REPUTATIONAL BARGAINING WITH INCOMPLETE INFORMATION ABOUT PREFERENCES

MARCIN PĘSKI

Abstract. We study a reputational war-of-attrition bargaining over a pie with heterogeneous parts, incomplete information over preferences, and behavioral types. To screen across preference types, each player may demand that the opponent chooses from an arbitrary menu of offers. When there is one-sided uncertainty about preferences (and two-sided about the behavioral type), there is a unique limit equilibrium, in which the player with known preferences proposes a menu of all allocations that give him at least his worst complete information bargaining payoff. The outcome is ex ante and ex post efficient. Being able to commit to a menu instead of a single-offer increases equilibrium payoffs of the player with known preferences. Multiple equilibria are possible with two-sided incomplete information about preferences.

1. Introduction

This paper studies bargaining over a heterogeneous pie, when the preferences over the relative value of the components of the bargaining object are unknown. Such uncertainty is a common feature of complex negotiations. For instance, the EU officials likely began the Brexit talks without fully understanding the relative value for their British counterparts of the Irish border issue, the access to the common market, or

---

Date: October 9, 2021.

I am grateful to V Bhaskar, Philippe Jehiel, Thomas Wiseman, participants of the talks at Washington University in St. Louis, UCL, CETC Montreal, SAET and SITE at Stanford, the editor, and the two referees for helpful comments. I gratefully acknowledge the financial support from the Insight Grant of the Social Sciences and Humanities Research Council of Canada and the hospitality of HEC Paris, where parts of this research were completed.

fishing rights. An employer negotiating wage and/or employment reduction may not know which of those two is more acceptable for a labor union.

A typical feature of complex negotiations is that parties make sophisticated offers, like presenting a menu of options to choose from.[1] For instance, British negotiators were offered a choice between full control over their borders but excluding N. Ireland or reduced sovereignty over larger area. With an incomplete information about preferences, menus have a natural advantage over simple offers. A simple offer can be acceptable to some, but not all preference types of the opponent; menus can be acceptable regardless of what is her true preference type. By relying on static screening methods, menus may reduce or eliminate complications and inefficiencies due to delay that comes with dynamic revelation of information.

The goal of this paper is to study consequences, costs and benefits of using menus as offers under uncertainty over opponent preferences. We are interested in the following questions. First, is uncertainty advantageous in bargaining? On one hand, the revelation of private information might help to identify Pareto-optimal trades (i.e., "deals" in the language of Jackson *et al.* (2020)). At the same time, players may want to conceal or misreport their private information to improve their bargaining position. Second, what is the role of menus? Do they change the outcome of the bargaining? Are they beneficial for the parties? And how does their role interact with uncertainty? Finally, we are interested in the Nash program and connections between strategic and cooperative bargaining models. A typical axiomatic solution under uncertainty is presented as a mechanism, rather than an simple offer (Harsanyi and Selten 1972, Myerson 1984).

[1]The author had an opportunity to observe the bargaining over a pension plan reform that took place in 2016/18 between three Ontario universities and the representatives of faculty and staff. Among others, the parties negotiated the size of the spousal benefit, early retirement options, inflation indexation, etc. It was understood that the universities care only about the total actuarial cost. The preferences of the labor side were uncertain, mostly due to the heterogeneity of the labor side (for instance, the staff, but not the faculty, valued the early retirement more than the spousal benefit). The negotiations were preceded by months of meetings and consultations. The bargaining itself was very fast and it took a weekend in a hotel in downtown Toronto. In the end, the universities proposed a menu of options, and the labor side chose one of the options.

Does expanding of the space of offers pushes strategic models of bargaining towards their axiomatic counterparts?

We study these questions in a reputational model of war-of-attrition bargaining, where players build reputation for being stubborn, in order to increase the likelihood that their opponent concedes to their offer (Kambe (1999), Abreu and Gul (2000)). The behavior in such models has a very intuitive interpretation and it is also quite tractable. Such models have also proved very successful in fulfilling the Nash program in the case of complete information about preferences.

In our model, Alice and Bob want to divide a heterogeneous pie with $N \geq 2$ parts (chocolate, vanilla, etc.). See Figure 1.1; here, Alice's utility increases in the NE direction, and Bob's utility in the SW direction. For the main result, we assume that Bob's preferences are publicly known and he has beliefs about Alice's preferences. The players, one after another, choose their demands. Each demand takes form of a menu of allocations, with the interpretation that, if accepted, the opponent is free to choose any allocation in the menu. After the menu choice, the players learn whether they are rational or stubborn (Kambe (1999)). Next, a war-of-attrition bargaining commences. The first player to concede accepts the menu of the opponent, in which case the game ends. The stubborn type never concedes. Importantly, once chosen, the players do not have an opportunity to revise their demands. We believe that this is a reasonable assumption in situations when the object of bargaining is very complicated, preparing an offer takes significant resources (time, lawyers, consultations with stakeholders), and the bargaining process itself is fast (see footnote 1). We are interested in the limit, where the discrete time decisions in the war of attrition are taken more and more frequently. Further, we focus on the rational limit, where the probability of the stubborn type is arbitrarily small. The role of the behavioral types is to provide an equilibrium selection in the war of attrition.

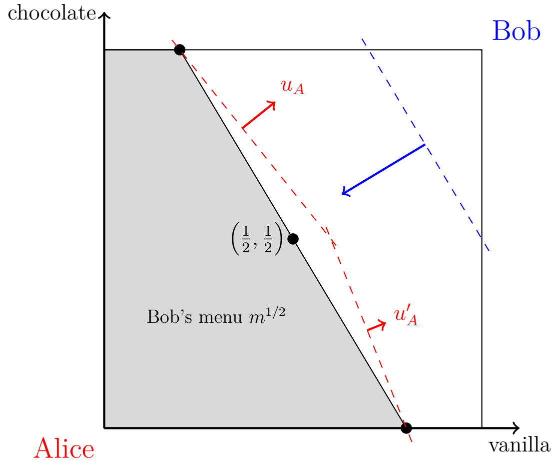

FIGURE 1.1. Bargaining environment.

A special case of the model is when there is a complete information about Alice's preference type (with an incomplete information about players' behavioral types remaining). In such a case, like in the earlier reputational literature (Kambe (1999), Abreu and Gul (2000)), in the unique limit of equilibria, both players receive their Nash solution payoffs.

The main result of the paper characterizes the limits of equilibrium payoffs when the distribution of Alice's types has a full support. There is essentially unique equilibrium outcome. The limit behavior does not depend on any details of the distribution, apart from its support. There is no significant delay in equilibrium, and the outcome is ex ante and ex post efficient.

The limit equilibrium outcome has a simple description: Bob proposes a menu of all allocations that gives him at least his worst complete information payoff across all Alice's types, and Alice chooses her favorite allocation from the menu. If we normalize Bob's payoffs so that he values the whole pie at 1, then his worst payoff is equal to $\frac{1}{2}$, and we refer to Bob's equilibrium menu as $m^{1/2}$. When $N = 2$; the equilibrium menu is drawn on Figure 1.1.

In the limit, neither Alice's or Bob's payoffs depend on his beliefs, as long as the beliefs have full support. Comparing to the complete information case, Alice's benefits and Bob loses from uncertainty about her preferences.

We describe the intuition for the result. Abreu and Gul (2000) define a strength of a player as the ratio of the payoff from winning divided by the payoff from losing the war of attrition. The stronger player must concede fast to make the weaker opponent indifferent; to compensate, the weaker opponent must begin the game with an atomic concession. We can extend the notion to one-sided incomplete information. Alice's types sort by strength, with weaker types conceding first. Bob's strength is a priori not well-defined as his payoff from winning depends on which allocation in his menu is chosen by Alice, that, in turn, depends on her type. Instead, we define Bob's strength *as if* he faces the strongest type of Alice. This definition works because most of the war-of-attrition learning is spent when Bob believes that her preference type is close to her strongest type (see Fanning (2016) and Abreu *et al.* (2015) for a similar observation). In the rational limit, the weaker player must concede with a probability close to 1 in the initial periods of the game. We show that, if the demand of the other player is incompatible, Bob's choice of menu $m^{1/2}$ or Alice's choice of her favorite allocation from this menu always makes them stronger.

Because in the limit equilibrium Bob ignores his beliefs and behaves as if facing the worst (for him) type of Alice, the result has a "Coasian" flavor (Gul *et al.* 1986). At the same time, there are important differences. First, in the Coasian bargaining literature, the ordering of types is exogenous, typically from the highest to the lowest value. Here, the sorting by strength as well as the strongest type of Alice depends endogenously on the menus demanded by each player. It is only in the equilibrium that the strongest type coincides with the ex ante worst type for Bob. Second, because the concession game is two-sided, both players keep some amount of bargaining power and Bob receives positive share of surplus.

Third, the Coasian result from this paper arises in a setting with some commitment power due to reputational types. Interestingly, this commitment power seems important. In a companion paper Peski (2019), we analyze the same bargaining environment with one-sided incomplete information, but with two differences: (a) there are no reputational types and (b) the bargaining protocol takes the form of alternating offers, where an offer takes form of an arbitrary mechanism. (If the offer is accepted, the players implement the offered mechanism, which determines the final allocation. The mechanisms include singleton offers, menus, but can be more general.) There is essentially a unique equilibrium outcome, which is non-Coasian: Bob proposes an optimal (for him) screening menu m^* that gives each type of Alice at least the same payoff as under complete information about her type. Bob is better-off under m^* than under $m^{1/2}$; each Alice's type is worse off. In other words, Bob is better-off without reputational commitment.

To understand the role of menus, we consider a version of the model where players are only able to make singleton offers. In such a case, Bob's payoff is generically strictly less than $\frac{1}{2}$. Thus, Bob suffers from not being able to offer a menu. There is a simple intuition for this result: Bob must commit himself to an offer before learning Alice's type. If Bob is only able to commit to a single offer, he faces a risk that the offer is not acceptable or efficient against the true Alice's type. This risk eliminates his commitment power and reduces his bargaining payoff.

With two-sided incomplete information about preferences, there is no natural notion of strength and no a priori sorting. In fact, we show with a two-type example that the war-of-attrition stage can have multiple equilibria. On the other hand, when $N = 2$ and the types are drawn from a continuum, and each player's menu is linear, we show that partial sorting can be restored and there is a unique equilibrium in the war-of-attrition. We define the strength of a player as the winning/concession ratio under the restriction that, when conceding, the player must choose an allocation that belongs to the diagonal (i.e., each part of the pie is divided using the same ratio). Because of linearity of preferences, the strength does not depend on the player's type. In the equilibrium, the weaker player concedes in the early periods of the game with a probability arbitrarily close to 1.

A substantial literature studies bargaining under uncertainty. The strategic literature either focuses on one-dimensional or two-type cases, with a special kind of uncertainty, including the uncertainty about values (Gul *et al.* 1986), the discount factor and time preferences (Rubinstein 1985, Abreu *et al.* 2015), or bargaining postures (Myerson (1991), Abreu and Gul (2000), Kambe (1999), Fanning 2016).

Abreu and Gul (2000) study a generalized protocol of the alternating-offer bargaining (Rubinstein 1982) with a two-sided possibility of behavioral types who never accept any offer worse than their fixed demand. The behavior in the game looks like the war of attrition that ends when one of the players reveals herself to be rational. When that happens, an earlier result by Myerson (1991) shows that the revealed player will concede quickly in any equilibrium. Kambe (1999) studies a model where players learn their commitment type after the initially chosen menu and the strategic types are not able to revise their offers upwards. The main difference with our model is that we assume that the (rational) players cannot revise their offers. We do not know whether a version of Myerson's result holds in our context.

The Nash program (originated in Nash (1953)) studies strategic foundations of cooperative games. Among others, the papers from Rubinstein 1982 through Myerson 1991 to Abreu and Gul (2000) provide such foundations for the complete information Nash bargaining solution Nash Jr (1950) (see Serrano (2004) for the overview of the literature). Recently, de Clippel *et al.* (2019) proposed a strategic model to implement the solution from Myerson 1984 under the assumption of verifiable types. The message coming from this paper is more confusing. First, the menu $m^{1/2}$ is not part of the axiomatic solutions proposed in Harsanyi and Selten 1972 and Myerson 1984; we do not know of any axiomatic model with one-sided incomplete information for which $m^{1/2}$ arises as a solution. Second, as we explain above, the companion paper (Peski (2019)) shows that an alternating-offer bargaining without reputational types have yet a different solution. We conclude that the incomplete information breaks the connections between the axiomatic approaches as well different versions of the bargaining game.

Our paper is not the first one to use menus in bargaining. In the context of the Coasian bargaining, Wang 1998 studies a similar bargaining environment with two-dimensional allocations, two types for Alice, and with Bob (the uninformed side) making all offers. He shows that, in the unique equilibrium, Bob separates the two-type of Alice by using an optimal screening contract. In particular, the Coase conjecture fails as Bob keeps all power subject to the incentive compatibility constraints. More recently, Strulovici 2017 works with a closely related model, but he assumes that, instead of ending the game, any accepted offer becomes a status quo for future bargaining. He shows that the uninformed player is unable to offer an inefficient payoff to type u_1' in order to screen out the more extreme type u_1'' and a Coasian result prevails. An alternating-offer bargaining with menus is studied in Sen 2000 (see also Inderst 2003), who studies the two-type case and establishes the uniqueness of equilibrium that satisfies a perfect sequential refinement of Grossman and Perry (1986).

Jackson *et al.* (2020) studies an alternating offer bargaining game and with menus as offers. Similarly as in our paper, the equilibrium outcome is efficient, but for different reasons. Although the authors allow for incomplete information on both sides, they make a strong assumption that the total value of bargaining surplus is commonly known. This assumption implies that there are no incentive problems that stop agents from truthfully revealing their information. In the unique equilibrium, the agents use menus to implement information revelation in a single round of bargaining.

2. Model

Two players, Alice and Bob, $i = A, B$, bargain over a heterogeneous pie with $N \geq 2$ parts. An allocation is defined as $x \in X := [0,1]^N$. Each player has a linear preference over allocations $u_i \in \mathcal{U} := \left\{u \in \mathbb{R}_+^N : \sum u^n = 1\right\}$. (The normalization is w.l.o.g.) The payoffs from allocation x is equal to $u_A(x) = \sum_n u_A^n x_n$ for Alice's type u_A and $u_B(x) = 1 - \sum_n u_B^n x_n$ for Bob's type u_B. We assume that spaces X and $\mathcal{U}$ are equipped with the "sup" metric.

2.1. **Bargaining.** In the baseline model, Bob's preferences u_B are publicly known. Initially, Alice has private and partial information about her preferences in the form of a signal $s \in \mathcal{U}$ drawn from distribution π_A.

The game has two stages: a menu choice followed by a war-of-attrition bargaining. In the first stage, player $k = A, B$, followed by $-k$ publicly chooses her or his bargaining demand. The demand of each player i takes form of a menu $m_i \subseteq X$, with the interpretation that $-i$ is offered to pick any allocation in m_i. After the choice of menu, Alice privately learns her preference type u_A drawn from distribution $\pi_S(.|s) \in \Delta\mathcal{U}$. Simultaneously and independently, each player learns whether she or he is the stubborn type. The probability of the stubborn type is the same and equal to $\lambda \in (0, 1)$.

Next, the war of attrition commences. In alternating periods, starting with player k in period 1, the normal type of each player i decides whether to continue or concede. If he or she continues, the game moves to the next period and the other player. If she or he concedes, she must choose an allocation x from menu m_{-i}. The stubborn type never concedes. The players maximize the expected utility and they discount with a common factor $\mathrm{e}^{-\Delta}$, where Δ represents the length between two subsequent decision points.

Alice's imperfect information about her own type alleviates the signaling problem inherent to bargaining with private information. We make the following assumption:

Assumption 1. *(Common support) For each s, $\mathrm{supp}\pi_A = \mathrm{supp}\pi_S(.|s) = \mathcal{U}_A \subseteq \mathcal{U}$ for some closed convex set $\mathcal{U}_A$. There $K < \infty$ such that for each s, $\pi_S(.|s)$ has a Lipschitz continuous density with constant K with respect to the Lebesgue measure on $\mathcal{U}$.*[2]

[2]For each convex subset $U \subseteq \mathbb{R}^N$, one can find its affine hull, i.e., the intersection of all affine spaces that contain U. The Lebesgue measure on the affine hull assigns positive mass to U. Whenever we mention "the Lebesgue measure on U", we refer to the restriction of such a measure to set U.

The first part of the assumption says that the support of Alice's type distribution given signal does not depend on the signal. It follows that the support of Bob's posterior beliefs after she chooses her menu does not depend on the menu. The second part of the assumption ensures that the posterior beliefs are sufficiently regular: the posterior belief over Alice's types is Lipschitz.[3] The regularity plays an important role in the proofs of subsequent results, including the proof of Lemma 2. The assumption does not eliminate signaling completely, as Bob's beliefs may (and, typically, will) depend on Alice's choice.

