Optimally Stubborn

How long to hold and who will fold?*

Anna Sanktjohanser†

December 26, 2025

Abstract

I consider a model of reputational bargaining in which the stubborn type can choose their initial demand. There are two types of players: rational and stubborn. The game has two stages: a demand stage and a concession stage. Types can pool or separate in equilibrium for any fixed probability of facing a stubborn type. When the probability of facing a stubborn type is small, any feasible payoff can be achieved in equilibrium for either type. When the probability is large, there is either immediate agreement or long delays.

*I am very grateful to Marina Halac and Johannes Hörner for many insightful conversations and comments. I would like to thank the two referees, Dilip Abreu, James Best, Ben Brooks, Olivier Compte, Joyee Deb, Péter Esö, Jack Fanning, Faruk Gul, Philippe Jehiel, Kyungmin Kim, Chiara Margaria, Margaret Meyer, Sujoy Mukerji, Juan Ortner, Daniel Quigley, Rajiv Sethi, Jean Tirole, Juuso Välimäki, Alex Wolitzky, and William Zame, as well as various seminar and conference audiences for helpful discussions and comments. I gratefully acknowledge funding from the ERC Grant N^o 340903 and from ANR under grant ANR-17-EURE-0010 (Investissements d'Avenir program).

†Toulouse School of Economics, 1 Esplanade de l'Université, 31080 Toulouse, France, anna.sanktjohanser@gmail.com.

1 Introduction

In political bargaining, leaders often make public commitments to bolster their position with foreign counterparts and to signal resolve to domestic audiences. During the Brexit negotiations, Prime Minister Theresa May publicly articulated a set of "red lines" (e.g., leaving the single market and the customs union).[1] Walking these back risked severe domestic penalties in Parliament and within her party. Leaders typically have better private information about the domestic political costs of reneging, whereas counterparts may be uncertain about how damaging any backtracking would be to the leader's credibility at home. How do such strategic public postures shape the outcome of trade?

This paper explores the middle ground between fully rational agents and behavioral agents, providing a framework for studying bargaining dynamics when players can strategically commit to postures. Motivated by examples like May during the Brexit negotiations, I propose a model of reputational bargaining in which the behavioral type can choose their initial demand – essentially determining the posture they wish to project. This is a departure from the strategy restriction on behavioral types typically made in the literature on reputational bargaining (Myerson 1991, Abreu and Gul 2000 [AG] and follow-up papers), where behavioral types cannot choose their initial demand (and more generally, have no choices to make). Specifically, I consider a bargaining game with two types of players: rational and stubborn. The game has two stages: a demand stage and a concession stage. Players simultaneously make demands, and the game concludes when one player concedes to the other's demand. Rational players can concede at any point in time, whereas stubborn players can choose their initial demand but cannot concede thereafter.

I establish the existence of both pooling and separating equilibria for any fixed probability of facing a stubborn type. When this probability is low, I show that any

[1] *The New York Times*, "In 'Brexit' Speech, Theresa May Outlines Clean Break for U.K.," January 17, 2017

payoff, for both rational and stubborn players, can be achieved in equilibrium. In fact, this is true both for pooling and separating equilibria. When the probability is high, there is either immediate agreement or there are prolonged delays, causing payoffs for both players, regardless of type, to be arbitrarily small.

The intuition behind these results is as follows. A rational player never benefits from making a demand known to be compatible with the opponent's, since they can always concede later. In contrast, a stubborn player risks losing any chance of agreement if their demand is incompatible with that of a similarly stubborn opponent. This difference in preferences allows for type separation in equilibrium.

However, types can also pool over multiple demands. As in AG, players (regardless of their type) face a tradeoff between the amount received if the opponent concedes and the speed with which the opponent concedes. However, this trade-off is not the same for the two types. When demands are compatible, the two types receive the same payoff. When demands are incompatible, there is a cost of being stubborn (relative to being rational). This cost is smaller the higher the demands. Higher demands imply a longer war of attrition and hence, the stubborn type's cost of not being able to concede is paid "far in the future."[2] Appropriate punishment with off-path beliefs imply that deviations to other demands are not profitable for either type. This together implies that types can pool over multiple demands in equilibrium.

