^i_n serves as an exogenous randomization device that is independent from the history of the game, and stopping by player i occurs when x^i_n is less than π^i_n. The definitions are consistent because the probability that x^i_n is less than π^i_n is equal to π^i_n.[6]

[5] For our purposes, independence from the filtration means that x_n is independent from all random variables adapted to $\mathcal{F}_n$ such as the odds of winning or the history of the game.

[6] The statement of Yasuda's (1985) theorem uses behavioral strategies and do not require the equivalent concept of random stopping times associated with a behavioral strategy. However, we make this definition because it is the one Yasuda uses in his proof.

If π_n^i is equal to 0 or 1 for every possible history, than the random stopping time associated with the behavior strategy is equivalent to a pure stopping time.

For each player $i = 1, 2$, let x_n^i be i.i.d. random variables of uniform distribution on $[0, 1]$ and independent of the original filtration $\mathcal{F}_n$. Let $\mathcal{G}_n$ be the σ-field generated by $\mathcal{F}_n, (x_n^i)_{i \in \{1,2\}}$. Consider the pair of behavior stopping times given by $(\{\pi_n^1\}_{n \in \mathbb{N}}, \{\pi_n^2\}_{n \in \mathbb{N}}) = (\pi^1, \pi^2)$.

Further let $\Pi_n^i = \{\pi^i : \pi_1^i = \pi_2^i = \dots \pi_{n-1}^i = 0\}$. That is the set Π_n^i correspond to the set of behavior strategies where player i never starts a conflict before n.

With some abuse of notation let $\tau(\pi^1, \pi^2)$ denote $\min(\tau(\pi^1), \tau(\pi^2))$, the time of the conflict.

Observation 1. *For any t, and any realization p_t, $w_t^2 = v - l - w_t^1$.*

By observation 1 we can write any payoff in period t, in terms of w_t^1.

Finally let

$$\begin{aligned}
\bar{\gamma}_n^1 &= ess \inf_{\Pi_n^2} ess \sup_{\Pi_n^1} E(\beta^{\tau(\pi^1,\pi^2)-n} w^1_{\tau(\pi^1,\pi^2)} | \mathcal{F}_n) \qquad (6.1.1)\\
\underline{\gamma}_n^1 &= ess \sup_{\Pi_n^1} ess \inf_{\Pi_n^2} E(\beta^{\tau(\pi^1,\pi^2)-n} w^1_{\tau(\pi^1,\pi^2)} | \mathcal{F}_n)\\
\bar{\gamma}_n^2 &= ess \inf_{\Pi_n^1} ess \sup_{\Pi_n^2} E(\beta^{\tau(\pi^1,\pi^2)-n} (v - l - w^1_{\tau(\pi^1,\pi^2)}) | \mathcal{F}_n)\\
\underline{\gamma}_n^2 &= ess \sup_{\Pi_n^2} ess \inf_{\Pi_n^1} E(\beta^{\tau(\pi^1,\pi^2)-n} (v - l - w^1_{\tau(\pi^1,\pi^2)}) | \mathcal{F}_n)
\end{aligned}$$

We now note that, for any pair of equilibrium strategies $\tau(\pi^1), \tau(\pi^2)$, confrontation occurs, almost surely, in finite time. That is, equilibrium payoffs must be achieved by fighting strategies, so that $\bar{\gamma}_n^i$ and $\underline{\gamma}_n^i$, $i \in \{1, 2\}$ are well-defined bounds.[7] To see this, first note that payoffs never exceeds v and so, continuation payoff has to be less than βv, $\beta < 1$. Let $\hat{p}^1(\beta) < 1$ and $\hat{p}^2(\beta) > 0$ be such that $\hat{p}^1(\beta)(v+l) - l > \beta v$ and $-\hat{p}^2(\beta)(v+l) + v > \beta v$. Then, in equilibrium, player 1 must confront wherever the probability of victory exceeds $\hat{p}^1(\beta)$, and player 2 must confront wherever the probability of victory

[7] The essential suprema and the essential suprema are obtained as limits of finite time payoffs and hence requires stopping times to be almost surely finite.

exceeds $1-\hat{p}^2(\beta)$. This follows because in these cases the expected payoffs obtained by immediate confrontation exceeds the maximum discounted payoffs that can be achieved by not confronting. Thus, in equilibrium, confrontation occurs either before or at any period t such that p_t exceeds $\hat{p}^1(\beta)$ or is below $1-\hat{p}^2(\beta)$. Now, $\{\tilde{p}_t\}$ has full support on $[0,1]$ and a continuous density. So, every period, p_t exceeds $\hat{p}^1(\beta)$ or is below $1-\hat{p}^2(\beta)$ with positive probability. Given that $\{\tilde{p}_t\}$ is i.i.d., by the Borel-Cantelli lemma, almost surely, p_t exceeds $\hat{p}^1(\beta)$ or is below $1-\hat{p}^2(\beta)$. Hence, almost surely, in equilibrium, confrontation happens in finite time.

We now reproduce the Theorem 3.1 of Yasuda (1985) (and its proof) adapted to our setting.

Theorem 3 (Yasuda)**.** *Assume that*

$$E[\sup_n |w_n^i|] < \infty,\ E[\sup_n (w_n^i)^-] < \infty \textit{ and } E[\sup_n (w_n^i)^+] < \infty,\ \textit{for each } i \in \{1,2\}.$$

Then for each i, $\bar{\gamma}_n^i$ and $\underline{\gamma}_n^i$ coincide for all n.

Proof. We reproduce the proof of Yasuda (1985) because we believe the proof is insightful. The approach taken by Yasuda (1985) is to define a value operator that every equilibrium value needs to satisfy and show that there exists a unique set of values that can satisfy this operator. The analysis in our case is simplified since in our game the payoffs arising from a confrontation are bounded.The reader may skip it and go directly to corollary 3. However, the notation in the original representation is challenging. Here we present the argument in a notation which is (to us) simpler to follow.

First observe that the assumptions in theorem 3 implies:

$$ess \sup_{\Pi_{n+1}^1} E(\beta^{\tau(\pi^1,\pi^2)-n} w^1_{\tau(\pi^1,\pi^2)} | \mathcal{F}_n) = \beta E(ess \sup_{\Pi_{n+1}^1} E(^{\tau(\pi^1,\pi^2)-n} w^1_{\tau(\pi^1,\pi^2)} | \mathcal{F}_{n+1}) | \mathcal{F}_n)$$

Now, consider the following two by two game:

	$Conf$	$No\ Conf$
$Conf$	$w_n^1, v-l-w_n^1$	$w_n^1, v-l-w_n^1$
$No\ Conf$	$w_n^1, v-l-w_n^1$	$\beta E(\gamma_{n+1}^1 \vert \mathcal{F}_n), \beta E(\gamma_{n+1}^2 \vert \mathcal{F}_n)$

Let VAL denote the equilibrium values of the bimatrix game. Consider the functional equation associated with VAL operator,

$$(\gamma_n^1,\gamma_n^2) \in VAL \begin{bmatrix} w_n^1, v-l-w_n^1 & w_n^1, v-l-w_n^1 \\ w_n^1, v-l-w_n^1 & \beta E(\gamma_{n+1}^1|\mathcal{F}_n), \beta E(\gamma_{n+1}^2|\mathcal{F}_n) \end{bmatrix} \qquad (6.1.2)$$