Because $\mathcal{U}_A$ can be a singleton, a complete information about Alice's preferences is a special case of the model.

The role of the behavioral types is to pin down the equilibrium in the war-of-attrition stage; it is well known that, without them, the war-of-attrition games have a continuum of equilibria. See Kambe (1999) for a similar approach to the behavioral types.

2.2. **Equilibrium.** Let $\mathcal{C}_X$ be the space of closed subsets of X with the topology of the Hausdorff distance. Let $M_i \subseteq \mathcal{C}_X$ be a compact set of possible menu choices. A strategy profile in the menu choice game is a pair of measurable mappings $m_k : \mathcal{U}_k \to \Delta M_k$, $m_{-k} : \mathcal{U}_{-k} \times M_k \to \Delta M_{-k}$, where we take $\mathcal{U}_B = \{u_B\}$. After Alice chooses her menu, Bob updates his beliefs about her prior signal and forms beliefs about Alice's true preferences.

Let T_i be the set of periods in which player i makes a decision in the war of attrition. A strategy of the (normal type of) player i is a pair $\sigma_i = \left(\sigma_i^T, \sigma_i^M\right)$ of measurable stopping time $\sigma_i^T : \mathcal{U}_i \to \Delta T_i$ and a choice $\sigma_i^M : \mathcal{U}_i \to \Delta m_{-i}$. In each period, each player has a belief about the probability that the opponent is stubborn; Bob also has beliefs about Alice's type.

[3]Indeed, suppose that $p\,(.|m) \in \Delta\mathcal{U}$ are the posterior beliefs over Alice's signals after her choice of menu m. Belief p will typically depend on Alice's strategy in the menu choice game. Bob's beliefs about Alice's preference type are given by $\pi\,(W) = \int \pi_S\,(W|s)\,p\,(ds|m)$ for each measurable set $W \subseteq \mathcal{U}_A$. Clearly, π have support equal to $\mathcal{U}_A$ and Lebesgue density that is continuous and Lipschitz with constant K.

A *Perfect Bayesian equilibrium* is a profile of strategies and beliefs such that players best respond to each other and the beliefs are derived from strategies through the Bayes formula whenever possible and player's beliefs do not change after his or her own out-of-equilibrium action.

Let $E_i(u_i; \Delta, \lambda, M_., \pi, \pi_S, k)$ be the set of equilibrium payoffs of player i type u_i in the game where player k makes the first choice. We are interested in the following limits:

- the initial information becomes approximately perfect: $\pi_S(.|s) \to \delta_s$, weakly, for each $s \in \mathcal{U}_A$; we write $\pi_S \to \delta$,
- the game approximates continuous time $\Delta \to 0$, and players become fully rational, $\lambda \to 0$.

Define the limit set of equilibrium payoffs of type $u_i$of player i as

$$E_i(u_i; M_., \pi, k) = \lim \sup_{\pi_S \to \delta} \lim \sup_{\lambda \to 0, \Delta \to 0} E_i(u_i; \Delta, \lambda, M_., \pi, \pi_S, k).$$

The order of the last two limits is not important.

2.3. **Nash allocations and Coasian menu.** Before we proceed with the analysis of the model, we define two important objects. For each type u_A of Alice, define a Nash allocation $x^{\text{Nash}}(u_A)$ as the maximizer of the product of Alice's and Bob's utilities:

$$x^{\text{Nash}}(u_A) \in \arg\max_{x \in X} u_A(x) u_B(x).$$

Nash Jr (1950) proposed such an allocation as the unique solution to the bargaining problem that satisfies a series of axioms. Later, Rubinstein (1982) showed that Nash allocations arise in a strategic alternating-offer bargaining model.

It is instructive to illustrate Nash allocations in a special case when $N = 2$. See Figure 2.1. Assume w.l.o.g. that $u_B^1 > u_B^2$. There are four distinct cases:

- If $u_A^1 = u_B^1$, i.e. Alice preferences are the same as Bob, then the Nash solution awards payoff of $\frac{1}{2}$ to each player; any allocation on Bob's indifference line

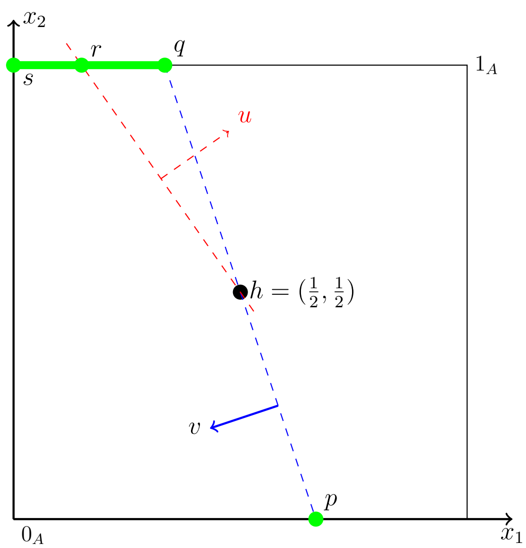

FIGURE 2.1. Nash allocations.

corresponding to payoff $\frac{1}{2}$ (the dashed line between allocations $p = \left(\frac{1}{2u_B^1}, 0\right)$[4] and $q = \left(1 - \frac{1}{2u_B^1}, 1\right)$) is a solution to the Nash bargaining problem.

- If $u_A^1 > u_B^1$,, i.e. Alice likes the first part of the pie more than Bob, she is going to get her favorite allocation subject to the constraint that Bob's payoff is at least $\frac{1}{2}$, i.e., allocation p. In such a case, Bob's payoff is $\frac{1}{2}$ and Alice gets a payoff that is strictly larger than $\frac{1}{2}$.
- If $\frac{1}{2} \leq u_A^1 < u_B^1$, i.e. Alice prefers the first part of the pie to the second, but the intensity of her preference for the first part is smaller than Bob's, Bob receives his favorite allocation subject to the constraint that Alice's payoff is at least $\frac{1}{2}$, i.e., allocation $r = \left(1 - \frac{1}{2u_c}, 1\right)$. The allocation and Bob's payoff depends on Alice's preference; Alice's payoff is $\frac{1}{2}$.
- Finally, if $u_A^1 < \frac{1}{2}$, i.e., Alice prefers the second part, each player receives his or her favorite part of the pie (allocation s).

[4]Here, and in the rest of the paper, unwasteful allocations are described by Alice's shares.

Notice that if Alice's preference type is in the third and fourth cases above, i.e., $u_A^1 < u_B^1$, Alice prefers allocation q to allocations r or s. Because allocation q can be obtained as an outcome of Nash bargaining if Bob thinks that Alice preferences are the same as his, such types of Alice could benefit from Bob not knowing their true preferences.

The incentive problem is also present more generally, when $N > 2$. In such a case, each Nash allocation leads to a payoff of at least $\frac{1}{2}$ for Bob. The worst Bob's payoff is attained when Alice's types has exactly the same preferences as Bob. Each type of Alice would like Bob to believe that her preferences are as close to his as possible.

Suppose that we ignore the incentive problem, ask Alice to report her type, and then implement the Nash allocation given her report. This is equivalent to allowing Alice to mimic arbitrary preference type, including a type with the same preferences as Bob's. If $N = 2$, a generic type of Alice would choose between allocations q or p. More generally, when $N \geq 2$, Alice would choose optimally from menu

$$m^{1/2} = \left\{x : u_B\left(x\right) \geq \frac{1}{2}\right\}.$$

Menu $m^{1/2}$ consists of allocations in which Bob receives at least his worst payoff among all Nash allocations and types of Alice. If implemented, such a menu has three features: (a) if Alice chooses optimally, the final allocation is ex-post efficient, (b) the chosen allocations are the worst Nash allocations from Bob's point of view, and (c) for Alice, they constitute the best choices among all allocations that can be obtained as Nash solution for some Alice's preference type. Such an outcome resembles the famous result from the Coasian bargaining literature on the durable good monopolist without commitment (Gul *et al.* (1986)). In that literature, in the gap case, the solution shares the same three features. For this reason, we refer to menu $m^{1/2}$ as the *Coasian menu* and 1/2 as the *Coasian payoff.* [5]

[5]The Coasian literature focuses on bargaining models where all the offers are made by the uninformed party (i.e., Bob). We believe that it is worthwhile to separate bargaining power from incomplete information, and it is the latter that is central to the Coasian logic.

3. Analysis

This section is divided into four parts. First, we characterize the solution in the special case of complete information. Second, we turn into incomplete information case. We define a notion of strength for each player that is appropriate for incomplete information game. That notion will be useful in the analysis of the war-of-attrition game in the second part. Finally, we prove our main result - the characterization of the limit equilibrium payoffs in the menu-choice game.

3.1. Complete information about Alice preferences.

We start with discussing the special case of our model, where $\mathcal{U}_A = \{u_A\}$ for some preference type u_A, or, in other words, where Alice's preferences are known to both agents before the choice of menu.

Theorem 1. *Suppose that* $\mathcal{U}_A = \{u_A\}$*. Then, for each player* $k = A, B$*, each* $j = A, B$*,*

$$E_k\left(u_k; \mathcal{C}_X, \pi, j\right) = u_k\left(x^{Nash}\left(u_A\right)\right). \tag{3.1}$$

If Alice's preferences are known, then in the continuous time and rational limit, both players receive their Nash payoffs. This result coincides with the message of reputational literature, including Kambe (1999) and Abreu and Gul (2000). The earlier papers only allow players to make simple offers. Allowing for more sophisticated offers in the form of menus does not change the outcome of bargaining.

The proof is standard. We remind it in order to draw the comparison with the incomplete information case. Abreu and Gul (2000) define a strength of a player as the ratio of the payoff from winning (i.e., the payoff if the opponent concedes) divided by the payoff from losing (i.e., conceding) the war of attrition. The stronger player must concede fast to make the weaker opponent indifferent; to compensate, the weaker opponent must begin the game with a large probability concession. The notion of strength extends to menus as offers in a natural way, where a winning and losing payoff depends on the optimal choice from a menu.

Suppose that one of the players chooses a single-element menu $\left\{x^{\text{Nash}}\left(u_A\right)\right\}$. If the other player responds with a menu that gives the first player payoff not much smaller than the payoff from the Nash allocation, the first player receives her Nash payoff. Alternatively, if the other player chooses a menu in which the best payoff of the first player is significantly smaller than her Nash payoff, we can show that the first player is stronger, and the second player must start with a concession. In the rational and patient limit, the probability of the initial concession is arbitrary close to 1. On the other hand, if the first player makes an offer that does not guarantee the second player his Nash payoff, the second player has a counter-offer that ensures him the Nash payoff and simultaneously makes him stronger, ensuring his victory in the bargaining game.

In any case, both players are able to guarantee the Nash payoffs. Because the Nash payoffs are efficient, none of the players can do better. The details can be found in Appendix B.

3.2. **Strength.** Next, we turn to the incomplete information case and we assume that set $\mathcal{U}_A$ has a non-empty interior. We begin with the war-of-attrition stage. A generic Alice's menu m_A contains the unique Bob's optimal allocation $x_A \in \arg\max_{x \in m_A} u_B\left(x\right)$. In such a situation, x_A is the only allocation chosen by Bob when he concedes and that plays any role in the bargaining game. As a first approximation, we assume that Alice's menu consists of a single element, $m_A = \{x_A\}$; we deal with the non-generic case later. We refer to x_A as Alice's winning allocation.

Given the assumption, we define the strength of Alice's type u_A as

$$S_A\left(u_A\right) = \frac{u_A\left(x_A\right)}{\max_{x \in m_B} u_A\left(x\right)}.$$

As in Abreu and Gul (2000), the strength is defined as a ratio of the payoff from winning the war of attrition (i.e., from the winning allocation x_A) versus the payoff from losing, that is equal to the best payoff that type u_A can attain in menu m_B.

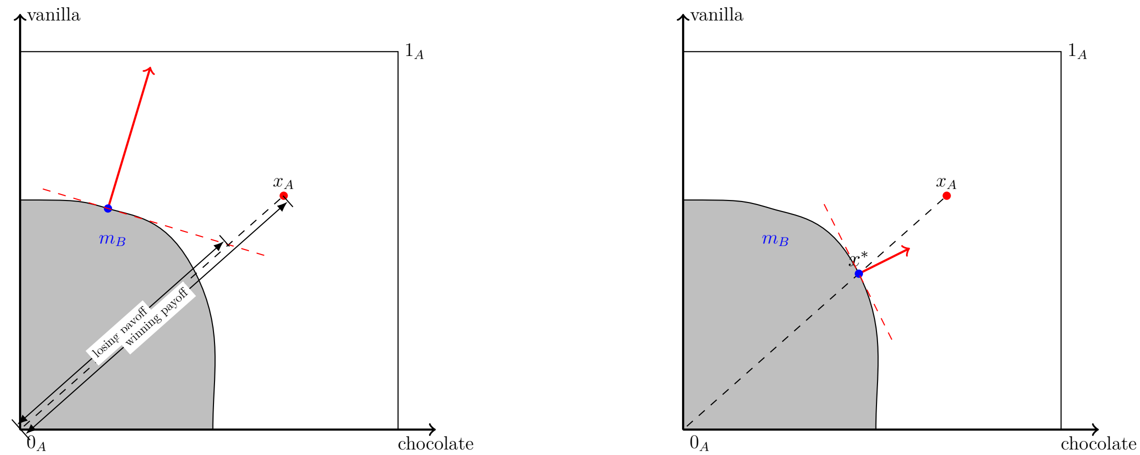

FIGURE 3.1. Alice's strength and general menus.

Define the strength of Alice the player as the strength of the strongest type in her support:

$$S_A^* = \max_{u_A \in \mathcal{U}_A} S_A(u_A).$$

A geometric intuition on how to find the strongest type is illustrated on the left side of Figure 3.1. The dashed ray connects allocations $\mathbf{0}_A$ (i.e., Alice gets nothing) and the winning allocation x_A. The win/loss ratio is equal to the ratio of the length of the ray and the distance between allocations $\mathbf{0}_A$ and the intersection of the ray and Alice's indifference hyperplane that corresponds to her losing payoff. By moving the losing indifference hyperplane along the menu boundary, we can see that the ratio is maximized when the indifference hyperplane touches the menu exactly at the ray. More formally, let

$$m_B^* = \bigcap_{u_A \in \mathcal{U}_A} \left\{ x : u_A(x) \leq \max_{x' \in m_B} u_A(x') \right\} \tag{3.2}$$

be the largest menu that gives each of Alice's types exactly the same utility as the menu m_B. We say that m_B^* is a *completion* of m_B. Further, let

$$\kappa^* = \sup\left\{\kappa \in [0,1] : (1-\kappa)\,\mathbf{0}_A + \kappa x_A \in m_B^*\right\}, \text{ and} \tag{3.3}$$
$$x^* = (1-\kappa^*)\,\mathbf{0}_A + \kappa^* x_A.$$

Allocation x^* is the optimal allocation that belongs to menu m_B^* and to the ray connecting allocations $\mathbf{0}_A$ and x_A (see the right panel of Figure 3.1).

Given the above preparation, notice that the strength of an arbitrary type u_A is equal to

$$S_A(u_A) = \frac{u_A(x_A)}{\max_{x\in m_B^*} u_A(x)} = \frac{\frac{1}{\kappa^*}u_A(x^*)}{\max_{x\in m_B^*} u_A(x)} \leq \frac{1}{\kappa^*},$$

where the inequality follows from the fact that $\max_{x\in m_B^*} u_A(x) \leq u_A(x^*)$ (due to $x^* \in m_B^*$). The upper bound is reached by a type $u_A^* \in \mathcal{U}_A$ such that $x^* \in \arg\max_{x\in m_B^*} u_A^*(x)$. (Such a type exists, due to the menu m_B^* being complete.) Thus, x^* is a (possibly, one of many) allocation(s) of the (possibly, one of many) strongest type(s), and

$$S_A^* = \frac{1}{\kappa^*}. \tag{3.4}$$

We define Bob's strength as the ratio of the payoff from allocation x^* (i.e., the winning allocation against the strongest type) and his losing payoff:

$$S_B^* = \frac{u_B(x^*)}{u_B(x_A)}. \tag{3.5}$$

As we show soon (see Lemma 2 below), the strength comparison plays an important role in determining the outcome of the war-of-attrition. The next result describes an important special case. Alice is stronger when she makes an arbitrary Coasian offer to Bob (i.e., an offer that belongs to the Coasian menu) and that offer is not included in his (completed) menu.

Lemma 1. *Suppose that Alice's offer* $m_A = \left\{x^A\right\}$ *consists of a single allocation such that* $x^A \in m^{1\backslash 2}$ *and* $x^A \notin m_B^*$*. Then,* $S_A^* > S_B^* > 1$.