A low probability of facing a stubborn type drives equilibrium multiplicity, as the benefits of making incompatible demands (and leaving nothing on the table) outweigh the risk of facing a similarly stubborn opponent. This creates a force toward multiple equilibria, where any payoff can be achieved in equilibrium. In contrast, a high probability of facing a stubborn type incentivizes players to either reach immediate agreement or escalate to high demands, prolonging delays and polarizing outcomes. Thus, the proportion of behavioral types significantly shapes negotiation dynamics

[2] A rational player is willing to wait to concede only so long as he is uncertain about the opponent's type. Hence, the length of the war of attrition determines the payoff difference between the two types.

and payoffs.

Much of the literature on reputational bargaining focuses on the limiting case where the probability of encountering a stubborn opponent is small. In contrast, the analysis here considers a fixed ex-ante probability of stubbornness, which aligns with empirical evidence suggesting that stubborn behavior is not uncommon. For example, Backus et al. (2020) examine behavior patterns in bilateral bargaining using data from eBay's Best Offer platform. They find that a significant portion of negotiations end in disagreement after a delay. Such outcomes are difficult to reconcile with reputational bargaining models that assume a low probability of stubbornness. In my model, however, when the probability of facing a stubborn opponent is high, the likelihood of perpetual disagreement is substantial.

The results on prolonged delays resonate with documented real-world phenomena across various applications. In US politics, for instance, Binder (1999) provides evidence that intrabranch conflict (and hence, uncertainty as to whether a branch is willing or able to make concessions) is critical in shaping deadlock. Similarly, Card (1990) provides evidence that longer strikes in labor disputes are associated with lower wage settlements – potentially reflecting employers' stubbornness.

My results emphasize the importance of defining behavioral types based on the specific context, as different economic applications may call for different approaches. This model is suited to analyzing situations where agents privately know whether making demands or threats will limit their flexibility, though this constraint is not common knowledge. For instance, a political leader may issue demands during international negotiations while privately aware that backtracking later would incur significant political costs (or be impossible). This creates a so-called audience cost: a domestic political penalty when foreign policy actions are perceived as unsuccessful.[3] Such scenarios are not adequately captured by standard models of reputational bargaining, where players with audience costs (behavioral types) are not strategic players. In contrast, here behavioral types strategically choose their demands while

[3]See Fearon (1994) and follow-up papers, Ozyurt (2014) and Ozyurt (2015a,b).

anticipating the impossibility of future concessions. This allows opponents to update their beliefs about a leader's constraints based on the demands made. For example, during the 2023 US debt ceiling crisis, Republicans insisted on spending cuts as a condition for raising the debt ceiling, while Democrats argued for a "clean bill" without preconditions. Both sides strategically weighed the political costs of conceding against the benefits of holding firm on their demands.[4]

This paper contributes to the literature on bargaining with two-sided incomplete information. In bargaining with two-sided private information about valuations for a good, a player's offer can serve as a signal of their information, which can lead to multiplicity of equilibria. Signaling allows a player to be "punished with beliefs" for deviating from a proposed equilibrium path. This can support a wide variety of behavior, ranging from no trade (Ausubel and Deneckere 1992) to Myerson and Satterthwaite's (1983)'s constrained efficient bounds (Ausubel and Deneckere 1993).[5]

When the two-sided private information is about a player's ability to concede rather than their value for the good, this multiplicity disappears. Models on reputational bargaining (Myerson 1991, AG and follow-up papers) have shown that the so-called Rubinstein-Stahl outcome is the unique outcome in a large class of bargaining protocols when the game is perturbed with simple behavioral types who are committed to a fixed stance.[6] This clear prediction of reputational bargaining models therefore stands in stark contrast to models where the private information is about a players' values for the good. In reputational bargaining models, belief-based punishments do not arise, despite two-sided incomplete information, because committed types have no choices to make and are therefore immune to such punishments: they insist on their pre-specified demands (regardless of their actual preferences over such

[4] *The Guardian*, "Danger and deja vu: what 2011 can tell us about the US debt ceiling crisis," April 30, 2023.