Notice that any equilibrium payoff has to satisfy 6.1.2. Now let, $(\hat{\gamma}_n^1,\hat{\gamma}_n^2)$ be a solution to 6.1.2 and $\hat{\pi}^1=(\hat{\pi}_n^1), \hat{\pi}^2=(\hat{\pi}_n^2)$ be an associated strategy. That is for each n

$$\begin{aligned}
\hat{\gamma}_n^1 &= (\hat{\pi}_n^1\hat{\pi}_n^2+\hat{\pi}_n^1(1-\hat{\pi}_n^2)+\hat{\pi}_n^2(1-\hat{\pi}_n^1))w_n^1+\beta(1-\hat{\pi}_n^1)(1-\hat{\pi}_n^2)E(\hat{\gamma}_{n+1}^1|\mathcal{F}_n) \\
\hat{\gamma}_n^2 &= (\hat{\pi}_n^1\hat{\pi}_n^2+\hat{\pi}_n^1(1-\hat{\pi}_n^2)+\hat{\pi}_n^2(1-\hat{\pi}_n^1))(v-l-w_n^1)+\beta(1-\hat{\pi}_n^1)(1-\hat{\pi}_n^2)E(\hat{\gamma}_{n+1}^2|\mathcal{F}_n)
\end{aligned}$$

Lemma 1. *For each n the following equalities hold:*

$$E(\beta^{\tau(\hat{\pi}^1,\hat{\pi}^2)-n}w^1_{\tau(\hat{\pi}^1,\hat{\pi}^2)}|\mathcal{F}_n)=\hat{\gamma}_n^1$$
$$E(\beta^{\tau(\hat{\pi}^1,\hat{\pi}^2)-n}(v-l-w^1_{\tau(\hat{\pi}^1,\hat{\pi}^2)})|\mathcal{F}_n)=\hat{\gamma}_n^2$$

Proof of Lemma. For each n

$$\begin{aligned}
& E(\beta^{\tau(\hat{\pi}^1,\hat{\pi}^2)-n}w^1_{\tau(\hat{\pi}^1,\hat{\pi}^2)}|\mathcal{F}_n)-\hat{\gamma}_n^1 \\
= \; & \beta^{m-n+1}E\left[[\prod_{k=n}^{m}(1-\hat{\pi}_k^1)(1-\hat{\pi}_k^2)](E(\beta^{\tau(\hat{\pi}^1,\hat{\pi}^2)-m-1}w^1_{\tau(\hat{\pi}^1,\hat{\pi}^2)}|\mathcal{F}_{m-1})-\hat{\gamma}_{m+1}^1)\right]
\end{aligned}$$

for any $m\geq n$. Since $\beta<1$ letting $m\to\infty$ we have

$$E(\beta^{\tau(\hat{\pi}^1,\hat{\pi}^2)-n}w^1_{\tau(\hat{\pi}^1,\hat{\pi}^2)}|\mathcal{F}_n)=\hat{\gamma}_n^1$$

and analogously

$$E(\beta^{\tau(\hat{\pi}^1,\hat{\pi}^2)-n}(v-l-w^1_{\tau(\hat{\pi}^1,\hat{\pi}^2)})|\mathcal{F}_n)=\hat{\gamma}_n^2$$

□

Now, consider a strategy $\pi^{2,(m)}=(\pi_1^{2,(m)},\pi_2^{2,(m)},\ldots)$ defined as follows: for each k

$$\pi_k^{2,(m)}=\begin{cases}\hat{\pi}_k^2 \text{ if } k>m\\ \pi_k^2 \text{ if } k\leq m\end{cases}$$

for an arbitrary strategy $\pi^2 = (\pi^2_1, \pi^2_2, \ldots)$. By lemma 1 it must be that

$$E(\beta^{\tau(\hat{\pi}^1,\pi^{2,(m)})-m} w^1_{\tau(\hat{\pi}^1,\pi^{2,(m)})}|\mathcal{F}_m) = E(\beta^{\tau(\hat{\pi}^1,\hat{\pi}^2)-m} w^1_{\tau(\hat{\pi}^1,\hat{\pi}^2)}|\mathcal{F}_m) = E(\hat{\gamma}^1_{m+1}|\mathcal{F}_m).$$

On the other hand since $\hat{\pi}^2$ satisfies 6.1.2, it has to be the case that

$$\hat{\gamma}^1_m \leq E(\beta^{\tau(\hat{\pi}^1,\pi^{2,(m)})-m} w^1_{\tau(\hat{\pi}^1,\pi^{2,(m)})}|\mathcal{F}_m)$$

But then iteratively $\hat{\gamma}^1_n \leq E(\beta^{\tau(\hat{\pi}^1,\pi^{2,(m)})-n} w^1_{\tau(\hat{\pi}^1,\pi^{2,(m)})}|\mathcal{F}_n)$ for each $m \geq n$. Since π^2 was arbitary letting $m \to \infty$ we have

$$\hat{\gamma}^1_n \leq ess \inf_{\Pi^2_n} E(\beta^{\tau(\hat{\pi}^1,\pi^2)-n} w^1_{\tau(\hat{\pi}^1,\pi^2)}|\mathcal{F}_n) \leq ess \sup_{\Pi^1_n} ess \inf_{\Pi^2_n} E(\beta^{\tau(\hat{\pi}^1,\pi^2)-n} w^1_{\tau(\hat{\pi}^1,\pi^2)}|\mathcal{F}_n) = \underline{\gamma}^1_n$$

Symmetric arguments deliver $\hat{\gamma}^1_n \geq \bar{\gamma}^1_n$ yielding $\bar{\gamma}^1_n = \underline{\gamma}^1_n$. An identical argument also delivers $\hat{\gamma}^2_n \leq \underline{\gamma}^2_n$ and $\hat{\gamma}^2_n \geq \bar{\gamma}^2_n$, yielding $\bar{\gamma}^2_n = \underline{\gamma}^2_n$. □

Corollary 3. *Our stopping games with behavior strategies has a unique pair of values that can be sustained in equilibrium.*

Proof. The inequalities in Yasuda (1985)'s assumptions hold because for all n, $-l < w^i_n < v$. Hence, $\bar{\gamma}^i_n = \underline{\gamma}^i_n$, $i \in \{1, 2\}$. The corollary is now achieved because any equilibrium payoffs (for player i at period n) given by $\hat{\gamma}^1_n$ and $\hat{\gamma}^2_n$ must be between $\underline{\gamma}^i_n$ and $\bar{\gamma}^i_n$. and thus follows directly from the last line in Yasuda's (1985) proof that $\hat{\gamma}^i_n = \bar{\gamma}^i_n = \underline{\gamma}^i_n$, $i \in \{1, 2\}$. □

Notice that corollary 3 only asserts uniqueness of sustainable equilibrium values. There are potentially multiple equilibrium strategies that attain these values.