Proof. The assumptions imply that there is $r > 0$ such that $B\left(x^A, r\right) \cap m_B^* = \emptyset$, where $B\left(x^A, r\right)$ is a ball with center at x^A and radius r. Let x^* and κ^* be defined as above. Then, $\kappa^* \leq 1 - r$ and $u_B\left(x^*\right) = 1 - \kappa^* + \kappa^* u_B\left(x_A\right)$. (To see the latter, notice that

$$\begin{aligned} 1 - u_B\left(x^*\right) &= 1 - \sum_n u_B^n\left(1 - x_n^*\right) = \sum_n u_B^n x_n^* \\ &= \sum_n u_B^n \kappa^* x_{A,n} = \kappa^* - \sum_n u_B^n \kappa^*\left(1 - x_{A,n}\right) = \kappa^*\left(1 - u_B\left(x_A\right)\right).) \end{aligned}$$

We compute Bob's strength:

$$S_B^* = \frac{u_B\left(x^*\right)}{u_B\left(x_A\right)} = \frac{\left(1-\kappa^*\right) + \kappa^* u_B\left(x_A\right)}{u_B\left(x_A\right)} = \frac{1-\kappa^*}{u_B\left(x_A\right)} + \kappa^* \leq \frac{1-\kappa^*}{1/2} + \kappa^* = 2 - \kappa^* \geq 1 + r.$$

The inequality comes from the fact that $x_A \in m^{1/2}$, hence $u_B\left(x_A\right) \geq \frac{1}{2}$. But because $\kappa^* \leq 1 - r$, we have

$$S_A^* - S_B^* = \frac{1}{\kappa^*} - 2 + \kappa^* \geq \frac{1}{1-r} + 1 - r - 2 = \frac{r^2}{1-r} > 0.$$

□

Remark. The proof of the Lemma establishes a slightly stronger claim: for each $r > 0$, there exists $\delta > 0$ such that if $m_A = \{x_A\}$ for some allocation x_A such that $B\left(x_A, r\right) \cap m^{1/2} \backslash m_B^* = \emptyset$, then $S_A^* > S_B^* > 1$ and $\min\left(S_A^* - S_B^*, S_B^* - 1\right) > \delta$.

Remark. An equally simple argument shows that if $m_B^* \supseteq m^{1/2}$ and $m_A = \{x_A\}$ for some $x_A \notin m_B^*$, then $S_A^* < S_B^*$.

3.3. **War of attrition.** The notion of strength leads to the following simple partial characterization of the behavior during the war-of-attrition.

Lemma 2. *Suppose that Assumption 1 holds, $\mathcal{U}_A$ has an open interior, $u_B \in \mathcal{U}_A$, and Bob's beliefs about Alice's type in the beginning of the war-of-attrition are derived from Alice's behavior in the menu choice game. Additionally, suppose that $m_A = \{x_A\}$, $x_A \notin m_B^*$, and that $0 \ll x_A \ll 1$. If $S_A^* > S_B^* > 1$, then, for each $\delta > 0$, there exist $\lambda^*, \Delta^* > 0$ such that if $\lambda \leq \lambda^*$ and $\Delta \leq \Delta^*$, then, there is $T < \infty$ such that $\mathrm{e}^{-\Delta T} > 1 - \delta$ and, in any equilibrium, Bob concedes with a probability of at least $1 - \delta$ before the end of period T.*

The Lemma is a generalization of the argument sketched in section 3.1 to an incomplete information. If Alice is stronger, $S_A^* > S_B^*$, Bob loses the war of attrition, i.e., there is a high probability that he concedes in the initial periods of the game. The payoffs depend on the support (through the definition of Alice's strength), but not on any other details of Bob's beliefs.

3.3.1. *Equilibrium description.* We describe the intuition behind the Lemma 2. Initially, we make a simplifying assumption that, as in the right panel of Figure 3.1, the strongest of Alice's types is unique and it has the unique optimal allocation x^*. In such a case, types that are close to the strongest one pick an allocation that is close to x^*.

We start with reminding the basic structure of equilibria in the war-of-attrition games. In each period, a player who has not yet revealed herself or himself to be stubborn must concede with a positive probability. The last concessions take place in two consecutive periods and occur in finite time. Let T^* be the last period in which a concession occurs.

While concessions take place, players learn about the rationality of their opponent; additionally, Bob learns about Alice's preference type. Let $p_B(t)$ be Bob's concession rate, i.e., the probability that Bob concedes in period $t \in T_B$ conditionally on reaching

t. Alice's type u_A gain from waiting from period $t-1$ to $t+1$ is equal to

$$\begin{aligned}&\left[\mathrm{e}^{-\Delta}p_B\left(t\right)\left(u_A\left(x_A\right)\right)+\mathrm{e}^{-2\Delta}\left(1-p_B\left(t\right)\right)\max_{x\in m_B^*}u_A\left(x\right)\right]-\max_{x\in m_B^*}u_A\left(x\right)\\&=\left[\mathrm{e}^{-\Delta}p_B\left(t\right)\left(S_A\left(u_A\right)-\mathrm{e}^{-\Delta}\right)-\left(1-\mathrm{e}^{-2\Delta}\right)\right]\max_{x\in m_B^*}u_A\left(x\right),\end{aligned}\tag{3.6}$$

and it is single-crossing in strength. It follows that her weaker types must concede before the stronger ones. Note that the strength order as well as the identity of the strongest type depends on the menu.

For each $t\in T_B$ and $t<T^*$, let $S_A\left(t\right)$ be the strength of Alice's weakest type who has not conceded before period t. In equilibrium, Bob's concession rate $p_B\left(t\right)$ makes the type $S_A\left(t\right)$ indifferent between conceding now and waiting till the next opportunity. The above formula implies that

$$p_B\left(t\right)=\frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}}\frac{1}{S_A\left(t\right)-\mathrm{e}^{-\Delta}}.$$

Similarly, for each $t\in T_A$ and $t<T^*$, let $p_A\left(t\right)$ denote Alice's concession rate. Analogous calculations show that

$$p_A\left(t\right)=\frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}}\frac{1}{\frac{u_B\left(w_A\left(t\right)\right)}{u_B\left(x_A\right)}-\mathrm{e}^{-\Delta}},$$

where $w_A\left(t\right)$ is the average allocation chosen by types of Alice who concede in period t:

$$w_A\left(t\right)=\frac{\int_{u_A:S_A(t-1)\le S_A(u_A)<S_A(t+1)}\left(\arg\max_{x\in m_B^*}u_A\left(x\right)\right)d\Pi\left(u_A\right)}{\Pi\left\{u_A:S_A\left(t-1\right)\le S_A\left(u_A\right)<S_A\left(t+1\right)\right\}},$$

and Π are Bob's beliefs at the beginning of the war-of-attrition. Although for some Alice's types the optimal allocation is not unique, it is unique for almost all types wrt. Lebesgue measure and Assumption 1 implies that the integral in the numerator is well-defined (note that the Assumption ensures that Π is absolutely continuous wrt. Lebesgue measure).

To complete the description of the equilibrium, notice that Alice's concession rate determines the types that are indifferent between conceding in the two consecutive periods:

$$p_A(t) = \frac{(1-\lambda)\Pi\{u_A : S_A(t-1) \leq S_A(u_A) < S_A(t+1)\}}{(1-\lambda)\Pi\{u_A : S_A(t-1) \leq S_A(u_A)\} + \lambda}.$$

In the second step, we divide the time in the game into three zones:

- At the very beginning of the war-of-attrition, there might be Alice's types who prefer to concede immediately than receive immediate Bob's concession (i.e., types whose strength is smaller than 1). We show that all such types concede very fast. More precisely, Lemma 9 in the Appendix shows that, for each $\delta > 0$, if Δ is sufficiently small, then there exists T_0 such that $\mathrm{e}^{-\Delta T_0} \geq 1-\delta$, and, either Bob concedes with probability close to $1-\delta$ before period T_0 (in which case, the thesis of Lemma 2 holds, or Alice's types with strength smaller than $1+\frac{1}{2}\delta^2$ concede before period T_0. If the former, the claim we are trying to prove holds. From now on, we assume the latter. We refer to periods $t < T_0$ as the *early game.*
- *Middle game* starts at T_0 and continues to $T_A^\eta = \min\{t : S_A(t) \geq S_A^* - \eta\}$ for some small η. T_A^η is the first period after which all remaining types of Alice are η-close to her strongest type. Notice that Alice's concession rate in the middle game is bounded by b
- We refer to the time after T_A^η as the *late game.* If x^* is the unique choice of the unique strongest type of Alice, then, for sufficiently small η, the optimal choices of the types that concede in the late game are close to x^*, and, as a result, $w(t) \approx x^*$. Hence the concession rates of the two players are close to

  $$p_A(.) \approx p_A^* = \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{S_B^* - \mathrm{e}^{-\Delta}} \text{ and } p_B(.) \approx p_B^* = \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{S_A^* - \mathrm{e}^{-\Delta}}.$$

  Because $S_A^* > S_B^*$, Alice concedes at faster rate than Bob.

The rest of the argument relies on two observations. First, as $\lambda \to 0$, the logic of the Bayesian updating implies that the length of the late game dominates earlier stages, i.e., $(T^* - T_A^\eta) \gg T_A^\eta$. (The idea why the late game dominates the early game is not new and played an important role in Abreu *et al.* (2015) and Fanning (2016).) For a simple intuition, treat Bob's concession rate in the late game as constant and approximately equal to $\simeq p_B^*$. The Bayes formula implies that

$$(1 - p_A^*)^{T^* - T_A^\eta} \simeq \frac{\lambda}{\lambda + (1-\lambda)\Pi\left(u : S_A(u) \geq S_A^* - \eta\right)},$$

where the numerator is the probability that Alice never concedes and the denominator is the probability that Alice does not concede before the late game. The concession rate in the middle game is not smaller than $p_A^{\min} \approx \frac{1-e^{-2\Delta}}{e^{-\Delta}} \frac{1}{\frac{1}{u_B(x_A)} - e^{-\Delta}}$, which implies that

$$\left(1 - p_A^{\min}\right)^{T_A^\eta} \geq \lambda + (1-\lambda)\Pi\left(u : S_A(u) \geq S_A^* - \eta\right).$$

By taking logarithms, we obtain

$$\begin{aligned} T^* - T_A^\eta &\simeq \frac{1}{p_A^*}\left[\log\left(\lambda + (1-\lambda)\Pi\left(u : S_A(u) \geq S_A^* - \eta\right)\right) - \log\lambda\right], \\ T_A^\eta &\leq \frac{1}{p_A^{\min}}\left[\log\left(\lambda + (1-\lambda)\Pi\left(u : S_A(u) \geq S_A^* - \eta\right)\right)\right]. \end{aligned}$$

For sufficiently small λ, the second quantity is much smaller and $T^* - T_A^\eta \gg T_A^\eta$. Note that the argument relies heavily on the ability to keep the probability $\Pi\left(u : S_A(u) \geq S_A^* - \eta\right)$ bounded away from 0, or, in other words, on the Assumption 1.

Because Alice's concession rate in the late game is higher than Bob's, familiar arguments from Abreu and Gul (2000) imply that the ratio of the probability that Alice has not conceded before is to the analogous probability for Bob at the onset of the late game is arbitrarily large as $\lambda \to 0$. During the middle game, Bob may concede faster. But, Bob's middle-game concession rate is bounded from above by $p_B^{\max} \leq \frac{1-e^{-2\Delta}}{e^{-\Delta}} \frac{1}{\frac{1}{2}\delta^2}$ (due to the assumption about what happens in the early game stated above), and, as we observed above, Alice's middle game concession rate is not lower than $p^A_{\min}$. Because the length of the middle game is much smaller than the length of the late game for sufficiently small λ, the late-game effect dominates and the probability that Alice has not conceded before the beginning of the middle game is arbitrarily larger (as $\lambda \to 0$) than the probability that Bob has not conceded.

As a result, Bob must concede with a probability arbitrarily close to 1 during the early game.

3.3.2. *Complications.* The estimate of Bob's concession rate in the late game relies on the assumption that the late-game types lose with an allocation close to x^*. This assumption is satisfied for menus m_B, for which there is the unique strongest type of Alice and that such a type has the unique optimal allocation.

In general, there are two complications. The first one arises when the strongest type is in the boundary of the type distribution $\mathcal{U}_A$. See the left panel of Figure 3.2, where Alice chooses from a single-element menu $\{x_B\}$. In such a case, the optimal choice of each Alice's type is x_B, and not x^*. Nevertheless, the thesis of the Lemma holds. The reason is that replacing x^* by x_B does not affect Alice's strength, hence Bob's concession rates in the late game. (To make this argument general, we need to assume that $u_B \in \mathcal{U}_A$.) At the same time, because Bob prefers x^* to x_B, to keep Bob indifferent, Alice must concede at an even faster rate. Hence, the original argument applies. (The fact that $u_B \in \mathcal{U}_A$ is necessary to deal properly with this complication.)

Because of the complication, the converse to Lemma 2, i.e., a claim that Bob wins the war of attrition if he is stronger, does not generally hold. However, the converse holds in a special case. The next result shows that if Bob's menu is close to the Coasian menu $m^{1/2}$ and Alice's demand gives Bob a payoff that is significantly lower than $\frac{1}{2}$, then Alice loses the war of attrition. (Notice that Bob is stronger in such a case, due to the Remark after Lemma 1.)

Lemma 3. *For any $\eta > 0$, there is $\varepsilon > 0$ such that for each $\delta > 0$, there exist $\lambda^*, \Delta^* > 0$ such that if $\lambda \leq \lambda^*$ and $\Delta \leq \Delta^*$, and if*

$$\max_{x \in m_A} u_B(x) < \frac{1}{2} - \eta \text{ and } d_H\left(m_B, m^{1/2}\right) \leq \varepsilon,$$

then Alice concedes with a probability of at least $1 - \delta$ in his first period of the game.

The second complication is when the menu is linear in the neighborhood of x^*. In such a case, the optimal allocation of almost all types can be significantly far away from x_i^*, even during the late game. Nevertheless, we show that in such a case, the average allocation chosen by Alice, conditional on her conceding, converges to x^*. The intuition is described on the right panel of Figure 3.2. Almost all types of player i pick one of two y^1, y^2 optimal allocations. The dotted lines represent the indifference hyperplanes of the strongest types of player i that concede in periods $t-2$, t, and $t+2$. The areas between the hyperplanes contain Alice's types who concede in periods t and $t+2$. The intersection of the indifference curves belongs to the dashed ray because, as we explained above, the strength of any type can be parameterized by the distance between the zero allocation and the intersection of the ray with an indifference curve. In the late game, only the types with strength close to S_A^* survive. Due to the regularity assumptions, and specifically to the continuity of the density, in the late game, the conditional probabilities of the two optimal allocations are proportional to the angles between the two consecutive indifference curve. A simple geometric intuition shows that, in such a case, the weighted average concession allocation is close to x^*.

3.3.3. *Proof intuition - continued.* Our proof in the Appendix deals with two complications at the same time and it necessarily differs from 2-dimensional intuitions described above. Note that the argument sketched in Section 3.3.1 holds if, during the late game, Alice's concession rate is not (significantly) smaller than p_A^*, or, if Bob's utility from the average allocation chosen by conceding types of Alice $w(t)$ is not (significantly) higher than his utility from allocation x^*. In order to show the latter, we pay a closer

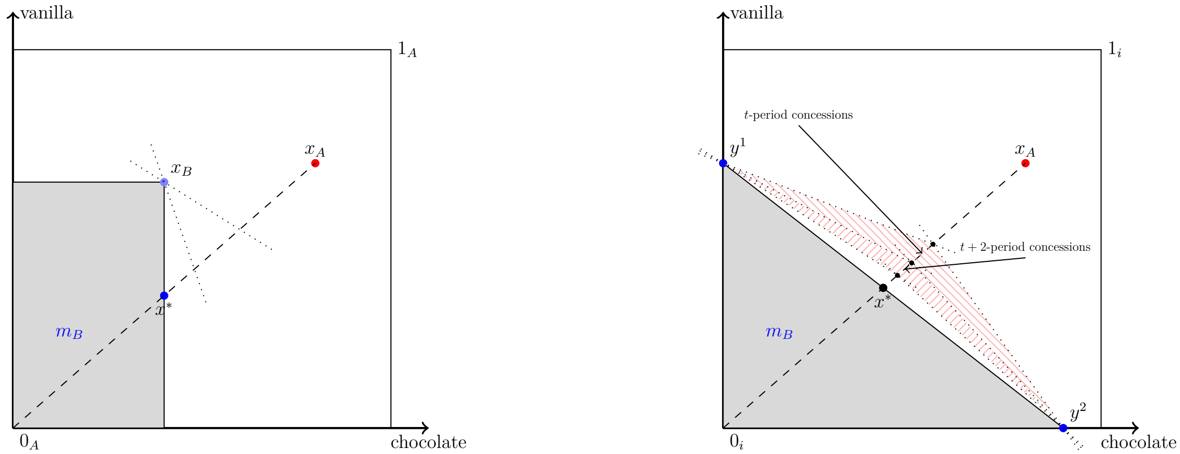

FIGURE 3.2. Special cases.

attention to Alice's optimal choices:

$$x_u \in \arg\max_{x \in m_B^*} u(x).$$

We explain how we show the latter. As we have observed above, those choices are unique almost all types $u \in \mathcal{U}$.