[5] For a helpful discussion, see the survey by Fanning and Wolitzky (2022).

[6] Among others, see Abreu and Pearce (2007), Abreu, Pearce, and Stacchetti (2015), Abreu and Sethi (2003), Atakan and Ekmekci (2014), Compte and Jehiel (2002), Fanning (2016, 2018, 2021), Kambe (1994), Kim (2009), Wolitzky (2012).

demands), forcing behavior onto the equilibrium path.

A closely related paper is Kambe (1994), who also departs from AG by allowing demands to be chosen rather than assume an exogenously fixed set of commitment demands. Specifically, players do not know at the time of choosing whether they are rational or committed: they first select a demand, and only afterward may discover that they are bound to it. This preserves the payoff predictions, because off-path punishments cannot be freely used once demands are precommitted in this way.

My model is closer to AG in that types are realized before demands are chosen, but differs crucially in allowing stubborn types to choose strategically. This flexibility by the stubborn type has two implications. First, there are more deviations to consider than in AG, as the stubborn type also needs to be appropriately incentivized. This in itself is a force towards unique predictions. However, there is a second implication: conditional on being able to appropriately incentivize the stubborn type, there is greater flexibility on which demands can be assigned positive probability as coming from a behavioral type *on path*. This means it is easier to incentivize the rational type to be willing to make certain demands. In order to "force" behavior by the stubborn type onto the equilibrium path for such a variety of demands, we need the possibility of belief-based punishments *off-path*: without them, we could not deter these additional deviations by the stubborn type. In fact, these off-path belief-based punishments are not enough to deter the stubborn type when the probability of facing a stubborn type is high. Hence, the flexibility of the stubborn types together with the possibility of belief-based punishments off-path are two crucial features of my environment that allow for Folk theorem like payoff multiplicity when the probability of facing a stubborn type is small.

This paper explores the middle ground between fully rational agents and fully committed agents who have no strategic choices to make. When behavioral types are given the ability to choose their initial stance, the possibility of belief-based punishments plays an important role in establishing a Folk Theorem when the ex-ante probability of encountering a behavioral type is small. Reducing the probability of being behavioral in classical models of reputational bargaining is a force towards efficiency. Here, it is a force toward multiplicity. Equilibria retain a war of attrition structure, but uncommitted players no longer necessarily mimic behavioral types. When the ex ante probability of behavioral types is high, the Folk Theorem breaks down: the force of belief-based punishments is not enough to deter stubborn types from deviating (to compatible demands).

2 Model

Time is continuous, and the horizon is infinite. Two players decide on how to split a unit surplus. At time $t = -1$, players i and j simultaneously announce demands, α^i and α^j, with $\alpha^i, \alpha^j \in [0,1]$. If $\alpha^i + \alpha^j \leq 1$, the demands are said to be *compatible*. In this case, the game ends. If $\alpha^i + \alpha^j > 1$, the demands are *incompatible*. In this case, a concession game starts at $t = 0$. The game ends when one player concedes. Concession means agreeing to the opponent's demand.