6.2 Known Strength

6.2.1 Informal Description of the Proof Strategy for Theorem 1

By Yasuda (1985), there is a unique pair of equilibrium values. By Mertens (2002), we know there exists an equilibrium in behavior Markov strategies. When both players use a Markovian strategy, the value functions associated with any pair of Markovian strategies has a time independent structure due to the i.i.d. odds of victories. We use the i.i.d. distributions (on the odds of victory) to show that, in our setting, any equilibrium in behavior Markov strategies is an equilibrium in hitting strategies with a fixed threshold. Thus, there exists an equilibrium in hitting strategies with fixed thresholds. The computation of the (equilibrium) threshold is straightforward. While these computations only characterizes a particular equilibrium that Mertens (2002) has shown to exist, the equilibrium payoffs that it delivers are the only ones in this game as shown by Yasuda (1985).

6.2.2 Proof of Theorem 1

Step 1: Equilibria in Markovian strategies

Definition 11. *A behavior Markovian strategy, is a behavior stopping time, where* $(\pi_n|\mathcal{F}_n) \equiv (\pi_n|p_n)$.

The interpretation of a behavior Markovian strategy is that when the true state of the world is ω the probability that the player stops at time n is $\pi_n(\omega)$ *conditional* on stopping occurring after time $n-1$, only depends on the n^{th} coordinate of ω. That is, the probability that the player stops at time n conditional on stopping occurring after time $n-1$, only depends on the realization p_n.

Note that a behavior Markovian strategy defines a mapping $\pi^i : [0,1] \to [0,1]$, where if, at period n, the odds of victory is p_n and no confrontation has occurred until period n, then player i stops at period n with probability $\pi^i(p_n)$. Thus, given that $\{\tilde{p}_n\}$ is i.i.d., the following observation holds.

Observation 2. *When both players employ behavior Markovian strategies, the probability of a confrontation happening at period* $t+1$*, conditional on stopping occurring after time* t *does not depend on the history before and at period* t *and therefore this probability is a number (denoted* χ*) given by:*

$$\begin{aligned}\chi &= P(\tau^i \wedge \tau^j = t+1|\mathcal{F}_t) \\ &= \int_0^1 [\pi^i(p)\pi^j(p) + \pi^i(p)(1-\pi^j(p)) + (1-\pi^i(p))\pi^j(p)]dp\end{aligned}$$

Observation 3. *When both players employ behavior Markovian strategies, expected payoff from a confrontation at period $t+1$, conditional on stopping occurring after time t, are also independent of history until period t, including the realization at period t therefore these payoffs are numbers (denoted ξ^1 and ξ^2) given by:*

$$\begin{aligned}\xi^1 &= \int_0^1 \{p[\pi^i(p)\pi^j(p) + \pi^i(p)(1-\pi^j(p)) + (1-\pi^i(p))\pi^j(p)](v+l) - l\}dp \\ \xi^2 &= \int_0^1 \{-p[\pi^i(p)\pi^j(p) + \pi^i(p)(1-\pi^j(p)) + (1-\pi^i(p))\pi^j(p)](v+l) + v\}dp\end{aligned}$$

Observation 4. *When both players employ behavior Markovian strategies, expected continuation payoffs denoted by $E(V^i(\tau(\pi^i), \tau(\pi^j))|\mathcal{F}_t)$ are also independent of history until period t, including the realization at period t and is given by:*

$$\begin{aligned}E(V^i(\tau(\pi^i), \tau(\pi^j))|\mathcal{F}_t) &= \sum_{k=0}^{\infty} \beta^k [1-\chi]^k \left(\xi^i\right) \\ &= \frac{\xi^i}{1-\beta[1-\chi]}\end{aligned}$$

Definition 12. *A pure strategy τ^i is called a hitting strategy, if $\exists C^i \subseteq [0,1]$ such that*

$$\tau^i = \inf\{t \geq 0 | p_t \in C^i\}$$

That is player i stops the first time the realization p_t of the process $\tilde{p}_t$ belongs to C^i. The C^i is fixed for all periods t. If C is of the form $[k,1]$ or $[0,k]$ we call it a *hitting strategy with fixed threshold* k. By definition hitting strategies are Markovian.

Definition 13. *The best response of player i against player j's random stopping time associated with a behavior strategy $\tau(\pi^j)$ denoted by $b^i(\tau(\pi^j))$ $= (b^i_n(\tau(\pi^j)))_{n\in\mathbb{N}}$ is a behavior stopping time that satisfies:*

$$b^i_n(\tau(\pi^j)) \in \arg\max_{\Pi^i} E(\beta^{\tau(b^i(\tau^j))\wedge\tau^j - n} w^i_{\tau(b^i(\tau^j))\wedge\tau^j} | \mathcal{F}_n)$$

That is, $b^i(\tau(\pi^j))$ is a behavior stopping time of player i that is a best response to a (potentially behavior) stopping time $\tau(\pi^j)$ of player j.

Lemma 2 holds in our stopping game with an i.i.d. uniform process $\{\tilde{p}_t\}$.

Lemma 2. *In equilibrium if both players employ a behavior Markov strategy, then both players must be employing a hitting strategy with a fixed threshold.*

Proof. Suppose there exists an equilibrium in behavior Markovian strategies. Let $\tau(\pi^j)$ be a randomized stopping time associated with a behavior Markovian strategy and let $b_t^i(\tau(\pi^j))$ be a best response to $\tau(\pi^j)$ that is a behavior Markovian strategy. Let $V^i(b^i(\tau(\pi^j)), \tau(\pi^j))$ denote the continuation payoff for player i when player i employs the best response $b^i(\tau(\pi^j))$ and player j employs $\tau(\pi^j)$. Since both players employ strategies that are Markovian, due to observation 4 it must be the case that for any period t, $E(V^i(b^i(\tau(\pi^j)), \tau(\pi^j))|\mathcal{F}_t)$ is a constant real number. That is $\exists v^i(b^i(\tau(\pi^j)), \tau(\pi^j)) \in [-l, v]$ such that $E(V^i(b^i(\tau(\pi^j)), \tau(\pi^j)^j)|\mathcal{F}_t) = v^i(b^i(\tau(\pi^j)), \tau(\pi^j))$ for all t. But then for any period t and any realization w_t^i, the problem of player i is as follows:

$$\max_{b_t^i(\tau(\pi^j))\in[0,1]} [(1-b_t^i(\tau(\pi^j)))\beta v^i(b^i(\tau(\pi^j)), \tau(\pi^j)) + b_t^i(\tau(\pi^j))w_t^i](1-\mathrm{P}(\tau(\pi^j)=t))$$

$$+[\mathrm{P}(\tau(\pi^j)=t)E(w_t^i|\tau(\pi^j)=t)].$$

Since $v_i(\tau^i(\tau(\pi^j)), \tau(\pi^j))$ is bounded above and below by respectively v and $-l$ there exists a unique $\bar{p}_i$ such that $\beta v_i(b^i(\tau(\pi^j)), \tau(\pi^j)) = \bar{w}_t^i$. Thus, wlog letting player i be player 1.

$$\arg\max_{b_t^1(\tau(\pi^2))\in[0,1]} [(1-b_t^1(\tau(\pi^2)))\beta v_1(b^1(\tau(\pi^2)), \tau(\pi^2)) + b_t^1(\tau(\pi^2))(p_t v-(1-p_t)l)] =$$

$$\begin{cases} 1 \; if \; p_t > \bar{p}_1 \\ 0 \; if \; p_t < \bar{p}_1 \\ [0,1] \; if \; p_t = \bar{p}_1 \end{cases}$$

But then $\tau^1 = t \iff p_t > \bar{p}_1$ thus the best response to a behavior Markov strategy is a hitting strategy with a fixed threshold, with a measure 0 indifference in $\bar{p}_1$, which we break in favor of a confrontation. Given that hitting strategies with a fixed thresholds are Markovian, there exists an equilibrium in hitting strategies with fixed thresholds. □

In step 2, we now identify an equilibrium in hitting strategies with fixed thresholds.