To in such a We show that a strength of a type is monotonically decreasing in the value of a certain convex function of a type (see equations (A.2) and (A.3) in the Appendix), and at each period, the remaining Alice's types belong to the lower contour sets of this function. Then, Lemma 7 uses convex analysis techniques to show that, in the late game, Alice concedes not slower than as if she had strength S_A^*.

3.4. **Menu choice game.** The above results lead to a straightforward characterization of the limit payoffs at the menu choice stage.

Theorem 2. *Suppose that Assumption 1 hold, $\mathcal{U}_A$ has an open interior, and $u_B \in \mathcal{U}_A$. Then, for each $u_A \in \mathcal{U}_A$, each $\pi \in \Delta\mathcal{U}$, each initial player $k = A, B$*

$$E_A(u_A; \mathcal{C}_X, \pi, k) = \max_{x \in m^{1/2}} u_A(x) \ \text{ and } E_B(u_B; \mathcal{C}_X, \pi, k) = \frac{1}{2}. \tag{3.7}$$

Theorem 2 says that Alice receives her optimal payoff from the Coasian menu, i.e., the optimal payoff subject to the constraint that Bob receives at least $\frac{1}{2}$. The latter is equal to his worst payoff in a game in which Alice's type is known (or Nash solution payoff) across all possible Alice's types.[6] Analogously, each of her opponent types receives her best payoff across all possible outcomes in complete information equilibria (i.e., across all Alice's types). The equilibrium payoffs are ex post efficient.

Although the statement focuses on the payoffs, note that the only way to obtain such payoffs is when Bob proposes menu $m^{1/2}$, which is accepted.

The assumption that $\mathcal{U}_A$ has non-empty interior ensures a non-trivial incomplete information about preferences.

The assumption $u_B \in \mathcal{U}_A$ plays an important role in the proof of Lemma 2. Without it, the thesis of Theorem 2 may not hold. (To see why, suppose that $\mathcal{U}_A \to \{u_A\}$ in the Hausdorff sense. Then, we expect Bob's equilibrium payoff to converge to his Nash payoff from bargaining with type u_A. For a large set of Alice's types u_A, Bob's Nash payoff is strictly larger than $\frac{1}{2}$.) In such a case, we hypothesize that a generalized version of Theorem 2 may hold, in which Bob's equilibrium payoff is the worst Nash payoff across all possible types of Alice. We leave this hypothesis for future research.

Proof. We sketch the proof when $k = B$, or when Bob is the first to propose a menu. The details and the remaining part is in the Appendix. We are going to show that Alice can guarantee herself any payoff from any allocation in the Coasian menu $m^{1/2}$. In particular, her payoff cannot be smaller than (3.7).

Indeed, Lemmas 1 and 2 show that either Bob's offer includes all allocations in $m^{1/2}$, or Alice can pick any allocation in $m^{1/2} \backslash m^*_B$ and expect that Bob will concede and accept her offer. In any case, she is guaranteed to receive her best payoff from $m^{1/2}$.

[6] For $u_{-i,1} \geq u_{-i,1}$, the Nash allocations are presented in Figure 1.1, and the Nash payoffs are given by function

$$\mathcal{N}_A\left(u_i\right) = \max\left(\frac{1}{2u_{-i,}}u_1, \frac{1}{2}, \frac{1}{2}\left(1-u_1\right)\right). \tag{3.8}$$

On the other hand, Lemma 3 implies that either Alice's offer gives Bob a payoff that is arbitrarily close to $\frac{1}{2}$, or Bob has a counteroffer $m_B = m^{1/2}$ which leads to Alice's concession and payoff of $\frac{1}{2}$. In any case, Bob is guaranteed the payoff of $\frac{1}{2}$.

4. Discussion

4.1. **Relation to axiomatic solutions.** The solution obtained in Theorem 2 can be compared with axiomatic solutions to bargaining under incomplete information. Harsanyi and Selten (1972) consider the space $\mathcal{G}$ feasible (incentive compatible and individually rational) allocation mechanism g with an equilibrium with expected interim payoffs $g_j(u_j)$ to player j and type u_j.[7] They propose that such a solution should maximize

$$\log g_B(u_B) + \int \log g_A(u_A)\, d\rho_A(u_A) \tag{4.1}$$

among all $g \in \mathcal{G}$. They derive the formula as an extension of the Nash solution (which maximizes $\log g_A + \log g_B$ in the complete information analogue).

Myerson (1984) proposes an alternative bargaining solution that avoids some issues with handling state-dependent utility present in Harsanyi and Selten (1972) (these issues are irrelevant for the known-own payoff environment that we study in this paper). The solution is presented as a certain fixed point problem.

The Coasian menu $m^{1/2}$ is feasible. Nevertheless, it is easy to show that $m^{1/2}$ is neither part of the Harsanyi-Selten nor Myerson's solutions. Both solution concepts rely on axioms that compare bargaining problems across various environments, including

[7]In an allocation mechanism, players choose actions, and the mechanism determines an allocation. In the context of the bargaining environment of this paper, it is easy to show (see, for instance Peski (2019)) that, g is a feasible allocation mechanism if and only if there exists a menu $m \subseteq X$ such that $g_i(u_i) = \max_{x\in m} u_i(x)$ and $g_{-i}(u_{-i}) \leq \int \left(\max_{x\in \arg\max_{x'\in m} u_i(x')} u_{-i}(x)\right) d\mu(u_i)$, where μ are Bob's beliefs, and, alternatively, only if (a) $g_i(.)$ is convex, (b) if $D_{u_i} g_i$ is the set of supporting hyperplanes of $g_i(.)$ at u_i (i.e., l is affine, $l(u_i) = g_i(.)$ and $l(.) \leq g_i(.)$), then $l(u_i') \in [0,1]$ for each $l \in D_{u_i} g_i$ and each u_i, u_i', and (c) $g_{-i}(u_{-i}) \leq \int \left(1 - \max_{l \in D_{u_i} g_i} l(u_{-i})\right) d\mu(u_i)$.

interdependent value cases. Hence, because our solution is only defined for the known-own payoff heterogeneous pie case, it is difficult to point out a single axiom that is violated.

4.2. **Simple offers.** A natural question in our model is whether the ability to commit to menus is valuable, as compared with a more standard model, where players can only demand simple allocation. Consider a version of the model where players can only choose singleton menus, $M_i = \{\{x\} : x \in X\} \simeq X$.

Proposition 1. *Suppose that (a) Bob moves first, (b) Bob's preferences are known,* $\mathcal{U}_B = \{u_B\}$*, (c) Bob likes at least two parts of the pie,* $u_B^i, u_B^j > 0$ *for* $i \neq j$*, and (d) Alice's preferences have full support,* $\operatorname{supp} \pi = \mathcal{U}$*. Then,*

$$\sup E_B\left(u_B; X, \pi, B\right) < \frac{1}{2}. \tag{4.2}$$

Comparing to Theorem 2, Bob is strictly worse off and his equilibrium payoff is strictly below his worst possible complete information payoff. Thus, Bob values being able to use menus rather than singletons.

The result is a straightforward corollary to the proof of Theorem 2, which shows that Bob can ensure a payoff of $\frac{1}{2}$ only if he offers $m^{1/2}$. But Bob cannot offer such a menu if he is restricted to singletons and he likes at least two different parts of the pie. (If he liked only one part of the pie, then an offer where he keeps half of that part for himself and leaves the rest of the pie to Alice would be accepted.)

To see this intuition a bit more clearly, consider an arbitrary Bob's offer x. We can find a set $Y(x)$ of offers y such that, if offered by Alice, there is at least one type $u_A \in \mathcal{U}$ of Alice who is stronger than Bob:

$$\frac{u_A(y)}{u_A(x)} \geq \frac{u_B(x)}{u_B(y)}.$$

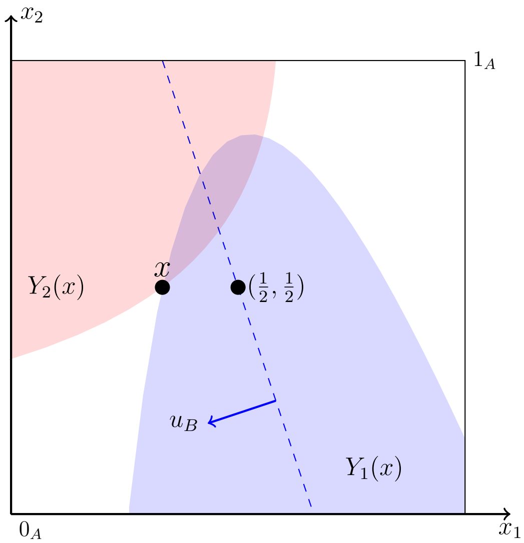

FIGURE 4.1. Allocations in sets $Y_1(x)$ (blue) and $Y_2(x)$ (red).

It is easy to see that, if there is a type that satisfies the above inequality, then there is also an extreme type that does so; we can replace the left-hand side by $\max_n \frac{y_n}{x_n}$. Thus,

$$Y(x) = \bigcup_n \{y : y_n u_B(y) \geq x_n u_B(x)\} = \bigcup_n Y_n(x),$$

where sets $Y_n(x)$ are defined as the sets of allocations y such that $y_n u_B(y) \geq x_n u_B(x)$, or Alice's offers for which the Alice preference type who only likes part n is stronger. Figure 4.1 illustrates on an example.

Any allocation from set $Y(x)$ makes Alice at least as strong as Bob in the following war-of-attrition. If λ and Δ are small, then any such choice y (or, more precisely, any choice in the interior of set $Y(x)$) will ensure that Alice wins and receives payoff $u_A(y)$, with a payoff $u_B(y)$ for Bob. Lemma 2 implies that Alice's menu choice and the continuation equilibrium in the war-of-attrition must result in a payoff that is not much smaller than $\max_{y \in Y(x)} u_A(y)$.

Calculations like in Lemma 1 show that, if $u_B(x) > \frac{1}{2}$, then, for each Alice's type, there is an allocation y in set $Y(x)$ that gives her more than her best payoff from menu $m^{1/2}$. If $u_B(x) \leq \frac{1}{2}$, then, depending on x, some types are better, or strictly better off by accepting allocation x, but there are others, who have a stronger counteroffer. In short, each Alice's type is able to ensure herself a payoff that is no worse, and sometimes strictly better than her best payoff from menu $m^{1/2}$. But because menu $m^{1/2}$ was efficient, Bob's payoff must be strictly smaller than $\frac{1}{2}$.

4.3. **Two-sided incomplete information.** With two-sided incomplete information about preferences, there is no natural notion of strength and no a priori sorting. In fact, it is possible to construct an example with two types for both players, where the war-of-attrition stage has two equilibria with different payoffs and different order of conceding types.

At the same time, when $N = 2$, the types are from a continuum, and when each player's menu is linear, some partial sorting can be restored. If menu m_{-i} is linear and there are only two dimensions, conceding player i picks one of the two extreme allocations from the menu. One shows that types that choose the same allocation can be sorted: the types that care strongly about one or the other part of the pie concede first; the type with an indifference curve aligned with the boundary of the menu concedes the last. The war-of-attrition stage has a unique equilibrium, and it has a simple characterization. Define the strength of a player as the winning/concession ratio under the restriction that, when conceding, the player must choose an allocation that belongs to the diagonal of the allocation space X. Because of linearity of preferences, the strength does not depend on the player's type. In the equilibrium, the weaker player concedes in early periods of the game with a probability arbitrarily close to 1.

Details on the two-type example and the above result can be found in the online Appendix.

Appendix A. Proof of Lemma 1

We start with an overview of the proof of the Lemma.

Subsection A.2 is devoted to notation and some preparatory results. Subsection A.2 introduces a convenient re-normalization of Alice's preference types so that the payoff of each type from allocation x^* is constant. Lemma 6 in Section A.3 lists basic properties of the war-of-attrition game, including sorting by strength of Alice's types and a characterization of concession rates. The next two results establish bounds in the late game. First, the key step of the proof is contained in Lemma 7, where we show that Bob's payoffs from Alice's concessions in the late game are not much higher than the payoff from x^*. Second, this observation is used in Lemma 8 to show that Alice's concedes at rates that are not slower than as if Bob was facing a type who always chooses x^* as her optimal choice. Lemma 9 deals with the early game and shows that either Bob or all Alice's types with a strength smaller than $1+\varepsilon$ concede very fast. Subsection A.7 uses these results to finish the proof of Lemma 1.

The assumption that $0 \ll x_A \ll 1$ implies that $d = \min_n (x_{A,n}, 1 - x_{A,n}) > 0$. Let $r_X = \frac{1}{4}d^2 > 0$. We divide the proof of the Lemma into two cases: either $B(\mathbf{0}_A, r_X) \subseteq m_B^*$, or not. The bulk of the proof is devoted to the former case; the latter case is dealt with in Subsection A.8.

A.1. **Constants.** The assumptions of the Lemma imply that $S_A^* - S_B^*, S_B^* - 1, u_B(x_A) > 0$. (Note that $u_B(x_A) \geq d$.) In the course of the proof, we define various constants, all of which depend only on the parameters of the model and

$$m_0 = \min(r_X, S_A^* - S_B^*, S_B^* - 1, d) > 0,$$

but not on any other properties of menu m_B, allocation x_A, or Bob's beliefs. From now on, we assume that Δ is sufficiently small so that $e^{\Delta} - 1 < \frac{1}{2}m_0$.

Let $\operatorname{diam}(\mathcal{U}_A)$ be the diameter of set $\mathcal{U}_A$. Because $\mathcal{U}_A$ has nonempty interior, there exists $r_U > 0$ and $u_0 \in \mathcal{U}_A$ such that $B(u_0, r_U) \subseteq \mathcal{U}_A$.

Let Λ be the Lebesgue measure on $\mathcal{U}_A$ (it exists, as $\mathcal{U}_A$ is a convex subset of $\mathbb{R}^N$ with a non-empty interior). Let $\Pi \in \Delta\mathcal{U}_A$ be Bob's beliefs. By the assumption, Bob's beliefs in the war-of-attrition stage are derived from updating on Alice's choice in the menu game. Let $\mu \in \Delta S$ would be Bob's posterior beliefs about Alice's signals. Then, the density of Bob's beliefs Π wrt. Λ is equal to $\pi(.) = \int \pi_S(.|s)\, d\mu(s)$. Assumption 1 implies that π is strictly positive on $\mathcal{U}_A$ and Lipschitz continuous.

For each player $j = A, B$, define the "limit" concession rates as if facing the opponents with strengths S^*_{-j}:

$$p^*_j = \frac{1 - \mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{S^*_{-j} - \mathrm{e}^{-\Delta}} > 0, \tag{A.1}$$

A.2. **Re-normalization.** It is convenient to re-normalize Alice's preference types. Recall that $\mathcal{U}_A \subseteq \mathcal{U} = \left\{u \in \mathbb{R}^N_+ : \sum u_n = 1\right\}$. Define function $\rho : \mathcal{U}_A \to \mathcal{U}' = \left\{u' \in \mathbb{R}^N_+ : \sum u'_n x^*_n = 1\right\}$ by $\rho(u) = \frac{1}{\sum u_n x^*_n} u$. The re-normalization allows us to represent the strength (see Section 3.2) of a type $u' = \rho(u)$ as

$$S_A(u) = \frac{u(x_A)}{\max_{x \in m^*_B} u(x)} = \frac{u(x_A)}{u(x^*)} \frac{1}{\max_{x \in m^*_B} \frac{1}{u(x^*)} u(x)} = \frac{1}{\kappa^*} \frac{1}{h(\rho(u)) + 1}, \tag{A.2}$$

where we use the fact that, by construction, $\kappa^* = \frac{u(x_A)}{u(x^*)}$, and where we define function $h : \mathcal{U}'_A \to \mathbb{R}_+$ by

$$h(u') = \max_{x \in m^*_B} u' \cdot (x - x^*). \tag{A.3}$$

The strength of a type is decreasing with $h(\rho(u))$, i.e., the value of function h computed for the re-normalized type.

Because of the properties of mapping ρ, the support $\mathcal{U}'_A = \rho(\mathcal{U}_A)$ of Bob's induced beliefs Π' is convex, and Π' has a strictly positive density π'. The next result summarizes some properties and bounds on re-normalized objects.