Each player i is rational with probability $1 - z$ and stubborn with probability z, where $z \in (0,1)$. Before the game starts, each player privately learns whether he is stubborn or rational. A rational player $i = 1, 2$ can make any demand $\alpha^i \in [0,1]$ at time 0 and concede to his opponent at any point in time. Stubborn player i can choose his initial demand $\alpha^i \in [0,1]$ but cannot concede to his opponent. Note that this is unlike in AG, where a stubborn player cannot choose his initial demand. A strategy for a stubborn player, i, $\sigma^{s,i}$, is defined by a Borel probability measure s^i on $[0,1]$ (his demand α^i). A strategy for a rational player i, $\sigma^{r,i}$, is defined by a Borel probability measure r^i on $[0,1]$ (his demand α^i) and a collection of cumulative distributions $F^{r,i}_{k,\ell}$ on $\mathcal{R}_+ \cup \{\infty\}$ for each incompatible pair of realized demands $(\alpha^i, \alpha^j) = (\alpha_k, \alpha_\ell)$ with $\alpha_k + \alpha_\ell > 1$.[7] Throughout, I restrict attention to equilibria that involve strategies s^i and r^i with finite support.[8] For any $x \in [0,1]$, let δ_x denote the Dirac measure at x.

[7] The set of (demand) strategies (measures) is endowed with the topology of weak convergence. Throughout, this is how limits are understood.

[8] That is, I focus on equilibria such that this is the case, but I do not impose that players cannot deviate to arbitrary demand strategies.

Thus, a strategy with finite support $\mathcal{C} = \{\alpha_1, \ldots, \alpha_K\}$ and weights $(r^i_1, \ldots, r^i_K)$ for rational player i (resp. $(s^i_1, \ldots, s^i_K)$ for stubborn player i) is

$$r^{\,i} = \sum_{k=1}^{K} r^i_k\, \delta_{\alpha_k}, \qquad s^{\,i} = \sum_{k=1}^{K} s^i_k\, \delta_{\alpha_k},$$

with $r^i_k, s^i_k \geq 0$ and $\sum_{k=1}^{K} r^i_k = \sum_{k=1}^{K} s^i_k = 1$. Throughout, I use the notation that lower demands have lower subscripts, i.e., $\alpha_k < \alpha_{k+1}$.

For realized demands $\alpha^i = \alpha_k$ and $\alpha^j = \alpha_\ell$ with $\alpha_k + \alpha_\ell > 1$, let $F^{r,i}_{k,\ell}(t)$ denote the probability that a *rational* player i concedes to player j by time $t \geq 0$. The (unconditional) probability that player i concedes by time t is

$$F^i_{k,\ell}(t) = \big(1 - \pi^{\,i}(\alpha_k)\big)\, F^{r,i}_{k,\ell}(t),$$

where the posterior that i is stubborn after observing $\alpha^i = \alpha_k \in \mathcal{C}$ at time $t = -1$ (given $\sigma^{r,i}$ and $\sigma^{s,i}$) is

$$\pi^{\,i}(\alpha_k) = \frac{z\, s^i_k}{z\, s^i_k + (1-z)\, r^i_k}.$$

Therefore, if $\alpha^i = \alpha_k \in \mathcal{C}$,

$$\lim_{t\to\infty} F^i_{k,\ell}(t) \;\leq\; 1 - \pi^{\,i}(\alpha_k).$$

Note that $F^i_{k,\ell}(0)$ may be positive; this is the probability that i immediately concedes to j.

Player i's discount rate is $\rho > 0$, for $i = 1, 2$. The continuous-time bargaining problem is denoted $B = \{z, \rho\}$. If $\alpha^i + \alpha^j \leq 1$ at $t = 0$, player i receives α^i and $1 - \alpha^j$ with probability 1/2.