Step 2: Identifying an equilibrium in hitting strategies

Let $\tau^1 = \inf\{t \geq 0 \mid p_t \geq p^1\}$ and $\tau^2 = \inf\{t \geq 0 \mid p_t \leq p^2\}$ denote two hitting strategies, where without loss of generality we assume $p^1 \geq 1/2$ and $p^2 \leq 1/2$.

Letting $F(p)$ and $f(p)$ denote the cdf and pdf of the i.i.d. stochastic process governing the odds. Now we start with any i.i.d. distribution and we will plug the uniform density as we proceed.

With hitting strategies the overall payoffs to the players are given by:

$$\begin{aligned}
U^1(\tau) &= \sum_{t=0}^{\infty} \beta^t \mathrm{P}(\tau^1 \wedge \tau^2 = t) E(\tilde{p}_t(v+l) - l | \tau^1 \wedge \tau^2 = t) \\
&= \sum_{t=0}^{\infty} \beta^t (F(p^1) - F(p^2))^t \left(\int_0^{p^2} [p(v+l) - l] f(p) dp + \int_{p^1}^{1} [p(v+l) - l] f(p) dp \right), \\
U^2(\tau) &= \sum_{t=0}^{\infty} \beta^t \mathrm{P}(\tau^1 \wedge \tau^2 = t) E(-\tilde{p}_t(v+l) + v | \tau^1 \wedge \tau^2 = t) \\
&= \sum_{t=0}^{\infty} \beta^t (F(p^1) - F(p^2))^t \left(\int_0^{p^2} [-p(v+l) + v] f(p) dp + \int_{p^1}^{1} [-p(v+l) + v] f(p) dp \right).
\end{aligned}$$

In particular, for any history, the continuation payoffs with hitting strategies denoted $E(V^i(\tau)|\mathcal{F}_t)$ are also given by:

$$\begin{aligned}
E(V^1(\tau)|\mathcal{F}_t) &= \sum_{k=t}^{\infty} \beta^{k-t} \mathrm{P}(\tau^1 \wedge \tau^2 = k|\mathcal{F}_t) E(\tilde{p}_t(v+l) - l | \tau^1 \wedge \tau^2 = k|\mathcal{F}_t) \\
&= \sum_{k=t}^{\infty} \beta^{k-t} (F(p^1) - F(p^2))^t \left(\int_0^{p^2} [p(v+l) - l] f(p) dp + \int_{p^1}^{1} [p(v+l) - l] f(p) dp \right), \\
E(V^2(\tau)|\mathcal{F}_t) &= \sum_{k=t}^{\infty} \beta^{k-t} \mathrm{P}(\tau^1 \wedge \tau^2 = k|\mathcal{F}_t) E(-\tilde{p}_t(v+l) + v | \tau^1 \wedge \tau^2 = k|\mathcal{F}_t) \\
&= \sum_{k=t}^{\infty} \beta^{k-t} (F(p^1) - F(p^2))^t \left(\int_0^{p^2} [-p(v+l) + v] f(p) dp + \int_{p^1}^{1} [-p(v+l) + v] f(p) dp \right).
\end{aligned}$$

The underlying process is i.i.d. and the strategies are hitting strategies. Therefore the continuation payoffs are constant across time (including period

0) and independent of history. Hence with a slight abuse of notation we let $E(V^i(\tau)) = U^i(\tau) = E(V^i(\tau)|\mathcal{F}_t)$ for any t denote the continuation values. Furthermore since the hitting strategies are identified by two thresholds, we will simplify the notation by using notions of strategy and threshold interchangeably.

For a pair of thresholds (p^1, p^2) that belongs $(0.5, 1) \times (0, 0.5)$ to be an equilibrium the thresholds must satisfy the following system of 6 equations:

$$
\begin{aligned}
&\beta E(V^1(\tau)) = (p^1 v - (1-p^1)l) \qquad (6.2.1)\\
&\beta E(V^2(\tau)) = (-p^2 l + (1-p^2)v)\\
&E(V^1(\tau)) = \int_0^{p^2} [p(v+l) - l] f(p)dp + \int_{p^1}^1 [p(v+l) - l] f(p)dp + \mathrm{P}(p^2 < p < p^1)\beta E(V^1(\tau))\\
&E(V^2(\tau)) = \int_0^{p^2} [-p(v+l) + v] f(p)dp + \int_{p^1}^1 [-p(v+l) + v] f(p)dp + \mathrm{P}(p^2 < p < p^1)\beta E(V^2(\tau))\\
&E(V^1(\tau)) \geq \int_0^1 [p(v+l)] f(p)dp - l\\
&E(V^2(\tau)) \geq \int_0^1 [-p(v+l)] f(p)dp + v
\end{aligned}
$$

The first two equations show that at the threshold the players have to be indifferent between confrontation and not. The third and the fourth equations deliver the continuation values. The last two inequalities ensures that the payoffs associated with the thresholds have to be higher than ex-ante immediate confrontation.

The system of equations 6.2.1 can be re-written as follows:

$$\beta E(V^1(\tau)) = (p^1 v - (1-p^1)l)$$
$$\beta E(V^2(\tau)) = (-p^2 l + (1-p^2)v)$$
$$E(V^1(\tau)) = \int_0^{p^2} [p(v+l)]f(p)dp + \int_{p^1}^{1} [p(v+l)]f(p)dp + p^1(v+l)\int_{p^2}^{p^1} f(p)dp - l$$
$$E(V^2(\tau)) = \int_0^{p^2} [-p(v+l)]f(p)dp + \int_{p^1}^{1} [-p(v+l)]f(p)dp - p^2(v+l)\int_{p^2}^{p^1} f(p)dp + v$$
$$E(V^1(\tau)) \geq \int_0^{1} [p(v+l)]f(p)dp - l$$
$$E(V^2(\tau)) \geq \int_0^{1} [-p(v+l)]f(p)dp + v$$

So now we solve for $E(V^1(\tau))$ and $E(V^2(\tau))$ by combining the third and the fourth equation with the first and the second equations.