Lemma 4. *Let* $U(r_X) = \{u \in \mathcal{U}'_A : h(\rho(u)) \le r_X\}$*. There are constants* $0 < \pi_{\min}, \pi_{\max}, c_r, C_r, K_r, c_\Pi < \infty$ *such that*

(1) $U(r_X) \subseteq \{u \in \mathcal{U}'_A : u \cdot x^* \ge r_X\}$.

(2) For each $u \in U(r_X)$, we have $\pi_{\min} \leq c_r \pi(\rho(u)) \leq \pi'(\rho(u)) \leq C_r \pi(\rho(u)) \leq \pi_{\max}$.
(3) The induced (re-normalized) density π' is Lipschitz with a constant $K_r < \infty$ on $U(r_X)$.
(4) For each $\eta \leq r_X$, $\Pi\{u : h(\rho(u)) \leq \eta\} \geq c_\Pi \eta^{N-1}$.

Proof. For property 1, notice that $B(\mathbf{0}_A, r_X) \subseteq m_B^*$ implies that $u \cdot x_u \geq r_X$ for each preference type u. Hence, if $h(u) = u \cdot (x_u - x^*) \leq r_X$, then $u \cdot x^* \geq r_X$. Properties 2 and 3 follow from the fact that mapping ρ has an inverse with continuous derivative with non-disappearing Jacobian and from the fact that Bob's beliefs are derived under Assumption 1, hence, they have continuous density with respect to Lebesgue measure with uniform strictly positive lower and upper bounds.

For property 4, take any $u^* \in U$ such that x^* is an optimal choice of u^* in the set m_B^* (such a preference type exists, because x^* belongs to the boundary of set m_B^*). Then $h(\rho(u^*)) = 0$ and for each $u \in B(u^*, \eta)$ we have $h(\rho(u)) \leq \eta$. Consider a convexification of set $A = \text{con}(\{u^*\} \cup B(u_0, r_U))$. Simple calculations show that there is u and $r = \frac{r_U}{r_U + \text{diam}\mathcal{U}_A}\eta$ such that $B(u, r) \subseteq A \subseteq \mathcal{U}_A$. By the assumption,

$$\Pi\{u : h(\rho(u)) \leq \eta\} \geq \Pi(B(u, r)) \geq \pi_{\min}\Lambda(B(u, r)) = \text{const}\left(\frac{r_U}{r_U + \text{diam}\mathcal{U}_A}\eta\right)^{N-1}$$

for some constant. $\square$

From now on, we are going to work with the re-normalized space of preference types (including Bob type u_B). In the interest of saving on notational clutter, we drop the primes "$'$" from the re-normalized notation.

The reason for the re-normalization is that function h, that is used above to express the strength of Alice's types, has some nice properties. These properties are stated in the subsequent Lemma.

Lemma 5. *Properties of function h:*

(1) *Function h is continuous, convex, non-negative, and it attains a minimum of 0 at $u^* \in \mathcal{U}_A$.*

(2) *For almost all $u \in \mathcal{U}_A$, there exists unique x_u that solves $\max_{x \in m_B^*} u \cdot (x - x^*)$ and a derivative $Dh(u) = x_u - x^* \in \mathbb{R}^N$ such that for almost all $u \in \mathcal{U}_A$,*

$$u_B \cdot (x_u - x^*) = h(u) - Dh(u) \cdot (u - u_B).$$

(3) *For each $\eta > 0$,*

$$\Lambda \{u : h(u) = \eta\} = 0, \text{ and } \Lambda \{u : h(u) \leq \eta\} \leq 2^{N-1} \Lambda \left\{u : h(u) \leq \frac{1}{2}\eta\right\}.$$

Proof. Property 1 is immediate. Property 2 is a consequence of the Envelope Theorem. Property 3 is a standard property of convex functions. □

A.3. **War-of-attrition equilibrium.** For each player $i = A, B$, let $t_i^0 \in \{1, 2\}$ be the first decision period for player i. Fix player's strategies σ_i. Define

$$f^\sigma(U|t) = (1-\lambda) \int_U \sigma_A^T(t|u)\, \pi(du) \text{ for each } t \in T_A \text{ and measurable } U \subseteq \mathcal{U}_A,$$

$$f^\sigma(t) = (1-\lambda)\, \sigma_B^T(t) \text{ for each } t \in T_B.$$

For $t \in T_B$, $f^\sigma(t)$ is Bob's concession probability. Let $f^\sigma(t) = f^\sigma(\mathcal{U}_A|t)$ be the overall Alice's probability of concession in period $t \in T_A$. For each $i = A, B$ and $t \in T_i$, let

$$F^\sigma(t) = \lambda + \sum_{s \in T_i : s \geq t} f^\sigma(t), \text{ and } p_i^\sigma(t) = \frac{1}{F^\sigma(t)} f^\sigma(t),$$

be, respectively, the probability that player i has not conceded before period t and the concession rate in period t. Let $T_i^{*,\sigma} = \max\{t \in T_i : f^\sigma(t) > 0\}$ be the last period in which a strategic type of player i concedes.

For each $t \in T_A$, let

$$w^\sigma(t) = \int \sigma_A^M(u) \frac{1}{f^\sigma(t)} f^\sigma(du|t) \in X \text{ for each } t \in T_A,$$

be the expected allocation that Bob obtains in period t, conditional on Alice's concession in that period.

The superscripts σ in the above notation denotes the dependence on the strategy profile σ; the subscript i - on the player i. We drop the superscripts and/or the subscripts from the above notation whenever it does not lead to confusion.

The next result summarizes basic properties of the war-of-attrition game.

Lemma 6. *Suppose that σ is an equilibrium.*

(1) *Support of the concession behavior:*
   (a) *in each period before the last concession, a concession occurs with a positive probability: for each $t \leq \max_i T_i^{*,\sigma}$, $f^{\sigma}(t) > 0$.*
   (b) *players last concession are in the consecutive periods: $|T_i^{*,\sigma} - T_{-i}^{*,\sigma}| = 1$.*
   (c) *last concession happens in finite time: for each i, $T_i^{*,\sigma} < \infty$,*
   (d) *only stubborn types never concede: for each i, $F_i^{\sigma}(T_i^{*,\sigma} + 2) = \lambda$.*

(2) *Sorting: There exists T^0 and a strictly decreasing sequence $\eta_{t_A^0} > \eta_{t_A^0+2} \dots > \eta_{T^0} = 0$ such that in each equilibrium σ, $\sigma_A^T(u) = t$ if $\eta_{t-2} > h(u) > \eta_t$, and $\sigma_A^T(u) \geq T^0$ if $h(u) = 0$. Moreover, if $\pi(h^{-1}(0)) = \int_{h^{-1}(0)} d\pi(u) = 0$, then $T^0 = T_A^{*,\sigma}$.*

(3) *Concession rates:*

$$p_A^{\sigma}(t) = \frac{1 - \mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{u_B(w^{\sigma}(t))}{u_B(x_A)} - \mathrm{e}^{-\Delta}} \text{ for each } t \in T_A,$$

$$p_B^{\sigma}(t) = \frac{1 - \mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{1}{\kappa^*} \frac{1}{\eta_{t-1}+1} - \mathrm{e}^{-\Delta}} \text{ for each } t \in T_B.$$

Proof. Characterization of a best response behavior: We start with a preliminary step. For each $t \in T_B$, let $w^{\sigma}(t) = x_A$. Further, for each type $u \in \mathcal{U}_A$ of Alice, let

$$L_A(u) = \max_{x \in m_B} u(x) \text{ for each } u \in \mathcal{U}_A, L_B(u_B) = u_B(x_A), \text{ and}$$
$$S_i(u_i, t) = \frac{u_i(w^{\sigma}(t))}{L_i(u_i)} \text{ for each player } i \text{ and } t \in T_{-i}.$$

Here, $L_i(u)$ is the payoff received upon concession and S_i^{σ} is the strength ratio. Note that Bob's strength, but not Alice's, depends on time and the equilibrium behavior. The expected payoff of player i type u from conceding in period $t \in T_A$ given opponent strategies (σ) is equal to

$$U_i^{\sigma}(u, t) = \sum_{s: s<t, s \in T_{-i}} \mathrm{e}^{-s\Delta} f_{-i}^{\sigma}(s) (u_i(w^{\sigma}(s))) + \mathrm{e}^{-t\Delta} F_{-i}^{\sigma}(t+1) L_A(u_A).$$

For each $t \in T_i$, we have

$$\begin{aligned} &\mathrm{e}^{t\Delta}\left[U_i^{\sigma}(u_i, t+2) - U_i^{\sigma}(u_i, t)\right] \qquad \text{(A.4)} \\ =&\mathrm{e}^{-\Delta} f_{-i}^{\sigma}(t+1)(u_i(w_i^{\sigma}(t+1))) + \left[\mathrm{e}^{-2\Delta}\left(F_{-i}^{\sigma}(t+1) - f_{-i}^{\sigma}(t+1)\right) - F_{-i}^{\sigma}(t+1)\right] L_i(u_i) \\ =&F_{-i}^{\sigma}(t+1)\left[\mathrm{e}^{-\Delta} p_{-i}^{\sigma}(t+1)(u_i(w_i^{\sigma}(t+1))) - \left(\mathrm{e}^{-2\Delta} p_{-i}^{\sigma}(t+1) + 1 - \mathrm{e}^{-2\Delta}\right) L_i(u_i)\right]. \end{aligned}$$

We have the following corollary to the above formula and the definition of strength: For each type u_i of player i, each $t \in T_i$, $U_i^{\sigma}(u_i, t+2) \geq (\leq) U_i^{\sigma}(u_i, t)$ if and only if

$$p_{-i}^{\sigma}(t+1) \geq (\leq) \frac{1 - \mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{S_i^{\sigma}(u, t+1) - \mathrm{e}^{-\Delta}}. \qquad \text{(A.5)}$$

Part 1. The above implies that, if$f^{\sigma}(t) = 0$ for some $t \in T_{-i}$, then it is a strictly better response for almost any type u of player i to concede in period $t-1$ rather than to wait to period $t+1$. It follows that $f^{\sigma}(t+1) = 0$. An induction implies that $f^{\sigma}(t') = 0$ for each $t' > t$. This implies claims (1a) and(1b).

To see claim (1c), we take $\mathcal{U}_B = \{u_B\}$ and let $L_i^{\min} = \inf_{u \in \mathcal{U}_i} L_i(u)$. The serious offer assumption ensures that $L_i^{\min} > 0$. Because $f_i^{\sigma}(t) > 0$ for each $t \leq T_i^{*,\sigma}$, it must be that for each $t \in T_i$, if $t < T_i^{*,\sigma}$, there is a type $u \in \mathcal{U}_{-i}$ of player $-i$ such that $U_{-i}^{\sigma}(u_{-i}, t-1) \leq U_{-i}^{\sigma}(u_{-i}, t+1)$. Inequalities (A.5) implies that for each $t < T_i^{*,\sigma}$,

$$p_i^{\sigma}(t) \geq \left(1 - \mathrm{e}^{-\Delta}\right) \frac{1 + \mathrm{e}^{-\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\max_{u \in A_{-i}} S_{-i}^{\sigma}(u_{-i}, t) - \mathrm{e}^{-\Delta}} \geq \left(1 - \mathrm{e}^{-\Delta}\right) L_{-i}^{\min} > 0.$$

Hence, for each $t \leq T_i^{*,\sigma}$

$$\begin{aligned} F_i^{\sigma}(t) &= \left(1 - p_i^{\sigma}(t-2)\right) F_i^{\sigma}(t-2) \leq \left(1 - \left(1 - \mathrm{e}^{-\Delta}\right) L_{-i}^{\min}\right) F_i^{\sigma}(t-2) \\ &\leq \left(1 - \left(1 - \mathrm{e}^{-\Delta}\right) L_{-i}^{\min}\right)^{\left(t - t_i^0\right)/2}. \end{aligned}$$

Because $F_i^{\sigma}(t) \geq \lambda$, it must be that $T_i^{*,\sigma} - t_i^0 \leq \frac{\log \lambda}{\log\left(1 - \left(1 - \mathrm{e}^{-\Delta}\right) L_{-i}^{\min}\right)}$.

Claim (1d) follows from the fact that, because offers are serious, if there are no further concession from the other player, each player prefers to concede and receive a strictly positive payoff.

Part 2. We show first that Bob's concession rates must be decreasing. Take any type u for whom it is the best response to concede in period $t+1$. Hence, it must be that $U_A(u, t+1) \geq \max\left(U_A(u, t-1), U_A(u, t+3)\right)$. By the preliminary step, and due to the re-normalization and equality (A.2), we have

$$p_B^{\sigma}(t) \geq \frac{1 - \mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{1}{\kappa^*} \frac{1}{h(u)+1} - \mathrm{e}^{-\Delta}} \geq p_B^{\sigma}(t+2).$$

By the first part, in each period $t \in T_A$, a strictly positive mass of Alice's types must concede. By Lemma 5, either $p_B^{\sigma}(t) = p_B^{\sigma}(t+2) = \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{1}{\kappa^*} - \mathrm{e}^{-\Delta}}$,or $p_B(t) > \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{1}{\kappa^*} - \mathrm{e}^{-\Delta}}$ and $p_B(t) > p_B(t+2)$. In the latter case, a 0 mass of Alice's types is indifferent between conceding in period t rather than in the preceding or subsequent periods. The rest of the claim follows from the characterization of the best response behavior.

Part 3 follows from part 1. $\square$

The mass of conceding Alice's types in period $t \in T_A$ and Alice's concession rate are equal to, respectively,

$$f^{\sigma}(t) = \Pi\left(h^{-1}[\eta_t, \eta_{t-2}]\right) \text{ and } p_A^{\sigma}(t) = \frac{\Pi\left(h^{-1}[\eta_t, \eta_{t-2}]\right)}{\Pi\left(h^{-1}[0, n_{t-2}]\right)}.$$

We let $\eta_t = 0$ for each $t > T_i^0$ and $t \in T_i$. The above proof implies that types $u \in h^{-1}(\eta_t)$ are indifferent between stopping in period t and $t+2$.

A.4. **Late game: payoff bounds.** The next result is a key step of this part of the Appendix. It says that, in the late game, the average winning payoff for Bob is not significantly larger than $u_B(x^*)$.

Let $T_A^{\eta} = \min\{t \in T_A : \eta_t \leq \eta\}$ be the period after which the value of function h for all remaining Alice's types is smaller than η. Let $|x|_+ = \max(x, 0)$.

Lemma 7. *There exists a constant $C_0 < \infty$ such that for each $0 < \eta < r_X$,*

$$\sum_{t \in T_A : t > T_A^{\eta}} \left|u_B\left(w_B^{\sigma}(t)\right) - u_B(x^*)\right|_+ f^{\sigma}(t) \leq C_0 \eta \left(F^{\sigma}\left(T_A^{\eta}\right) - \lambda\right).$$

Proof. For each $\eta \geq 0$, let $U(\eta) = \{u \in \mathcal{U}_A : 0 \leq h(u) \leq \eta\}$. By Lemma 5, for each $t \in T_A$, we have

$$\begin{aligned}
&\left(u_B\left(w_B^{\sigma}(t)\right) - u_B(x^*)\right) f^{\sigma}(t) \\
= &\int\limits_{u \in \mathcal{U}_A : \eta_t \leq h(u) < \eta_{t-2}} u_B \cdot (x^* - x_u)\, \pi(u)\, du \\
= &\int\limits_{u \in \mathcal{U}_A : U(\eta_{t-2}) \setminus U(\eta_t)} \left[Dh(u) \cdot (u - u_B) - h(u)\right] \pi(u)\, du. \qquad \text{(A.6)}
\end{aligned}$$

The first equality comes from the fact that $u_B(x) = 1 - u_B \cdot x$.