Suppose that $(\alpha^i, \alpha^j) = (\alpha_k, \alpha_\ell)$ is observed at time 0, with $\alpha_k + \alpha_\ell > 1$. Given a strategy profile $\bar{\sigma} = (\sigma^{\,i}, \sigma^{\,j})$ with $\sigma^{\,i} = (\sigma^{r,i}, \sigma^{s,i})$, the expected payoff to *rational* player i from conceding at time t is

$$\begin{aligned} U^{\,i}(t, \sigma^{\,j} \mid k, \ell) = \alpha_k \int_{y<t} e^{-\rho y}\, \mathrm{d}F^j_{\ell,k}(y) \;+\; \frac{\alpha_k + 1 - \alpha_\ell}{2}\, \big(F^j_{\ell,k}(t) - F^j_{\ell,k}(t^-)\big) e^{-\rho t} \\ + (1-\alpha_\ell)\, \big(1 - F^j_{\ell,k}(t)\big) e^{-\rho t}, \end{aligned}$$

where $F^j_{\ell,k}(t^-) := \lim_{y\uparrow t} F^j_{\ell,k}(y)$.Thus player i receives (i) the discounted value of α_k if j concedes first, (ii) $(\alpha_k + 1 - \alpha_\ell)/2$ if they concede simultaneously at t, and (iii) $1 - \alpha_\ell$ if i concedes first. If i never concedes,

$$U^i(\infty, \sigma^j \mid k, \ell) = \alpha_k \int_{[0,\infty)} e^{-\rho y}\, \mathrm{d}F^j_{\ell,k}(y),$$

which coincides with the stubborn player's payoff when facing an incompatible demand.

Since $F^i_{k,\ell}$ describes the concession behavior of a player, unconditional on his type, a rational player i's concession behavior is described by is $F^{r,i}_{k,\ell} = 1/(1 - \pi^i(\alpha_k))\, F^i_{k,\ell}$. Therefore the rational player's expected utility from using $F^i_{k,\ell}$, conditional on (α_k, α_ℓ), is

$$U^i(\bar\sigma \mid k, \ell) = \frac{1}{1 - \pi^i(\alpha_k)} \int_{[0,\infty)} U^i(y, \sigma^j \mid k, \ell)\, \mathrm{d}F^i_{k,\ell}(y).$$

Finally, the rational player's ex-ante expected utility under $\bar\sigma$ is

$$U^i(\bar\sigma) = \sum_{k=1}^{K} r^i_k \Bigg[\underbrace{\sum_{\ell:\, \alpha_\ell \le 1-\alpha_k} \frac{\alpha_k + 1 - \alpha_\ell}{2} \big((1-z)\, r^j_\ell + z\, s^j_\ell\big)}_{\text{compatible at } t=-1} + \underbrace{\sum_{\ell:\, \alpha_\ell > 1-\alpha_k} U^i(\bar\sigma \mid k, \ell)\, \big((1-z)\, r^j_\ell + z\, s^j_\ell\big)}_{\text{incompatible, concession game}} \Bigg].$$

For later use, define the opponent's probability of facing α_k by

$$q^{(-i)}_k := (1-z)\, r^{(-i)}_k + z\, s^{(-i)}_k,$$

and define player i's *strength* at α_k by

$$\mu^i(\alpha_k) := \big(\pi^i(\alpha_k)\big)^{\frac{1}{1-\alpha_k}}.$$

Write $\mu^i_k := \mu^i(\alpha_k)$. In all formal statements and proofs the index notation μ^i_k is used, while numerical examples with explicit demands use the functional notation $\mu^i(\alpha)$; when $\alpha = \alpha_k$ this coincides with μ^i_k.

I follow the literature in modeling the bargaining as a war of attrition (rather than allowing players to revise their demands). This is inspired by Myerson's (1991) insight that revising one's demand reveals rationality, so that it is equivalent to conceding (in the context of his model, which is closely related, but not identical to mine).[9]

A crucial feature of the model – one that departs from AG – is that initial demands are chosen simultaneously. This assumption is essential for the existence of pooling equilibria involving more than one offer. If demands were instead made sequentially, symmetric pooling would not be robust, although Folk-theorem-like payoff multiplicity would still arise.