$$\frac{(p^1 v - (1-p^1)l)}{\beta} = \int_0^{p^2} [p(v+l) - l]f(p)dp + \int_{p^1}^{1} [p(v+l) - l]f(p)dp + (p^1 v - (1-p^1)l)\int_{p^2}^{p^1} f(p)dp$$
$$\frac{(-p^2 l + (1-p^2)v)}{\beta} = \int_0^{p^2} [-p(v+l) + v]f(p)dp + \int_{p^1}^{1} [-p(v+l) + v]f(p)dp + (-p^2 l + (1-p^2)v)\int_{p^2}^{p^1} f(p)dp$$

Letting $\phi(p^1, p^2) = F(p^1) - F(p^2)$ i.e. the total probability of no conflict, we have the following:

$$\int_0^{p^2} [p(v+l)]f(p)dp + \int_{p^1}^{1} [p(v+l)]f(p)dp = \frac{(p^1(v+l) - l)}{\beta} - (p^1(v+l))\phi(p^1, p^2) + l$$
$$\int_0^{p^2} [p(v+l)]f(p)dp + \int_{p^1}^{1} [p(v+l)]f(p)dp = \frac{(p^2(v+l) - v)}{\beta} - (p^2(v+l))\phi(p^1, p^2) + v$$

The system above is a necessary condition for equilibria with general i.i.d. distributions. We now plug in the uniform density to the system, i.e. $f(p) = 1$ and $F(p) = p$. With some algebra, the solution is

$$\bar{p}^1 = \frac{1}{2} + \frac{1}{4\beta} - \frac{\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}{4\beta(l+v)} \tag{6.2.2}$$

$$\bar{p}^2 = \frac{1}{2} - \frac{1}{4\beta} + \frac{\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}{4\beta(l+v)}$$

Finally, we now verify the inequalities in the fifth and sixth equations. They hold given that $v < l$ and $\beta < 1$. This follows because

$$\begin{aligned} E(V^1(\bar{\tau})) &= \frac{-l(1-\bar{p}^1+\bar{p}^2) + (\frac{v+l}{2})[(\bar{p}^2)^2 + 1 - (\bar{p}^1)^2]}{1-\beta(\bar{p}^1-\bar{p}^2)} \\ &= \frac{(\frac{v-l}{2})[1-2(\frac{1}{4\beta} - \frac{\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}{4\beta(l+v)})]}{1-\beta 2(\frac{1}{4\beta} - \frac{\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}{4\beta(l+v)})} > \frac{v-l}{2}. \end{aligned}$$

The inequality on the sixth equation (which refers to $E(V^2(\bar{\tau}))$) also holds and the calculation is completely analogous.

Proof of Corollary 1

Taking the derivative of p^1 in 6.2.2 with respect to β and simplifying yields

$$\frac{\partial p^1}{\partial \beta} = \frac{-2l\beta + 2v\beta + l + v - \sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}{4\beta^2\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}}$$

We now show that the derivative is negative whenever $v \leq l$.

To see that the numerator is negative:

$$-2l\beta + 2v\beta + l + v \leq \sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)} \text{ or}$$

$$(-2l\beta + 2v\beta + l + v)^2 \leq \left(\sqrt{(l+v)\left(l+v-4l\beta+4v\beta+4l\beta^2-4v\beta^2\right)}\right)^2$$

Which simplifies to:

$$8\beta^2 v(v-l) \leq 0 \text{ or}$$
$$v \leq l$$

Proof of Corollary 2

Now, we will see the effect of changing just one player's discount factor when initially they start from the same β. Here we consider small changes around the initial equilibrium we found. Using the first and third equations in 6.2.1 we construct $f_1(p^1, p^2, \beta^1, \beta^2)$ and using the second and fourth equations we construct $f_2(p^1, p^2, \beta^1, \beta^2)$ as follows:

$$\begin{aligned} f_1(p^1, p^2, \beta^1, \beta^2) = \beta^1 &\left(\int_0^{p^2} [p(v+l)]dp + \int_{p^1}^1 [p(v+l)]dp + p^1(v+l)(p^1 - p^2) - l \right) \\ &- (p^1 v - (1-p^1)l) \\ f_2(p^1, p^2, \beta^1, \beta^2) = \beta^2 &\left(\int_0^{p^2} [p(v+l)]dp + \int_{p^1}^1 [p(v+l)]dp + p^2(v+l)(p^1 - p^2) - v \right) \\ &- (p^2 l - (1-p^2)v) \end{aligned}$$

It is straightforward to see that if $\beta^1 = \beta^2 = \beta$ then setting $\bar{p}^1 = \bar{p}^1(\beta^1, \beta^2)$ and $\bar{p}^2 = \bar{p}^2(\beta^1, \beta^2)$ we have both f_1 and f_2 equal to zero. That is,

$$f_1(\bar{p}^1, \bar{p}^2, \beta, \beta) = 0 \text{ and } f_2(\bar{p}^1, \bar{p}^2, \beta, \beta) = 0.$$

With some algebra, the Jacobian can be explicitly computed and shown to be non-singular. Hence, we can use the implicit function theorem on the system of two equations above. Let $\bar{p}(\beta, \beta)$ denote the vector of thresholds $(\bar{p}^1(\beta, \beta), \bar{p}^2(\beta, \beta))$ where both discount factors are equal to β. By the implicit function theorem,

$$\begin{aligned} D_\beta \bar{p}(\beta, \beta) = &\frac{-1}{-\beta^2(v+l)^2(\bar{p}^1(\beta,\beta) - \bar{p}^2(\beta,\beta))^2 + (v+l)^2} \times \\ &\begin{pmatrix} \beta(v+l)(\bar{p}^1(\beta,\beta) - \bar{p}^2(\beta,\beta)) - (v+l) & -\beta(v+l)(\bar{p}^1(\beta,\beta) - \bar{p}^2(\beta,\beta)) \\ -\beta(v+l)(\bar{p}^1(\beta,\beta) - \bar{p}^2(\beta,\beta)) & \beta(v+l)(\bar{p}^1(\beta,\beta) - \bar{p}^2(\beta,\beta)) - (v+l) \end{pmatrix} \times \\ &\begin{pmatrix} d_{11} & 0 \\ 0 & d_{22} \end{pmatrix} \end{aligned}$$

Where

$$d_{11} = \int_0^{\bar{p}^2(\beta,\beta)} [p(v+l)]dp + \int_{\bar{p}^1(\beta,\beta)}^1 [p(v+l)]dp + \bar{p}^1(\beta,\beta)(v+l)(\bar{p}^1(\beta,\beta) - \bar{p}^2(\beta,\beta)) - l$$

and

$$d_{22} = \int_0^{\bar{p}^2(\beta,\beta)} [p(v+l)]dp + \int_{\bar{p}^1(\beta,\beta)}^1 [p(v+l)]dp + \bar{p}^2(\beta,\beta)(v+l)(\bar{p}^1(\beta,\beta) - \bar{p}^2(\beta,\beta)) - v$$

Plugging in the values of $\bar{p}^1(\beta,\beta)$ and $\bar{p}^2(\beta,\beta)$ from 6.2.2, we can precisely calculate the partial derivatives. In particular the signs of the partial derivatives are as follows.

$$\begin{pmatrix} \frac{\partial \bar{p}^1}{\partial \beta^1} & \frac{\partial \bar{p}^1}{\partial \beta^2} \\ \frac{\partial \bar{p}^2}{\partial \beta^1} & \frac{\partial \bar{p}^2}{\partial \beta^2} \end{pmatrix} = \begin{pmatrix} - & + \\ - & + \end{pmatrix}$$

The proof is now concluded because an increase in $\bar{p}^1$ is corresponds to a decrease in aggressiveness whereas an increase in $\bar{p}^2$ corresponds to an increase.