Polar coordinates. The goal of the main part of the proof is to provide bounds on the the integral in the last line of (A.6). For this purpose, we switch to polar coordinates with center at u_B. Let $S_{N-2} = \left\{\nu \in \mathbb{R}^N : \nu \cdot \nu = 1, \nu(x^*) = 0\right\}$ be a $N-2$-dimensional sphere. Let $|S_{N-2}|$be the Lebesgue measure of the sphere S_{N-2}. For each vector $\nu \in S_{N-2}$, each $\eta \geq 0$, let $\alpha^{\min}(\nu,\eta)$ and $\alpha^{\max}(\nu,\eta)$ be defined as

$$\alpha^{\min}(\nu,\eta) = \inf\{a \geq 0 : u_B + a\nu \in U(\eta)\},$$
$$\alpha^{\max}\ (\nu,\eta) = \sup\{a \geq 0 : u_B + a\nu \in U(\eta)\}.$$

In other words, these quantities are the distance coordinates of the first and the last preference type along vector v from set $U(\eta)$. If the set of such αs is empty, we take $\alpha^{\min}(\nu,\eta) = \alpha^{\max}(\nu,\eta) = 0$. Then, due to the convexity of function h, up to 0-measure sets, set $U(\eta_t)$ is equal to the union of the line segments $\{u_B + a\nu : \alpha^{\min}(\nu,\eta) \leq a \leq \alpha^{\max}(\nu,\eta)\}$ for all $\nu \in S_{N-2}$. Also, let

$$\alpha(\nu,\eta) = \alpha^{\max}(\nu,\eta) - \alpha^{\min}(\nu,\eta) = |\{a : u_B + a\nu \in U(\eta)\}|.$$

Notice that

$$\begin{aligned}\alpha(\nu,\eta_{t-2}) - \alpha(\nu,\eta_t) &= \alpha^{\max}(\nu,\eta_{t-2}) - \alpha^{\max}(\nu,\eta_t) + \alpha^{\min}(\nu,\eta_t) - \alpha^{\min}(\nu,\eta_{t-2}) \\ &= |\{a : u_B + a\nu \in U(\eta_{t-2}) \backslash U(\eta_t)\}|.\end{aligned}$$

For each $v \in S_{N-2}$, each $a \geq 0$, define

$$\begin{aligned} h_v(a) &= h(u_B + av), \qquad \text{(A.7)} \\ g_v(a) &= \pi(u_B + a\nu)\, a^{N-2}, \\ g_\nu^*(a) &= a g_\nu(a). \end{aligned}$$

Then, g is Lipschitz with constant $K_g = (\operatorname{diam}(\mathcal{U}_A))^{N-3}(K_r(\operatorname{diam}(\mathcal{U}_A)) + (N-2)\pi_{\max})$, where K_r is the Lipschitz constant from Lemma 4. For future reference, note that for each $s = \min, \max$,

$$\begin{aligned} h_\nu\left(\alpha^s\left(\nu,\eta\right)\right) &\leq \eta, \\ \left|h_\nu\left(\alpha^s\left(\nu,\eta_{t-2}\right)\right) - h_\nu\left(\alpha^s\left(\nu,\eta_t\right)\right)\right| &\leq \eta_{t-2} - \eta_t, \\ \left|a g_v'\left(a\right) + 2g\left(a\right)\right| &\leq D a^{N-2}. \end{aligned} \tag{A.8}$$

for $D = (N-2)\pi_{\max} + K_r$.

Computing (A.6). The integral in the last line of (A.6) is equal to

$$= \frac{1}{|S_{N-2}|}\int_{S_{N-2}}\left(\int_{\alpha^{\min}(\nu,\eta_{t-2})}^{\alpha^{\min}(\nu,\eta_t)}\left(ah_v'\left(a\right) - h_v\left(a\right)\right)g_v\left(a\right)da + \int_{\alpha^{\max}(\nu,\eta_t)}^{\alpha^{\max}(\nu,\eta_{t-2})}\left(ah_v'\left(a\right) - h_v\left(a\right)\right)g_v\left(a\right)da\right)d\nu.$$

By the integration by parts, the above is equal to

$$\begin{aligned} = &-\frac{1}{|S_{N-2}|}\int_{S_{N-2}}\left(\int_{\alpha^{\min}(\nu,\eta_{t-2})}^{\alpha^{\min}(\nu,\eta_t)}h_v\left(a\right)\left(ag_v'\left(a\right) + 2g_v\left(a\right)\right)da + \int_{\alpha^{\max}(\nu,\eta_t)}^{\alpha^{\max}(\nu,\eta_{t-2})}h_v\left(a\right)\left(ag_v'\left(a\right) + 2g_v\left(a\right)\right)da\right)d\nu \\ &+ \frac{1}{|S_{N-2}|}\int_{S_{N-2}}\left(h_v\left(a\right)g_v^*\left(a\right)\Big|_{\alpha^{\min}(v,\eta_{t-2})}^{\alpha^{\min}(v,\eta_t)} + h_v\left(a\right)g_v^*\left(a\right)\Big|_{\alpha^{\max}(v,\eta_t)}^{\alpha^{\max}(v,\eta_{t-2})}\right)d\nu. \end{aligned} \tag{A.9}$$

Bounding the terms. Due to inequalities (A.8), the first term is not larger than

$$\frac{1}{|S_{N-2}|}D\eta_t\int_{S_{N-2}}\left(\int_{\alpha^{\min}(\nu,\eta_{t-2})}^{\alpha^{\min}(\nu,\eta_t)}a^{N-2}da + \int_{\alpha^{\max}(\nu,\eta_t)}^{\alpha^{\max}(\nu,\eta_{t-2})}a^{N-2}da\right)d\nu \leq D\frac{1}{\pi_{\min}}\eta_t\left(F\left(t-2\right) - F\left(t\right)\right).$$

In order to find an upper bound on the second term of (A.9), we introduce the following notation: for any function f, each t, let

$$\begin{aligned} \Delta_t^s f &= f\left(\alpha^s\left(\nu,\eta_{t-2}\right)\right) - f\left(\alpha^s\left(\nu,\eta_t\right)\right), \text{ for } s = \min, \max, \\ \Delta^t f &= f\left(\alpha^{\max}\left(\nu,\eta_t\right)\right) - f\left(\alpha^{\min}\left(\nu,\eta_t\right)\right), \text{ and} \\ \Delta^t\left(\Delta f\right) &= \Delta_t^{\max} f - \Delta_t^{\min} f. \end{aligned}$$

The expression in the brackets of the second term of (A.9) is equal to

$$
\begin{aligned}
&\Delta^t \left(\Delta \left(h_v g_v^*\right)\right) = \Delta_t^{\max}\left(h_v g_v^*\right) - \Delta_t^{\min}\left(h_v g_v^*\right) \\
=&\Delta_t^{\max}\left(h_v\right) g_v^*\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right) - h_v\left(\alpha^{\max}\left(\nu,\eta_t\right)\right)\Delta_t^{\max}\left(g_v^*\right) - \Delta_t^{\min}\left(h_v\right) g_v^*\left(\alpha^{\min}\left(\nu,\eta_{t-2}\right)\right) + h_v\left(\alpha^{\min}\left(\nu,\eta_t\right)\right)\Delta_t^{\min}\left(g_v^*\right) \\
=&\Delta_t^{\max}\left(h_v\right) g_v^*\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right) - \Delta_t^{\min}\left(h_v\right) g_v^*\left(\alpha^{\min}\left(\nu,\eta_{t-2}\right)\right) - \left[h_v\left(\alpha^{\max}\left(\nu,\eta_t\right)\right)\Delta_t^{\max}\left(g_v^*\right) - h_v\left(\alpha^{\min}\left(\nu,\eta_t\right)\right)\Delta_t^{\min}\left(g_v^*\right)\right] \\
=&\left(\Delta_t^{\max} h_v - \Delta_t^{\min} h_v\right) g_v^*\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right) + \left(\Delta_t^{\min} h_v\right)\left(\Delta^{t-2} g_v^*\right) - \left[\left(\Delta^t h_v\right)\left(\Delta_t^{\max}\left(g_v^*\right)\right) + h_v\left(\alpha^{\min}\left(\nu,\eta_t\right)\right)\left(\Delta_t^{\max} g_v^* - \Delta_t^{\min} g_v^*\right)\right] \\
=&\Delta^t\left(\Delta h_v\right) g_v^*\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right) + \left(\Delta_t^{\min} h_v\right)\left(\Delta^{t-2} g_v^*\right) - \left(\Delta^t h_v\right)\left(\Delta_t^{\max}\left(g_v^*\right)\right) - h_v\left(\alpha^{\min}\left(\nu,\eta_t\right)\right)\Delta^t\left(\Delta g_v^*\right),
\end{aligned}
\tag{A.10}
$$

Notice that function g_v^* is Lipschitz with a Lipschitz constant $K(\nu) = (N-1)\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right)^{N-2}\pi_{\max}$ on the interval $\alpha \in \left[\alpha^{\min}\left(\nu,\eta_{t-2}\right), \alpha^{\max}\left(\nu,\eta_{t-2}\right)\right]$. This, and the properties of function h_v imply that

$$
\begin{aligned}
\left|\Delta_t^{\min} h_v\right|, \left|\Delta_t^{\min} h_v\right| &\leq \eta_{t-2} - \eta_t, \\
\left|\Delta^t h_v\right|, \left|h_v\left(\alpha^{\min}\left(\nu,\eta_t\right)\right)\right| &\leq \eta_t, \\
\left|\Delta^{t-2} g_v^*\right| &\leq K(\nu)\,\alpha\left(\nu,\eta_{t-2}\right), \\
\left|\Delta_t^{\max}\left(g_v^*\right)\right| &\leq K(\nu)\left|\alpha^{\max}\left(\nu,\eta_{t-2}\right) - \alpha^{\max}\left(\nu,\eta_t\right)\right| \leq K(\nu)\left(\alpha\left(\nu,\eta_{t-2}\right) - \alpha\left(\nu,\eta_t\right)\right), \\
\Delta^t\left(\Delta g_v^*\right) &\leq K(\nu)\left(\alpha\left(\nu,\eta_{t-2}\right) - \alpha\left(\nu,\eta_t\right)\right).
\end{aligned}
$$

Out of all terms of the last line of (A.10), we still need to bound the first term, $\Delta^t\left(\Delta h_v\right) g_v^*\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right)$. We consider the following cases:

- if $h_v(u_B) > \eta_{t-2}$, then, because $u_B \in \mathcal{U}_A$ (this is the only place in the proof where this assumption is used), $\alpha^{\min}\left(\nu,\eta_{t-2}\right) > 0$, and $h_v\left(\alpha^{\min}\left(\nu,\eta_{t-2}\right)\right) = \eta_{t-2}$:
  - if $\alpha^{\min}\left(\nu,\eta_t\right) = 0$, then $\alpha^{\max}\left(\nu,\eta_t\right) = 0$, and $h\left(\alpha^{\max}\left(\nu,\eta_t\right)\right) = h\left(\alpha^{\min}\left(\nu,\eta_t\right)\right)$,
  - if $\alpha^{\min}\left(\nu,\eta_t\right) > 0$, then $h\left(\alpha^{\min}\left(\nu,\eta_t\right)\right) = \eta_t$. If $h\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right) \geq \eta_t$, then $h\left(\alpha^{\max}\left(\nu,\eta_t\right)\right) = \eta_t$, in which case $h\left(\alpha^{\max}\left(\nu,\eta_t\right)\right) = h\left(\alpha^{\min}\left(\nu,\eta_t\right)\right)$,
  - otherwise, i f$\alpha^{\min}\left(\nu,\eta_t\right) > 0$, and $h\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right) < \eta_t$, then $h\left(\alpha^{\max}\left(\nu,\eta_{t-2}\right)\right) = h\left(\alpha^{\max}\left(\nu,\eta_t\right)\right)$.

In each of the above sub-cases, we have

$$\Delta^t(\Delta h_v) = h_v(\alpha^{\max}(\nu, \eta_{t-2})) - h_v(\alpha^{\max}(\nu, \eta_t)) + h_v\left(\alpha^{\min}(\nu, \eta_t)\right) - h_v\left(\alpha^{\min}(\nu, \eta_{t-2})\right) \leq 0,$$

which implies that $\Delta^t(\Delta h_v) g_v^*(\alpha^{\max}(\nu, \eta_{t-2})) \leq 0$,

- if $h_v(u_B) \leq \eta_{t-2}$, then $\alpha^{\min}(\nu, \eta_{t-2}) = 0$, and $\alpha^{\max}(\nu, \eta_{t-2}) = \alpha(\nu, \eta_{t-2})$. In such a case,

$$\begin{aligned}
&\Delta^t(\Delta h_v) g_v^*(\alpha^{\max}(\nu, \eta_{t-2})) \\
\leq& 2(\eta_{t-2} - \eta_t) g_v^*(\alpha^{\max}(\nu, \eta_{t-2})) \leq 2(\eta_{t-2} - \eta_t)(\alpha^{\max}(\nu, \eta_{t-2}))^{N-1} \pi_{\max} \\
=& 2(\eta_{t-2} - \eta_t)(\alpha(\nu, \eta_{t-2}))^{N-1} \pi_{\max} \leq K(\nu) \alpha(\nu, \eta_{t-2})(\eta_{t-2} - \eta_t).
\end{aligned}$$

Ultimately, collecting all the bounds, we obtain that the expression in the brackets of the integral of the second term of (A.9) is not larger than

$$2K(\nu) \alpha(\nu, \eta_{t-2})(\eta_{t-2} - \eta_t) + 2K(\nu) \eta_t (\alpha(\nu, \eta_{t-2}) - \alpha(\nu, \eta_t)).$$

Note that the above is non-negative.

Finishing the proof. Putting all the bounds together, we obtain

$$\begin{aligned}
&\sum_{t \in T_A : t > T_A^\eta} |u_B(w_B^\sigma(t)) - u_B(x^*)|_+ f^\sigma(t) \\
\leq& D \frac{1}{\pi_{\min}} \eta (F(T_A^\eta) - \lambda) + 4(N-1) \pi_{\max} \eta \frac{1}{|S_{N-2}|} \int_{S_{N-2}} \alpha(\nu, \eta)(\alpha^{\max}(\nu, \eta_{t-2}))^{N-2} d\nu.
\end{aligned}$$

The last term can be bounded by

$$\frac{1}{|S_{N-2}|} \int_{S_{N-2}} \alpha(\nu, \eta)(\alpha^{\max}(\nu, \eta_{t-2}))^{N-2} d\nu \leq \frac{1}{|S_{N-2}|} \int_{S_{N-2}} \left(\int_{\alpha^{\min}(\nu, \eta)}^{\alpha^{\max}(\nu, \eta)} \alpha^{N-2} d\alpha \right) d\nu,$$

where we use the fact that

$$\int_{\alpha^{\min}}^{\alpha^{\max}} a^{N-2}da = \frac{1}{N-1}\left(\left(\alpha^{\max}\right)^{N-1} - \left(\alpha^{\max}\right)^{N-1}\right)$$

$$= \left(\alpha^{\max} - \alpha^{\min}\right)\frac{1}{N-1}\left(\sum_{k=0}^{N-2}\left(\alpha^{\min}\right)^{k}\left(\alpha^{\max}\right)^{N-2-k}\right) \leq \left(\alpha^{\max} - \alpha^{\min}\right)\left(\alpha^{\max}\right)^{N-2}.$$

Because the density is not smaller than $\pi_{\min}$ and $\alpha^{\max} \leq \operatorname{diam}\left(\mathcal{U}_A\right)$, the above is not larger than

$$\leq \frac{1}{|S_{N-2}|}\frac{1}{\pi_{\min}}\int\limits_{S_{N-2}}\left(\int\limits_{\alpha^{\min}(\nu,\eta)}^{\alpha^{\max}(\nu,\eta)} \alpha^{N-2}\pi\left(u_B + a\nu\right)da\right)d\nu = \frac{1}{\pi_{\min}}\int\limits_{U(\eta)}\pi\left(u\right)d\nu$$

$$= \frac{1}{\pi_{\min}}\left(F^{\sigma}\left(T_A^{\eta}\right) - \lambda\right).$$

Hence, using the definition of the constant D, we get

$$\sum_{t\in T_A:t>T_A^{\eta}}\left|u_B\left(w_B^{\sigma}\left(t\right)\right) - u_B\left(x^*\right)\right|_+ f^{\sigma}\left(t\right) \leq \frac{1}{\pi_{\min}}\left(K_{\pi}\operatorname{diam}\left(\mathcal{U}_A\right) + 5\left(N-1\right)\pi_{\max}\right)\eta\left(F\left(T_A^{\eta}\right) - \lambda\right)$$

□

A.5. **Late game: bounds on concession rates.** The next result shows that, in the late game, the average concession rate is not significantly slower than the limit rate p_A^* (see definition (A.1) above).