For the analysis in $B = \{z, \rho\}$, I use the solution concept of (weak) Perfect Bayesian equilibrium (PBE). A PBE is a profile of strategies $\sigma^* = (\sigma^{1*}, \sigma^{2*})$ and a system of initial beliefs mapping demands into probabilities that a player is stubborn,

$$\pi^i : [0,1] \to [0,1] \quad \text{for} \;\; i = 1, 2,$$

such that (1) the strategy maximizes a player's expected utility (given beliefs), and (2) if an information set is reached with positive probability given the strategy profile, beliefs are formed according to Bayes' rule; and if an information set is not reached with positive probability given the strategy profile, beliefs are arbitrary probabilities that a player is stubborn.[10]

Henceforth, equilibrium refers to weak PBE (see Fudenberg and Tirole, 1991 for a definition).[11] I focus on symmetric equilibria. By symmetric equilibria I mean equilibria where $r^i = r^j$ and $s^i = s^j$ – i.e., the identity of a player does not matter. Only

[9] AG show that any convergent sequence of equilibrium outcomes within a broad family of discrete-time games must converge to the unique continuous-time equilibrium outcome as the maximum time between consecutive opportunities to revise demands goes to 0. Of course, the modeling of AG differs from mine in some respects. Moreover, types in AG do not separate in equilibrium. As we will see, they can do so here, and hence, modeling the bargaining as a war of attrition entails some loss of generality here.

[10] While PBE is permissive with regards to off-path beliefs, this flexibility turns out to matter only for "high" off-path demands.

[11] To the extent that the concession behavior is a direct consequence of the demands made, I refer to an equilibrium by its support and the probabilities associated with that support.

his type does. This suffices to establish payoff multiplicity.[12] To simplify notation, I omit superscripts indicating a player's identity unless clarification is necessary.

I denote an equilibrium by (z, r, s), where z is the probability of a stubborn type, r is the rational type's strategy, and s is the stubborn type's strategy. The corresponding equilibrium payoffs are $v_r(z, r, s)$ for the rational type and $v_s(z, r, s)$ for the stubborn type. For brevity, I often write simply v_r and v_s when the equilibrium is clear from context.

3 Main Results

The two main propositions establish existence of pooling and separating equilibria for a given ex ante probability of facing a stubborn type, z. I define an equilibrium as pooling if both types (of both players, by symmetry) make the same demands in equilibrium with positive probability, and as separating if each demand made in equilibrium perfectly reveals the player's type. Additionally, the propositions establish limits for payoffs and probabilities as the probability of encountering either type vanishes. Specifically, I establish Folk theorem-like payoff multiplicity when the stubborn type vanishes. Conversely, when the probability of the rational type vanishes, there is either immediate agreement or infinite delay.

Before turning to non-degenerate pooling, it is useful to note that there always exist degenerate single-demand pooling equilibria. If all types demand $\alpha = 1/2$, agreement is immediate and both players obtain the efficient payoff of 1/2. If all types demand $\alpha = 1$, then there is infinite delay and both players obtain a payoff of 0. We now turn to the existence of non-degenerate pooling equilibria, in which types pool on more than one demand, and which give rise to the Folk-theorem result.

To illustrate, let us begin with an example of a pooling equilibrium featuring two demands. Suppose the prior probability of the stubborn type is $z = 1/10$. Then the following is an equilibrium: the rational and the stubborn type randomize over 1/3

[12] Note that given the assumption of symmetry, multiplicity refers to delay rather than division of surplus.

and 4/5 with

$$r \approx \delta_{1/3}0.381 + \delta_{4/5}0.619, \text{ and } s \approx \delta_{1/3}0.230 + \delta_{4/5}0.770.$$

Once the initial demands have been made, the war of attrition proceeds exactly as in AG.[13] If both players demand 1/3, there is immediate agreement (which ends the game). If both players demand 4/5, there is no immediate concession, and the rational player concedes at a rate that keeps their opponent indifferent between waiting and conceding. If players make different demands, then the player demanding 4/5 concedes immediately with probability ≈ 0.72. Thereafter, players concede at a rate that keeps their opponent indifferent between waiting and conceding (meaning the player demanding 4/5 concedes at a slower rate than the player demanding 1/3). If a player faces an unexpected, incompatible demand $\alpha \notin \{1/3, 4/5\}$, he does not concede.