6.2.3 Informal Description of the Proof Strategy for Theorem 2

The proof of this theorem is in three steps. By Yasuda (1985), there is a unique set of equilibrium values. We then use a characterization of equilibrium values, due to (Ohtsubo 1987). The second step identifies martingale properties of equilibrium values. The final step utilizes a lemma (lemma 3) along with martingale properties in step 2 to obtain an ordering of stopping times (i.e., the comparative statics relating aggressiveness and strength). Similar to the i.i.d. case, the equilibrium values are unique, but there are potentially multiple equilibrium strategies that sustain these values.

6.2.4 Proof of Theorem 2

Step 1: Characterization of continuation values

The w_t^i in definition 2.4 is crucial to identify and characterize an equilibrium. In that direction for each $m \in \mathbb{N}$, let the following sequences of random variables $\{(\psi_n^m, \phi_n^m)\}_{n=0}^m$ be defined by backward induction in the following manner:

$$(\psi_m^m, \phi_m^m) = (w_m^1, w_m^2)$$

$$(\psi_n^m, \phi_n^m) = \begin{cases} (w_n^1, w_n^2) \text{ if } (w_n^1, w_n^2) \geq \left(\beta E\left(\psi_{n+1}^m | \mu_n\right), \beta E\left(\phi_{n+1}^m | \mu_n\right)\right) \\ \left(\beta E\left(\psi_{n+1}^m | \mu_n\right), \beta E\left(\phi_{n+1}^m | \mu_n\right)\right) \quad o/w \end{cases}$$

This construction is very similar to the construction of a Snell envelope, extended to accommodate 2 players. Here let us remind that the Snell envelope is the value of a discounted optimal stopping problem with a single decision maker. Hence, it is the smallest supermartingale majorant of the discounted process that is being stopped optimally. The calculation of the Snell envelope relies on a similar backward induction argument, where first the stopping problem is assumed to stop at an arbitrary time (m in the above construction) and then the values for any $n < m$ is calculated by taking the maximum of the current stopped value and the expected continuation payoff. Finally the envelope is identified as the limit of the values as m tends to infinity. Further details and properties of the construction can be found in Snell (1952). There is a crucial observation to be made here. Both the Snell envelope and the construction of Ohtsubo (1987) is the limit of a process that is eventually stopped. The main implication is that these constructions rely on stopping (confrontation) happening almost surely, which is true in our case due to confrontations being destructive.

It is shown in Ohtsubo (1987) that $\psi_n = \lim_{m\to\infty} \psi_n^m$ and $\phi_n = \lim_{m\to\infty} \phi_n^m$ are well defined. Similar to the Snell envelope, the P-limit of these sequences define the essential suprema, which was shown to be equal to the equilibrium payoffs in the stopping game of (Ohtsubo 1987).

So, Ohtsubo (1987) characterizes a pair of equilibrium values using backward induction, and shows associated stopping times that attain these values. Here we present Ohtsubo (1987) theorem without proof.

Theorem 4 (Ohtsubo)**.**

$$\bar{\tau}^1 = \inf\{k \geq 0 | \beta E\left(\psi_{k+1}|\mu_k\right) \leq w_k^1\} \tag{6.2.3}$$

$$\bar{\tau}^2 = \inf\{k \geq 0 | \beta E\left(\phi_{k+1}|\mu_k\right) \leq w_k^2\} \tag{6.2.4}$$

constitute an equilibrium of this game. Where, the sequences $\{\psi\}, \{\phi\}$ *correspond to the associated equilibrium values which are given by,*

$$\begin{aligned}
\psi_t &= ess\sup_{\tau^1} E\left(\beta^{\bar{\tau}^2\wedge\tau^1} w^1_{\bar{\tau}^2\wedge\tau^1}|\mu_t\right) \\
\phi_t &= ess\sup_{\tau^2} E\left(\beta^{\tau^2\wedge\bar{\tau}^1} w^2_{\tau^2\wedge\bar{\tau}^1}|\mu_t\right) \\
E(V^1(\tau)|\mathcal{F}_t) &= \psi_t = E\left(\beta^{\bar{\tau}^1\wedge\bar{\tau}^2} w^1_{\bar{\tau}^1\wedge\bar{\tau}^2}|\mu_t\right) \\
E(V^2(\tau)|\mathcal{F}_t) &= \phi_t = E\left(\beta^{\bar{\tau}^1\wedge\bar{\tau}^2} w^2_{\bar{\tau}^1\wedge\bar{\tau}^2}|\mu_t\right)
\end{aligned}$$

Ohtsubo (1987) shows that $(\bar{\tau}^1, \bar{\tau}^2)$ is an equilibrium. Here as is usual with Snell envelopes, a closed form characterization of the equilibrium strategies is not obtained. However, we can still obtain comparative statics results.

The next proposition characterizes the optimal continuation condition for equilibrium strategies in Ohtsubo (1987).

Proposition 1. *For an arbitrary belief μ_n and current probability p_n the following inequalities have to hold simultaneously for the game to continue to period $n+1$*[8]

$$vp_n + (1-p_n)l \leq \beta \left(E\left(\psi_{n+1}|\mu_n\right)\right)$$
$$-p_n l + (1-p_n)v \leq \beta \left(E\left(\phi_{n+1}|\mu_n\right)\right)$$

The continuation condition in proposition 1 does not involve any discount factors on the left hand side. Thus pinning down how the right hand side changes with respect to the discount factor is crucial for the desired comparative statics results. The next step highlights the martingale properties associated with the continuation values (seen in the right-hand side of the equation above). This simplifies the problem of obtaining comparative statics results.

Step 2: Martingale properties

Corollary 4. *w^i_m is a martingale with respect to μ_{m-1}. That is, $E(w^i_m|\mu_{m-1}) = E(w^i_{m+k}|\mu_{m-1})$ for any $k \in N$.*

Proof. Let us directly calculate $\mathrm{E}\left(w^i_m|\mu_{m-1}\right)$. Utilizing the fact that $f^1(p) =$

[8] breaking indifference in favor of conflict

$f^2(1-p)$,

$$\begin{aligned}
\mathrm{E}\left(w_m^1|\mu_{m-1}\right) =& (v+l)\int_0^1 p(f^1(p)\mu_{m-1}+f^2(p)(1-\mu_{m-1}))dp - l\\
=&(v+l)\left[\mu_{m-1}\int_0^{1/2} pf^1(p)+(1-p)f^2(p)dp\right.\\
&\left.+(1-\mu_{m-1})\int_0^{1/2}(1-p)f^1(p)+pf^2(p)dp\right]-l\\
=&\mu_{m-1}\left[(v+l)\int_0^{1/2}(f^1(p)(2p-1)+f^2(p)(1-2p))dp\right]\\
&+(v+l)\int_0^{1/2}((1-p)f^1(p)+pf^2(p))dp-l.
\end{aligned}$$

by a similar calculation we have

$$\begin{aligned}
\mathrm{E}\left(w_m^2|\mu_{m-1}\right) \;=\; & -\mu_{m-1}\left[(v+l)\int_0^{1/2}(f^1(p)(2p-1)+f^2(p)(1-2p))dp\right]\\
&-(v+l)\int_0^{1/2}((1-p)f^1(p)+pf^2(p))dp+v.
\end{aligned}$$