Lemma 8. *There exists a constant $C_1 < \infty$ such that for each $0 < \eta < r_X$,*

$$\prod_{t\in T_A:T_A^{\eta}<t\leq T_A^{*,\sigma}}\left(1 - p_i^{\sigma}\left(t\right)\right) \leq \mathrm{e}^{C_1\eta}\prod_{t\in T_A:T_A^{\eta}<t\leq T_A^{*,\sigma}}\left(1 - p_A^*\right).$$

Proof. By Lemma 6, if $t \in T_i$ and $t < T_B^{*,\sigma}$, Bob must be indifferent between conceding in periods $t-1$ and $t+1$. Moreover,

$$\begin{aligned} p_A^\sigma(t) =& \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{u_B\left(w_B^\sigma(t)\right)}{u_B(x_A)} - \mathrm{e}^{-\Delta}} \\ =& \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{S_B^* - \mathrm{e}^{-\Delta} + \frac{u_B\left(w_B^\sigma(t)\right) - u_B(x^*)}{u_B(x_A)}}. \end{aligned}$$

Let $D = u_B(x^*) - \mathrm{e}^{\Delta} u_B(x_A) = u_B(x_A)\left(S_B^* - \mathrm{e}^{\Delta}\right) \geq \frac{1}{2} m_0^2 > 0$. Using the definition of p_i^* from (A.1), we compute that

$$\begin{aligned} \frac{1-p_A^\sigma(t)}{1-p_A^*} &= 1 + \frac{p_A^* - p_A^\sigma(t)}{1-p_A^*} \\ &\leq 1 + \frac{1}{D} p_A^\sigma(t)\left(u_B\left(w_B^\sigma(t)\right) - u_B(x^*)\right). \end{aligned}$$

Thus,

$$\begin{aligned} \frac{\prod_{t\in T_A : T_A^\eta < t \leq T_A^{*,\sigma}} \left(1 - p_A^\sigma(t)\right)}{\left(1-p_A^*\right)^{\frac{1}{2}\left(T_A^{*,\sigma} - T_A^\eta\right)}} &\leq \prod_{t\in T_A : T_A^\eta < t \leq T_A^{*,\sigma}} \frac{1-p_A^\sigma(t)}{1-p_A^*} \\ &\leq \prod_{t\in T_A : T_A^\eta < t \leq T_A^{*,\sigma}} \left(1 + \frac{1}{D} p_A^\sigma(t) \left|u_B\left(w_B^\sigma(t)\right) - u_B(x^*)\right|_+\right) \\ &\leq \exp\left(\frac{1}{D} \sum_{t\in T_A : T_A^\eta < t \leq T_A^{*,\sigma}} p_A^\sigma(t) \left|u_B\left(w_B^\sigma(t)\right) - u_B(x^*)\right|_+\right). \end{aligned}$$

Finally, we show that if η is sufficiently small, then the summation expression in the brackets becomes arbitrarily small. Inductively, define a sequence: $t_0 = T_i^\eta$, and for each l,

$$t_t = \max\left(t > t_{l-1} : t \in T_i, \eta_t > \frac{1}{2}\eta_{t_{l-1}}\right).$$

The definition implies that $\eta_{t_{l+1}} < \eta_{t_l}$, and that $\eta_{t_{l+2}} \leq \frac{1}{2}\eta_{t_l}$. Hence, $\sum_{l\geq 1} \eta_{t_{l-1}} \leq 2\eta_{t_0} \leq 2\eta$.

By Lemma 5, the Lebesgue mass of set $\left\{u : h(u) \leq \frac{1}{2}\eta_{t_{l-1}}\right\}$ is at least $2^{-(N-1)}$ of the Lebesgue mass of set $\left\{u : h(u) \leq \eta_{t_{l-1}}\right\}$. It follows that

$$\frac{F^{\sigma}(t_{l-1})}{F^{\sigma}(t_l)} \geq \frac{\pi_{\min}}{\pi_{\max}} \frac{\Lambda\left\{u : h(u) \leq \frac{1}{2}\eta_{t_{l-1}}\right\}}{\Lambda\left\{u : h(u) \leq \eta_{t_{l-1}}\right\}} \geq \frac{\pi_{\min}}{2^{N-1}\pi_{\max}},$$

where Λ is the Lebesgue measure. Using the above bound and Lemma 7, we obtain for each $l \geq 1$,

$$\begin{aligned}
&\sum_{t \in T_A : t_{l-1} < t \leq t_l} p_A^{\sigma}(t) \left|u_B(w_B^{\sigma}(t)) - u_B(x^*)\right|_+ \\
\leq &\frac{1}{F(t_{l-1})} \sum_{t \in T_A : t_{l-1} < t \leq t_l} \frac{F^{\sigma}(t_{l-1})}{F^{\sigma}(t)} \left|u_B(w_B^{\sigma}(t)) - u_B(x^*)\right|_+ f^{\sigma}(t) \\
\leq &\frac{2^{N-1}\pi_{\max}}{\pi_{\min}} \frac{1}{F(t_l)} \sum_{t \in T_A : t_{l-1} < t \leq t_l} \left|u_B(w_B^{\sigma}(t)) - u_B(x^*)\right|_+ f^{\sigma}(t) \\
\leq &\frac{2^{N-1}\pi_{\max}}{\pi_{\min}} C_0 \eta_{t_{l-1}}.
\end{aligned}$$

Hence,

$$\sum_{t \in T_A : T_A^{\eta} < t \leq T_A^{*,\sigma}} p_A^{\sigma}(t) \left|u_B(w_B^{\sigma}(t)) - u_B(x^*)\right|_+ \leq \frac{2^{N-1}\pi_{\max}}{\pi_{\min}} C_0 \sum_{l \geq 1} \eta_{t_{l-1}} \leq \frac{2^{N}\pi_{\max}}{\pi_{\min}} C_0 \eta.$$

□

A.6. **Early game.** The next result discusses the concession behavior when Alice may still have very weak (i.e., with strength not much higher than 1) types. It says that, essentially, either Bob concedes early with a probability arbitrarily close to 1, or all the weak Alice's types concede early, where "early" here means with a small amount of discounting.

Lemma 9. *For each $\delta \in \left(0, \frac{1}{2}\right)$, there exists $\Delta^* > 0$ such that if $\Delta \leq \Delta^*$, then there exists T_0 such that $\mathrm{e}^{-\Delta T_0} \geq 1 - 2\delta$ and for each equilibrium σ, either (a) $F_B^{\sigma}(T_0) \leq \delta$, or (b) $\sigma_A^{T_0}(u_A) \leq T_0$ for all $u_A \in \mathcal{U}_A$ st. $\sup_{t \in T_B} S_A^{\sigma}(u_A, t) \leq 1 + \frac{1}{2}\delta^2$.*

Proof. Let $k^* = \lceil -\log_2 \delta \rceil \leq -\log_2 \delta + 1$. Notice that $\delta (1 - \log_2 \delta) \leq 1$ for $\delta < \frac{1}{2}$, which implies that

$$1 - \delta^2 \geq 1 - \frac{\delta}{1 - \log_2 \delta} \geq (1 - \delta)^{\frac{1}{k^*}}.$$

Fix $\Delta^* > 0$ so that $2\Delta^* (1 - \log_2 \delta) \leq \log \frac{1-\delta}{1-2\delta}$. For each $\Delta \leq \Delta^*$, let n_Δ be the smallest even integer such that $\mathrm{e}^{-\Delta n_\Delta} \leq 1 - \delta^2$. Then, $\mathrm{e}^{-\Delta n_\Delta} \geq (1 - \delta^2) \mathrm{e}^{-2\Delta}$. Take $T_0 = k^* n_\Delta$. Then,

$$\mathrm{e}^{-T_0 \Delta} \geq \left(1 - \delta^2\right)^{k^*} \mathrm{e}^{-2\Delta k^*} \geq (1 - \delta) \mathrm{e}^{-2\Delta (1 - \log_2 \delta)} \geq 1 - 2\delta.$$

Suppose that there is a type $u_A \in \mathcal{U}_A$ such that $S_A^\sigma (u_A, t) \leq 1 + \frac{1}{2}\delta^2$ for each $t \in T_B$, and suppose that $T \geq T_0$ is a best response stopping time for such type u_A. Then, it must be that for each $t \in T_A$,$t < T$, type u_A prefers to continue waiting till period T rather than conceding in period t:

$$F_A(t) L_A(u_A) \leq \sum_{t<s<T: s \in T_B} f_B(s) \mathrm{e}^{-(s-t)\Delta} [S_A(u_A, s) L_A(u_A)] + F_A(T) \mathrm{e}^{-(T-t)\Delta} L_A(u_A).$$

After some algebra, and taking into account that $S_A(u_A, s) \leq 1 + \frac{1}{2}\delta^2$, we get

$$0 \leq \sum_{s>t: s \in T_B} f_B(s) \left(\mathrm{e}^{-(s-t)\Delta} \left(1 + \frac{1}{2}\delta^2\right) - 1\right).$$

Due to the choice of n_Δ, for each $t \leq T - n_\Delta$, the above is not larger than

$$\begin{aligned} &\leq \sum_{t<s<t+n_\Delta: s \in T_B} f_B(s) \frac{1}{2}\delta^2 + \sum_{s>t+n_\Delta: s \in T_B} f_B(s) \left(\mathrm{e}^{-n_\Delta \Delta} \left(1 + \frac{1}{2}\delta^2\right) - 1\right) \\ &\leq \varepsilon \left(\sum_{t<s<t+n_\Delta: s \in T_B} f_B(s) - \sum_{s>t+n_\Delta: s \in T_B} f_B(s)\right). \end{aligned}$$

In the second inequality, we used the fact that $\mathrm{e}^{-\Delta n_\Delta} \left(1 + \frac{1}{2}\delta^2\right) \leq (1 - \delta^2) \left(1 + \frac{1}{2}\delta^2\right) \leq -\frac{1}{2}\delta^2 - \frac{1}{2}\delta^4 \leq -\frac{1}{2}\delta^2$. Thus, for any such t,

$$\sum_{t<s<t+n_\Delta: s \in T_B} f_B(s) \geq \frac{1}{2} \left(\sum_{t<s<t+n_\Delta: s \in T_B} f_B(s) + \sum_{s>t+n_\Delta: s \in T_B} f_B(s)\right) = \frac{1}{2} \sum_{t<s<T: s \in T_B} f_B(s).$$

It follows that

$$1 - F_B(s) = \sum_{s<T_0:s\in T_B} f_B(s) \geq \sum_{l=1}^{k^*} \frac{1}{2^l} = 1 - \frac{1}{2^{k^*}} \geq 1 - \delta.$$

□

A.7. **Proof of Lemma 2.** The idea of the proof is to divide the time in the game between the first and the last periods in which the rational types concede into three zones: early game (when Alice's types strength is not much higher than 1), late game (when the remaining types of Alice have strength close to the strongest type), and middle game (everything else). Lemma 9 implies that if Bob does not concede early, then all weak types of Alice must concede early. This allows us to bound from below Alice's probability of concession in the middle game. Next, we use the above-derived bounds to show that Alice must concede at a faster rate in the late game. We put those observations together to conclude that Bob does not concede sufficiently fast in the middle and late games; hence, he must concede with a large probability early.

We start with some notation. Recall that we assume that Δ is sufficiently small so that $\frac{1}{2}(S_B^* - 1) < \left(S_B^* - \mathrm{e}^{\Delta}\right)$. Let

$$\eta = \min\left(\frac{1}{2}(S_A^* - S_B^*), r_X, \frac{1}{C_1}\right) \text{ and } x = \frac{S_B^* - 1}{S_A^* - \eta - 1} < 1,$$

where $C_1 < \infty$ is the constant from Lemma 8 .

Lemma 6 implies that for each $t \in T_A$ st. $t < T_B^{*,\sigma}$, we have

$$p_A^{\sigma}(t) \geq \frac{1 - \mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{1}{u_B(x_A)} - \mathrm{e}^{-\Delta}} =: p_A^{\min}(\Delta).$$

Note that $p_A^{\min}(\Delta) \leq 4\Delta \frac{u_B(x_A)}{1-u_B(x_A)}$ for sufficiently small $\Delta > 0$.

For each sufficiently small $\delta > 0$, let

$$y(\delta) = 2\frac{\frac{1}{u_B(x_A)} - 1}{\frac{1}{2}\delta^2} \geq 1,$$

Find $\Delta_0(\delta)$ such that for each $\Delta \leq \Delta_0(\delta)$, we have (a) $\frac{1}{2}\frac{S_B^* - \mathrm{e}^{-\Delta}}{S_B^* - 1} \leq \frac{3}{4}$ and

$$1 - \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{1}{2}\delta^2} = 1 - y(\delta)\frac{1}{2}\frac{\frac{1}{u_B(x_A)} - \mathrm{e}^{-\Delta}}{\frac{1}{u_B(x_A)} - 1}\left(\frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}}\frac{1}{\frac{1}{u_B(x_A)} - \mathrm{e}^{-\Delta}}\right) \tag{A.11}$$
$$\geq 1 - y(\delta)\frac{3}{4}p_A^{\min}(\Delta) \geq \left(1 - p_A^{\min}(\Delta)\right)^{y(\delta)},$$

(the latter inequality holds if Δ is sufficiently small because $1 - y(\delta)\frac{3}{4}p \geq (1-p)^{y(\delta)}$ for a sufficiently small p), and (c) $1 - xp_A^* \geq (1-p_A^*)^x$, where recall that $p_A^* \leq 4\Delta\frac{1}{S_B^* - 1}$ has been defined in (A.1).

- *Early game*: By Lemma 9, for each $\delta > 0$, there exists $\Delta_1(\delta) \in (0, \Delta_0(\delta))$, such that for each $\Delta \leq \Delta_1(\delta)$, there exists T_0 such that $\mathrm{e}^{-\Delta T_0} \geq 1 - 2\delta$ and either (a) $F_B^\sigma(t) \leq \delta$, or (b) $\sigma_A^T(u) \leq T_0$ for all types $u \in \mathcal{U}_A$ such that $S_A(u) \leq 1 + \frac{1}{2}\delta^2$. If (a), the thesis of the Lemma holds. On the contrary, from now on, assume (b) and $F_B^\sigma(T_0) \geq \delta$.
- *Middle game* in periods t st. $T_0 \leq t < T_A^\eta$. In the middle game, $S_A(t) \geq 1 + \frac{1}{2}\delta^2$. Hence, by Lemma 6, we have

  $$p_A^\sigma(t) \geq p_A^{\min}(\Delta), \text{ and } p_B^\sigma(t) \leq \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}}\frac{1}{\frac{1}{2}\delta^2}.$$

  Then, inequality (A.11) implies that

  $$1 - p_B^\sigma(t) \geq \left(1 - p_A^{\min}(\Delta)\right)^{y(\delta)}. \tag{A.12}$$

- *Late game* in periods t st. $T_A^\eta \leq t < T^*$. By Lemma 8,

  $$\prod_{t \in T_A : T_A^\eta < t \leq T_A^{*,\sigma}} (1 - p_A^\sigma(t)) \leq \mathrm{e}^1 \prod_{t \in T_A : T_A^\eta < t \leq T_A^{*,\sigma}} (1 - p_A^*).$$

Moreover, the fact that $x<1$ implies that for each period t in the late game

$$\begin{aligned}1-p_B^\sigma\left(t\right) &\geq 1-\frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}}\frac{1}{S_A^*-\eta-\mathrm{e}^{-\Delta}} \\ &\geq 1-\frac{S_B^*-\mathrm{e}^{-\Delta}}{S_A^*-\eta-\mathrm{e}^{-\Delta}}p_A^*\geq 1-xp_A^*\geq\left(1-p_A^*\right)^x\end{aligned}$$

where the third inequality holds for $\Delta\leq\Delta_0\left(\delta,\eta\right)$.

Notice that for each player l and each t, $\lambda=F_l^\sigma\left(T^*\right)=F_l^\sigma\left(t\right)\prod_{s\in T_l:t\leq s\leq T^*}\left(1-p_l^\sigma\left(s\right)\right)$. The late-game estimates imply that

$$F_B^\sigma\left(T_A^\eta\right)=\frac{\lambda}{\prod_{s\in T_B:T_A^\eta\leq s\leq T^*}\left(1-p_B^\sigma\left(s\right)\right)}\leq\frac{\lambda}{\prod_{s\in T_B:T_A^\eta\leq s\leq T^*}\left(1-p_A^*\right)^x}=\lambda^{1-x}\left(F_A^\sigma\left(T_A^\eta\right)\right)^x\mathrm{e}^x.$$

Further,

$$\begin{aligned}\frac{F_B^\sigma\left(T_0\right)}{F_B^\sigma\left(T_A^\eta\right)} &= \frac{1}{\prod_{s\in T_B:T_0\leq s\leq T_A^\eta}\left(1-p_B^\sigma\left(s\right)\right)}\leq\prod_{s\in T_B:T_0\leq s\leq T_A^\eta}\left(1-p_A^{\min}\right)^{-y} \\ &\leq\left(\frac{F_A^\sigma\left(T_0\right)}{F_A^\sigma\left(T_i^\eta\right)}\right)^y\leq\left(F_A^\sigma\left(T_A^\eta\right)\right)^{-y},\end{aligned}$$

Together, we obtain,

$$F_B^\sigma\left(T_0\right)=\frac{F_B^\sigma\left(T_0\right)}{F_B^\sigma\left(T_A^\eta\right)}F_B^\sigma\left(T_A^\eta\right)\leq\lambda^{1-x}\left(F_A^\sigma\left(T_A^\eta\right)\right)^{x-y(\delta)}\mathrm{e}^x\leq\lambda^{1-x}\left(c_\pi\eta^{N-1}\right)^{x-y(\delta)}\mathrm{e}^x$$

where the last inequality comes from the fact that $F_i^\sigma\left(T_i^\eta\right)\geq\Pi\left\{u:h\left(u\right)\leq\eta\right\}\geq c_\pi\eta^{N-1}$ by Lemma 4. Because $x<1$, if λ is sufficiently small, we obtain the contradiction with $F_B^\sigma\left(T_0\right)\geq\delta$ for each $\Delta\leq\Delta_1$.