After the initial demands are made, a player demanding 1/3 is believed to be stubborn with probability $\pi(1/3) \approx 0.063$. A player demanding 4/5 is believed to be stubborn with probability $\pi(4/5) \approx 0.121$. If either player demands $\alpha \notin \{1/3, 4/5\}$, he is believed to be rational with probability 1.[14]

[13]Specifically, the war of attrition for incompatible demands unfolds as follows. The player making the higher demand concedes with positive probability at time 0. More precisely, player j's immediate concession probability when demanding $\alpha^j = \alpha_k$ and facing a demand of $\alpha^i = \alpha_\ell$ is given by

$$F^j_{k,\ell}(0) = \max\left\{1 - \left(\frac{\mu_k}{\mu_\ell}\right)^{1-\alpha_k}, 0\right\}.$$

In equilibrium $\mu_k \geq \mu_\ell$ if and only if $\alpha_k < \alpha_\ell$ – see Lemma 1 in the Appendix. Thereafter, players concede at a rate that keeps their opponent indifferent between waiting and conceding. Finally, there is a finite time by which the posterior probability of stubbornness reaches 1 simultaneously for both players and concessions by the rational type stop.

[14]Note that for any deviating demand $\alpha \leq 0.619$, it is sufficient that the deviating demand α is believed to come from the stubborn type with probability $\pi(4/5)^{(1-\alpha)/(1-4/5)}$ so that $\mu(4/5) = \mu(\alpha)$. This ensures that a demand 4/5 does not lead to concession with positive probability at time 0 to such an out-of-equilibrium demand, but simply to concession at a rate that keeps the opponent indifferent between waiting and conceding. For demands $\alpha > 0.619$, we ensure equilibrium existence

These strategies result in the following payoffs. A rational player demanding 1/3 receives 1/2 when facing a demand of 1/3, and receives a payoff of 0.2962 when facing a demand of 4/5. A rational player demanding 4/5 receives 0.6667 when facing a demand of 1/3 and a payoff of 0.2 when facing a demand of 4/5. Hence, fixing the opponent's strategy, the rational player expects the same payoff of 0.3708 from either of the equilibrium demands 1/3 and 4/5. A stubborn player demanding 1/3 receives 1/2 when facing a demand of 1/3, and receives of a payoff of 0.2946 when facing a demand of 4/5. A stubborn player demanding 4/5 receives 0.6426 when facing a demand of 1/3 and receives 0.19996 when facing a demand of 4/5. Hence, fixing the opponent's strategy, the stubborn player expects the same payoff of 0.3619 from either demands 1/3 or 4/5. The difference in the two types' payoffs comes from the fact that the stubborn type is unable to concede even when he puts probability 1 on being faced with a stubborn opponent. This difference in the two types' payoffs is smaller the smaller the ex ante probability of facing a stubborn type. As a result, in the limit, as the stubborn type vanishes, the payoff to the two types is identical.

Figure 1 illustrates the expected payoff from an equilibrium demand (shown as red and blue dots for the rational and stubborn type respectively at $\alpha = 1/3, 4/5$) and contrasts it to the payoff a rational (stubborn) player could receive from deviating to any other demand. Since after a deviation an opponent never concedes to an incompatible demand, we can limit attention to the deviations that are exactly compatible with some demand of the opponent: 2/3 and 1/5. Neither player gains from such a deviation. Clearly, the payoff from making any demand less than 2/3 is strictly less than the equilibrium payoff; for demands above 2/3, the payoff is identical and hence, the rational type is willing to make the equilibrium demands. For the stubborn type, the equilibrium payoff is strictly higher than the stubborn type could get from any other demand.