It is well established that, the belief μ_t is a martingale, and consequently w_t^i are martingales w.r.t to μ_t since they are just linear functions of μ_t. Now, we observe that given the other player's strategy, the problem of a player is to maximize a stopped martingale, which is again a martingale by (Williams 1991). □

Step 3: Ordering of stopping times

Notice that the strategies identified by Ohtsubo (1987) are Markovian (where the state space is the space of beliefs). This follows because the strategies in 6.2.3 and 6.2.4 only depend on the current belief μ_n and the current odds of winning p_n.Thus, wlog (see (Puterman 2014)) for any given belief, we can rescale time and only consider time 0. Given the other player's strategy, and looking at the time 0 problem, we can rewrite the continuation values in incremental form (in the Lebesgue-Stieltjes sense) as follows,

$$\phi_0 = ess \sup_{\tau^1} E\left(\sum_0^{\tau^1} \beta^t(\Delta w_t^1) + w_t^1(\Delta\beta^t)|\mu_0, \tau^2\right)$$
$$\psi_0 = ess \sup_{\tau^2} E\left(\sum_0^{\tau^2} \beta^t(\Delta w_t^2) + w_t^2(\Delta\beta^t)|\mu_0, \tau^1\right)$$

where $\Delta w_t^i = w_{t+1}^i - w_t^i$, and $\Delta\beta^t = \beta^{t+1} - \beta^t$. The incremental form is convenient because with it, we can utilize the following lemma due to Quah and Strulovici (2013).

Lemma 3 (Strulovici & Quah)**.** *Let H be a regular stochastic process of bounded variation such that $E[H_t] \leq E[H_{\bar{t}}]$ for all $t \in [0, \bar{t})$, and let γ be a positive regular deterministic process. Then,*

$$E\left[\int_0^{\bar{t}} \gamma_s dH_s\right] \geq \gamma(0)E[H(\bar{t}) - H(0)].^9$$

Using the lemma 3 with $\beta \leq \hat{\beta}$ we get the following inequalities. First, letting $\gamma = \frac{\beta}{\hat{\beta}}$ and $dH = \hat{\beta}(\Delta w_t^1 + w_t\Delta\hat{\beta}^t)$ yields:

$$E\left[\sum_0^{\bar{t}} \beta^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right] \geq \frac{\beta}{\hat{\beta}} E\left[\sum_0^{\bar{t}} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right]. \quad (6.2.5)$$

Similarly letting $\gamma = \frac{\hat{\beta}}{\beta}$ and $dH = \beta(\Delta w_t^1 + w_t\Delta\beta^t)$

$$E\left[\sum_0^{\bar{t}} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right] \geq \frac{\hat{\beta}}{\beta} E\left[\sum_0^{\bar{t}} \beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right]. \quad (6.2.6)$$

Given that $\beta \leq \hat{\beta}$, $\Delta\hat{\beta}^t \geq \Delta\beta^t \quad \forall t$. So we have the following;

[9] The proof is identical to theirs, the change in assumptions only allow us to use integration by parts without assuming γ is increasing.

If $E[w_{t+1}|\mu_t] \geq 0$ then using inequality 6.2.6, we replace $\Delta\beta^t$ in the left hand side with $\Delta\hat{\beta}^t$ to get

$$E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right] \geq \frac{\hat{\beta}}{\beta} E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right] \geq 0.$$

The positivity in our last inequality is due to a result by ($Williams$ 1991).It shows that if a martingale has positive expectation then the discounted version of the same martingale also has positive expectation. Since the expressions above are all positive and $\frac{\hat{\beta}}{\beta} \geq 1$, it follows that the inequalities below hold.

$$E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right] \geq E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right] \geq 0.$$

If $E[w_{t+1}|\mu_t] \leq 0$ using inequality 6.2.5, we replace $\Delta\hat{\beta}^t$ in the left hand side with $\Delta\beta^t$ to get

$$0 \geq E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right] \geq \frac{\beta}{\hat{\beta}} E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right]$$

Due to negativity we can just remove $\frac{\beta}{\hat{\beta}}$ and the order will still remain the same. In a similar manner we have,

$$0 \geq E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right] \geq E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\beta)} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right]$$

Since the sums consist of martingales and martingale difference sequences, the inequalities above still hold even if we change the stopping rule for the second player. In particular, the inequalities above hold if we replace $\tau^2(\beta)$ with t, for any t. So, we also have:

If $E[w_{t+1}|\mu_t] \geq 0$ then

$$E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\hat{\beta})} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right] \geq E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\hat{\beta})} \beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right] \geq 0$$

Analogously, if $E[w_{t+1}|\mu_t] \leq 0$

$$0 \geq E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\hat{\beta})} \beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)|\mu_0\right] \geq E\left[\sum_{0}^{\tau^1(\beta)\wedge\tau^2(\hat{\beta})} \hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)|\mu_0\right]$$

The next step is to show that, almost surely, the stopping times are ordered in the same manner, i.e. $\tau^i(\hat{\beta}) \geq \tau^i(\beta)$ if $E[w_{t+1}^i|\mu_t] \geq 0$ and $\tau^i(\hat{\beta}) \leq \tau^i(\beta)$ if $E[w_{t+1}^i|\mu_t] \leq 0$.

For a contradiction suppose $\Psi = \{\omega : \tau^1(\beta) \geq \tau^1(\hat{\beta}), E[w_{t+1}|\mu_t] \geq 0\}$ has strictly positive probability. Hence, unless the other player fights with certainty, by a similar calculation as above, we get

$$\begin{aligned} &E\left(\sum_{\tau^1(\hat{\beta})}^{\tau^1(\beta)} (\hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)) I_{(\Psi)}|\mu_t, \tau^2(\beta)\right) \\ \geq &\frac{\hat{\beta}}{\beta} E\left(\sum_{\tau^1(\hat{\beta})}^{\tau^1(\beta)} (\beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)) I_{(\Psi)}|\mu_t, \tau^2(\beta)\right) \end{aligned}$$

Now, since $\tau^1(\beta)$ was optimal for discount factor β, the right hand side has to be weakly positive. So, the inequalities above still hold if we change the stopping rule for the second player. Hence, dropping $\frac{\hat{\beta}}{\beta}$ yields

$$\begin{aligned} &E\left(\sum_{\tau^1(\hat{\beta})}^{\tau^1(\beta)} (\hat{\beta}^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t)) I_{(\Psi)}|\mu_t, \tau^2(\hat{\beta})\right) \\ \geq &E\left(\sum_{\tau^1(\hat{\beta})}^{\tau^1(\beta)} (\beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t)) I_{(\Psi)}|\mu_t, \tau^2(\hat{\beta})\right) \geq 0 \end{aligned}$$

But then, waiting for $\tau^1(\beta)$ is weakly better off on Ψ. So, the stopping rule defined as $\max\{\tau^1(\beta), \tau^1(\hat{\beta})\}$ dominates $\tau^1(\hat{\beta})$ contradicting its optimality.