A.8. **If not** $B\left(\mathbf{0}_A,\frac{1}{4}d^2\right)\subseteq m_B^*$**.** Here, we show how the proof works if $B\left(\mathbf{0}_A,\frac{1}{4}d^2\right)\subseteq m_B^*$. We show that the elements of the above proof apply in this case and can be used to finish the proof of Lemma 2. Notice that Lemmas 6 and 9 do not require the assumption that $B\left(\mathbf{0}_A,\frac{1}{4}d^2\right)\subseteq m_B^*$. In particular, the proof Lemma 6 implies that Alice's types will sort by strength.

The fact that $B\left(\mathbf{0}_A, \frac{1}{4}d^2\right) \not\subseteq m_B^*$. implies that there is a type $u_0 \in \mathcal{U}_A$ such that $u_0 \cdot x_{u_0} \leq \frac{1}{4}d^2$. Let $W_0 = \left\{u \in \mathcal{U}_A : \sum |u_n - u_{0,n}| \leq \frac{1}{4}d^2\right\}$. For each type $u \in W_0$, we have $S_A(u) = \frac{u(x_A)}{u(x_u)} \geq \frac{d}{\frac{1}{2}d^2} = \frac{2}{d}$. Hence, $W_0 \subseteq W = \left\{u \in \mathcal{U}_A : S_A(u) \geq \frac{2}{d}\right\}$. Because $W_0 \subseteq W$ is a subset of $\mathcal{U}_A$ with a nonempty interior, it has a non-zero mass $\Pi(W) \geq \Pi(W_0) > 0$. On the other hand, $S_B^\sigma(t) \leq \frac{\max_{x\in X} u_B(x)}{u_B(x)} \leq \frac{1}{d}$ for any period t and any strategy σ. The latter implies that Lemma 6 implies that for each $t \in T_A$ st. $t < T_B^{*,\sigma}$, we have

$$p_B^\sigma(t) \geq \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{2}{d} - \mathrm{e}^{-\Delta}} = p_B^{\text{late}}.$$

Let T_A^{late} be the first period in which the first type $u \in W$ concedes. The estimates of Alice's strength imply that

$$\prod_{t\in T_A: T_A^\eta < t \leq T_A^{*,\sigma}} \left(1 - p_i^\sigma(t)\right) \leq \prod_{t\in T_A: T_A^\eta < t \leq T_A^{*,\sigma}} \left(1 - p_A^{\text{late}}\right),$$

where

$$p_A^{\text{late}} = \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{\frac{1}{d} - \mathrm{e}^{-\Delta}} > p_B^{\text{late}}.$$

This yields an equivalent of Lemma 8 for this case. An analogous argument as in the previous case finishes the proof of Lemma 2.

Appendix B. Proof of Theorem 1

Given the discussion in main body of the paper, the proof of Theorem is a consequence of the following result.

Lemma 10. *For each $\varepsilon > 0$ and $\delta > 0$, there is $\Delta^*, \lambda^* > 0$ such that if $\Delta < \Delta^*, \lambda < \lambda^*$, then for each player k, if $m_k = \left\{x^{Nash}(u)\right\}$ and*

$$\max_{x\in m_{-k}} u_k(x) < u_k\left(x^{Nash}(u)\right) - \varepsilon,$$

then player $-k$ concedes with a probability of at least $1-\delta$ in his first period of the game.

To prove the Lemma, define players' strengths as the ratios between winning and losing payoffs:

$$S_k = \frac{u_k\left(x^{\text{Nash}}(u)\right)}{\max_{x\in m_{-k}} u_k(x)}, S_{-k} = \frac{\max_{x\in \arg\max_{y\in m_{-k}} u_k(y)} u_{-k}(x)}{u_{-k}\left(x^{\text{Nash}}(u)\right)}.$$

In case of player $-k$'s strength, we choose as the winning payoff the best possible payoff that player $-k$ gets when player k concedes. (That plays a role for some, non-generic, menus, in which player k is indifferent between multiple optimal choices. Such a definition maximizes $-k$'s strength, and by the subsequent analysis, it creates the worst-case scenario for the Lemma. Also, we can assume that $S_{-k} > 1$, otherwise player $-k$ prefers to concede immediately.

Because the Nash allocation is unique, for each $\varepsilon > 0$, there is $r_\varepsilon > 0$ such that

$$\begin{aligned}\frac{S_k}{S_{-k}} &= \frac{u_k\left(x^{\text{Nash}}(u)\right) u_{-k}\left(x^{\text{Nash}}(u)\right)}{\max_{x\in m_{-k}} u_k(x) \max_{x\in \arg\max_{y\in m_{-k}} u_k(y)} u_{-k}(x)} \\ &\geq \frac{u_k\left(x^{\text{Nash}}(u)\right) u_{-k}\left(x^{\text{Nash}}(u)\right)}{\max_{x\in X \text{ st. } u_k(x)\leq u_k(x^{\text{Nash}}(u))-\varepsilon} u_k(x)\, u_{-k}(x)} \geq 1 + r_\varepsilon.\end{aligned}$$

Hence, $S_k \geq S_{-k} + r_\varepsilon$.

A simplified version of Lemma 6 holds. In particular, in each period in which a player i is still considered as potentially normal, the player concedes at rate

$$\begin{aligned} p^\sigma_{-k}(t) = p^*_{-k} &= \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{S_k - \mathrm{e}^{-\Delta}} \text{ for each } t \in T_{-k}, \\ p^\sigma_k(t) \geq p^*_k &= \frac{1-\mathrm{e}^{-2\Delta}}{\mathrm{e}^{-\Delta}} \frac{1}{S_{-k} - \mathrm{e}^{-\Delta}} \text{ for each } t \in T_k. \end{aligned}$$

It follows that, for sufficiently small Δ,

$$p^\sigma_k(t) \geq p^\sigma_{-k}(t) \frac{S_k - \mathrm{e}^{-\Delta}}{S_{-k} - \mathrm{e}^{-\Delta}} \geq p^\sigma_{-k}(t)\left(1 + \frac{r_\varepsilon}{S_{-k}}\right).$$

At this moment, the proof of Lemma 2 as stated in Section A.7 applies, which concludes the proof of the Lemma.

Appendix C. Proof of Lemma 3

Let $X_B = \arg\max_{x\in m_A} u_B(x)$ be the set of Bob's optimal choices from Alice's menu. Let $\upsilon_B = \max_{x\in m_A} u_B(x) < \frac{1}{2} - \eta$ be Bob's optimal concession payoff. Assume that $\varepsilon > 0$ is small enough that

$$\left(\frac{1}{2} - \varepsilon\right)^2 > \upsilon_B (1 - \upsilon_B).$$

For each v, let $X(v) = \{x \in X : u_B(x) \geq v\} = \{x \in X : u_B \cdot x \leq 1 - v\}$ be the set of player i's allocations of player i that leaves Bob with a payoff of at least v. (Recall that Bob's payoff is equal to $u_B(x) = 1 - u_B \cdot x$.) Notice that $X_B \subseteq X(\upsilon_B) \subseteq \frac{1-\upsilon_B}{1/2} X\left(\frac{1}{2}\right)$, where the last set is a subset of $\mathbb{R}^N$ obtained by multiplication of vectors in $X\left(\frac{1}{2}\right)$ by a constant $\frac{1-\upsilon_B}{1/2}$. Then, for each $u_A \in \mathcal{U}_A$, $u_A(x) = u_A \cdot x$, and

$$\max_{x\in X_B} u_A(x) \leq \max_{x\in X(\upsilon_B)} u_A \cdot x \leq \max_{x\in \frac{1-\upsilon_B}{1/2} X\left(\frac{1}{2}\right)} u_A \cdot x = \frac{1-\upsilon_B}{1/2} \max_{x\in X\left(\frac{1}{2}\right)} u_A(x) = \frac{1-\upsilon_B}{1/2} \max_{x\in m^{1/2}} u_A(x).$$

Suppose that Bob's offer m_B is ε-close to $m^{1/2}$. The above inequality implies that for any strategy profile, for each $t \in T_B$, and each $u_A \in \mathcal{U}_A$,

$$S_A^\sigma(u_A, t) \leq \frac{\max_{x\in X_B} u_A(x)}{\max_{x\in m_B} u_A(x)} \leq \frac{1-\upsilon_B}{1/2 - \varepsilon}.$$

Additionally, for each $t \in T_A$,

$$S_B^\sigma\left(u_i^B, t\right) \geq \min_{u_A\in\mathcal{U}_A} \frac{\min_{x\in \arg\max_{x\in m_B} u_A(x)} u_B(x)}{\max_{x\in m_A} u_B \cdot x} \geq \frac{\frac{1}{2} - \varepsilon}{\upsilon_B} > \frac{1-\upsilon_B}{1/2 - \varepsilon},$$

where the last inequality comes from the choice of ε. By Lemma 6, in each period, Alice's concession rate is strictly smaller than Bob's concession rate.

$$p_A^\sigma(t) \leq p_A^\Delta = \frac{1 - e^{-2\Delta}}{e^{-\Delta}} \frac{1}{\frac{\frac{1}{2}-\varepsilon}{\upsilon_B} - e^{-\Delta}} < p_B^\Delta = \frac{1 - e^{-2\Delta}}{e^{-\Delta}} \frac{1}{\frac{1-\upsilon_B}{\frac{1}{2}-\varepsilon} - e^{-\Delta}} \leq p_B^\sigma(t).$$

It follows (see, for instance, the proof of Lemma 2) that

$$\frac{F_A^\sigma\left(t_A^0+2\right)}{F_B^\sigma\left(t_B^0+2\right)}=\frac{\lambda/\prod_{t\in T_A:t_A^0\leq t\leq T_A^*}\left(1-p_A^\sigma\left(t\right)\right)}{\lambda/\prod_{t\in T_B:t_B^0\leq t\leq T_B^*}\left(1-p_B^\sigma\left(t\right)\right)}<\left(\frac{1-p_B^\Delta}{1-p_A^\Delta}\right)^{\frac{1}{2}T^*}\rightarrow 0,$$

as $T^*\rightarrow\infty$. Similar arguments to those used in the proof of Lemma 2 show that T^* is arbitrarily large if Δ and λ are small. Thus, for a sufficiently small Δ and λ, $f_A^\sigma\left(t_A^0\right)$ is arbitrarily close to 1.

Appendix D. Proof of Theorem 2

The order in which the two players choose menus in the menu choice game is not important for this argument. By a remark after Lemma1 1 and by Lemma 2, for each $d>0$ and each $\delta>0$, there is λ^*,Δ^* such that for each $\lambda\leq\lambda^*,\Delta\leq\Delta^*$, by choosing an arbitrary allocation $x_A\in\left\{x\in m^{1/2}:\forall_n d\leq x_n\leq 1-d\right\}$, Alice with a signal $s\in\mathcal{U}_A$ can ensure herself a payoff of at least $\left(1-\delta\right)\left(E_S\left(u\left(x_A\right)|s\right)-d\right)$ (either because Bob will concede with probability $1-\delta$, or because Bob will propose a menu with a payoff of at least $u\left(x_A\right)-d$ for Alice. The proof of Lemma 2 shows that constants $\lambda^*,\Delta^*>0$ can be chosen in a way that depends only on d and δ. By taking $d\rightarrow 0$ and $\delta\rightarrow 0$, we observe that Alice's payoffs converges to $\max_{x\in m^{1/2}}E_S\left(u\left(x\right)|s\right)$. Hence, As $\pi_S\rightarrow\delta$, we obtain $E_A\left(u_A;\mathcal{C}_X,\pi,k\right)=\max_{x\in m^{1/2}}u_A\left(x\right)$.

Lemma 3 shows that for each $\eta>0$ and $\delta>0$, there exists $\lambda^*,\Delta^*>0$ such that for each $\lambda\leq\lambda^*,\Delta\leq\Delta^*$, by choosing a menu $m^{1/2}$, Bob can ensure a payoff of at least $\left(1-\delta\right)\left(\frac{1}{2}-\eta\right)$. It follows that $E_B\left(u_B;\mathcal{C}_X,\pi,k\right)\geq\frac{1}{2}$. This concludes the proof of the Theorem.

Appendix E. Proof of Proposition 1

For each menu m, each Alice's type u_A, let $u_A\left(m\right)=\max_{x\in m}u_A\left(x\right)$. The proof of Theorem 2 shows that, for any Bob's offer m_B, Alice's type u_A can ensure a payoff

$$\max\left(u_A\left(m_B\right),u_A\left(m^{1/2}\right)\right).$$

If $u_A(m_B) < u_A\left(m^{1/2}\right)$, then Alice's counteroffer $\arg\max_{x \in m^{1/2}} u_A(x)$ (which gives her payoff $u_A\left(m^{1/2}\right)$) makes her stronger; by making her offer a bit more attractive for herself, she remains stronger and wins the war of attrition with a payoff strictly higher than $u_A\left(m^{1/2}\right)$). In other words, unless $u_A(m_B) = u_A\left(m^{1/2}\right)$, Alice is going to get a payoff that is strictly higher than $u_A\left(m^{1/2}\right)$.

Now, the claim follows from two simple observations. First, because menu $m^{1/2}$ is efficient, if a positive mass of Alice types gets a payoff that is strictly higher $u_A\left(m^{1/2}\right)$, Bob gets an expected payoff that is strictly lower than $\frac{1}{2}$. Second, if Bob likes at least two parts of the pie, $u_B^i, u_B^j > 0$ for $i \neq j$, then there is no simple offer $m_B = \{x_B\}$ such that $u_A(x_B) \geq \frac{1}{2}$ and $u_A(m_B) = u_A\left(m^{1/2}\right)$ for every single Alice type u_A. To see that notice that, for any offer x_B st. $u_B(x_B) = \frac{1}{2}$, there is offer x'_B that redistributes parts i and j in a way that keeps Bob's payoff the same and increases payoff of at least one of Alice's types. Hence the original allocation could not have been optimal for such Alice's type.

References

Abreu, D. and Gul, F. (2000). Bargaining and Reputation. *Econometrica*, **68** (1), 85–117.

—, Pearce, D. and Stacchetti, E. (2015). One-sided uncertainty and delay in reputational bargaining. *Theoretical Economics*, **10** (3), 719–773.

de Clippel, G., Fanning, J. and Rozen, K. (2019). Bargaining over contingent contracts under incomplete information. *Available at SSRN 3527102*.

Fanning, J. (2016). Reputational Bargaining and Deadlines. *Econometrica*, **84** (3), 1131–1179.

Grossman, S. J. and Perry, M. (1986). Perfect sequential equilibrium. *Journal of Economic Theory*, **39** (1), 97–119.

Gul, F., Sonnenschein, H. and Wilson, R. (1986). Foundations of dynamic monopoly and the coase conjecture. *Journal of Economic Theory*, **39** (1), 155–190.

Harsanyi, J. C. and Selten, R. (1972). A Generalized Nash Solution for Two-Person Bargaining Games with Incomplete Information. *Management Science*, **18** (5(2)), 80–106.

Inderst, R. (2003). Alternating-offer bargaining over menus under incomplete information. *Economic Theory*, **22** (2), 419–429.

Jackson, M. O., Sonnenschein, H., Xing, Y., Tombazos, C. and Al-Ubaydli, O. (2020). *The Efficiency of Negotiations with Uncertainty and Multi-Dimensional Deals*. Tech. rep., Social Science Research Network, Rochester, NY.

Kambe, S. (1999). Bargaining with Imperfect Commitment. *Games and Economic Behavior*, **28** (2), 217–237.

Myerson, R. B. (1984). Two-Person Bargaining Problems with Incomplete Information. *Econometrica*, **52** (2), 461–487.

— (1991). *Game Theory*. Harvard University Press.

Nash, J. (1953). Two-person cooperative games. *Econometrica: Journal of the Econometric Society*, pp. 128–140.

Nash Jr, J. F. (1950). The bargaining problem. *Econometrica: Journal of the Econometric Society*, pp. 155–162.

Peski, M. (2019). Alternating-Offer Bargaining with Incomplete Information.

Rubinstein, A. (1982). Perfect equilibrium in a bargaining model. *Econometrica*, **50** (1).

— (1985). A Bargaining Model with Incomplete Information About Time Preferences. *Econometrica*, **53** (5), 1151–1172.

Sen, A. (2000). Multidimensional Bargaining Under Asymmetric Information. *International Economic Review*, **41** (2), 425–450.

Serrano, R. (2004). Fifty years of the Nash program, 1953-2003.

Strulovici, B. (2017). Contract Negotiation and the Coase Conjecture: A Strategic Foundation for Renegotiation-Proof Contracts. *Econometrica*, **85** (2), 585–616.

Wang, G. H. (1998). Bargaining over a Menu of Wage Contracts. *The Review of Economic Studies*, **65** (2), 295–305.