Similarly for a contradiction suppose $\Phi = \{\omega : \tau(\beta) \leq \tau(\hat{\beta}), E[w_{t+1}|\mu_t] \leq 0\}$ has strictly positive probability. By an identical argument we are going to get

$$E\left(\sum_{\tau(\beta)}^{\tau(\hat{\beta})}(\beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t))I_{(\Phi)}|\mu_t, \tau^2(\beta)\right)$$
$$\geq \frac{\beta}{\hat{\beta}}E\left(\sum_{\tau(\beta)}^{\tau(\hat{\beta})}(\beta^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t))I_{(\Phi)}|\mu_t, \tau^2(\beta)\right)$$

Now, since $\tau^1(\beta)$ was optimal for β, the left hand side has to be weakly negative but then by changing the stopping rule and dropping $\frac{\beta}{\hat{\beta}}$, we have

$$0 \geq E\left(\sum_{\tau(\beta)}^{\tau(\hat{\beta})}(\beta^t \Delta w_t^1 + w_t^1(\Delta\beta^t))I_{(\Phi)}|\mu_t, \tau^2(\hat{\beta})\right)$$
$$\geq E\left(\sum_{\tau(\beta)}^{\tau(\hat{\beta})}(\beta^t \Delta w_t^1 + w_t^1(\Delta\hat{\beta}^t))I_{(\Phi)}|\mu_t, \tau^2(\hat{\beta})\right)$$

But then, stopping earlier at $\tau^1(\beta)$ is weakly better off on Φ, so the stopping rule defined as $\min\{\tau^1(\beta), \tau^1(\hat{\beta})\}$ dominates $\tau^1(\hat{\beta})$ contradicting its optimality.

References

ABREU, D., AND F. GUL (2000): "Bargaining and reputation," *Econometrica*, 68(1), 85–117.

ACEMOGLU, D., AND J. A. ROBINSON (2001): "A theory of political transitions," *American Economic Review*, pp. 938–963.

BERGEMANN, D., AND J. VÄLIMÄKI (1996): "Learning and strategic pricing," *Econometrica: Journal of the Econometric Society*, pp. 1125–1149.

BOLTON, P., AND C. HARRIS (1999): "Strategic experimentation," *Econometrica*, 67(2), 349–374.

COMPTE, O., AND P. JEHIEL (2002): "On the role of outside options in bargaining with obstinate parties," *Econometrica*, 70(4), 1477–1517.

DAL BÓ, E., P. HERNÁNDEZ, AND S. MAZZUCA (2015): "Before Institutions: Security, Prosperity, and the Rise and Fall of Civilizations," Working papers.

DYNKIN, E. B. (1969): "Game Variant of a Problem on Optimal Stopping," *Soviet Mathematics Doklady*, 10, 270 – 274.

EKSTROM, E., AND S. VILLENEUVE (2006): "On the Value of Optimal Stopping Games," *The Annals of Applied Probability*, 16(3), pp. 1576–1596.

FEARON, J. D. (1995): "Rationalist Explanations for War," *International Organization*, 49(3), pp. 379–414.

GARFINKEL, M. R., AND S. SKAPERDAS (2007): "Economics of conflict: An overview," *Handbook of defense economics*, 2, 649–709.

KLEIN, N., AND S. RADY (2011): "Negatively Correlated Bandits," *The Review of Economic Studies*.

MERTENS, J.-F. (2002): "Chapter 47 Stochastic games," vol. 3 of *Handbook of Game Theory with Economic Applications*, pp. 1809 – 1832. Elsevier.

MERTENS, J.-F., AND T. PARTHASARATHY (1991): *Nonzero-sum stochastic games*. Springer.

Mertens, J.-F., S. Sorin, and S. Zamir (2015): *Repeated games*, vol. 55. Cambridge University Press.

Neveu, J. (1975): *Discrete Parameter Martingales*. North-Holland. Amsterdam.

Ohtsubo, Y. (1987): "A Nonzero-Sum Extension of Dynkin's Stopping Problem," *Mathematics of Operations Research*, 12(2), pp. 277–296.

Osborne, M. J., and A. Rubinstein (1990): *Bargaining and markets*, vol. 34. Academic press San Diego.

Powell, R. (1993): "Guns, butter, and anarchy.," *American Political Science Review*, 87(01), 115–132.

——— (1999): *In the shadow of power: States and strategies in international politics*. Princeton University Press.

——— (2004): "The inefficient use of power: Costly conflict with complete information," *American Political Science Review*, 98(02), 231–241.

——— (2006): "War as a Commitment Problem," *International Organization*, 60, 169–203.

Puterman, M. L. (2014): *Markov decision processes: discrete stochastic dynamic programming*. John Wiley & Sons.

Quah, J. K.-H., and B. Strulovici (2013): "Discounting, values, and decisions," *Journal of Political Economy*, 121(5), 896–939.

Rosenberg, D., E. Solan, and N. Vieille (2001): "Stopping games with randomized strategies," *Probability Theory and Related Fields*, 119(3), 433–451.

——— (2007): "Social Learning in One-Arm Bandit Problems," *Econometrica*, 75(6), 1591–1611.

Roth, A. E. (1985): *Game-theoretic models of bargaining*. Cambridge University Press.

Serrano, R. (2007): "Bargaining," Working Papers 2007-06, Instituto MadrileÃśo de Estudios Avanzados (IMDEA) Ciencias Sociales.

Shapley, L. S. (1953): "Stochastic games," *Proceedings of the National Academy of Sciences of the United States of America*, 39(10), 1095.

Shmaya, E., and E. Solan (2004): "Two-Player Nonzero-Sum Stopping Games in Discrete Time," *The Annals of Probability*, 32(3), pp. 2733–2764.

Snell, J. L. (1952): "Applications of martingale system theorems," *Transactions of the American Mathematical Society*, pp. 293–312.

Solan, E., B. Tsirelson, and N. Vieille (2012): "Random stopping times in stopping problems and stopping games," *arXiv preprint arXiv:1211.5802*.

Solan, E., and N. Vieille (2002): "Correlated equilibrium in stochastic games," *Games and Economic Behavior*, 38(2), 362–399.

Sorin, S. (1986): "Asymptotic properties of a non-zero sum stochastic game," *International Journal of Game Theory*, 15(2), 101–107.

Szajowski, K. (1993): "Double Stopping by Two Decision-Makers," *Advances in Applied Probability*, 25(2), pp. 438–452.

Williams, D. (1991): *Probability with Martingales*. Cambridge University Press.

Wiseman, T. (2017): "When Does Predation Dominate Collusion?," *Econometrica: Journal of the Econometric Society (forthcoming)*.

Yasuda, M. (1985): "On a randomized strategy in Neveu's stopping problem," *Stochastic Processes and their Applications*, 21(1), 159 – 166.