Before stating the first main result formally, given a pair (α_1, α_2), with $\alpha_1 \in$

by assigning probability 1 to the deviating demand coming from a rational type, which deters the stubborn type from deviating.

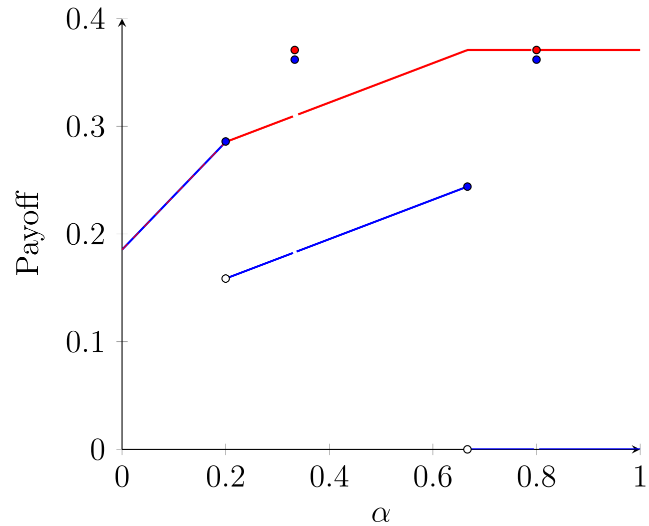

Figure 1: Expected payoff from making demand α for the rational (red) and stubborn (blue) type (pooling equilibrium). Here $z = 1/10$ and the equilibrium demands are $\alpha \in \{1/3, 4/5\}$.

$(0, 1/2)$ and $\alpha_2 \in (1 - \alpha_1, 1]$, let

$$\bar{z}(\alpha_1, \alpha_2) = \left(\frac{2(1-\alpha_1)(\alpha_1+\alpha_2-1)}{(1-2\alpha_1)(1-\alpha_2)} \frac{1-m}{m} + m^{\frac{2\alpha_2-1}{1-\alpha_2}} \right)^{-\frac{1-\alpha_2}{\alpha_2-\alpha_1}}$$

where

$$m = \min \left\{ \left(\frac{2(1-\alpha_1)(\alpha_1+\alpha_2-1)}{(1-2\alpha_1)(2\alpha_2-1)} \right)^{\frac{1-\alpha_2}{\alpha_2}}, \frac{(\alpha_2-\alpha_1)(\alpha_1+\alpha_2-1)}{(1-\alpha_2)^2+2\alpha_1\alpha_2-\alpha_1-\alpha_1^2} \right\}.$$

For $\alpha_1 = 1/2$ and $\alpha_2 > 1/2$, let $\bar{z}(1/2, \alpha_2) = 1$. Note that $\bar{z} > 0$ for all $\alpha_1 \leq 1/2$ and $1 > \alpha_2 > 1 - \alpha_1$.

Proposition 1. *1. [**Existence of Two Demand Pooling Equilibria**] Fix any $\alpha_1 < 1/2$ and $1 > \alpha_2 > 1 - \alpha_1$. There exists a pooling equilibrium (z, r, s) with finite support $\mathcal{C} = \{\alpha_1, \alpha_2\}$,*

$\forall z \in [0, \bar{z}]$.

*2. [**Convergence as stubborn types vanish**] Let (z^n, r^n, s^n) be a convergent sequence of pooling equilibria with finite support $\mathcal{C}^n = \{\alpha_1, \alpha_2\}$ and $\lim_{n\to\infty} z^n =$*

0. *Then along any such sequence,*

$$\lim_{n\to\infty} r^n = \delta_{\alpha_1}\frac{2(\alpha_1+\alpha_2-1)}{2\alpha_2-1} + \delta_{\alpha_2}\frac{1-2\alpha_1}{2\alpha_2-1},$$

and

$$\lim_{n\to\infty} s^n = \delta_{\alpha_1}\frac{1-\alpha_2}{2-\alpha_1-\alpha_2} + \delta_{\alpha_2}\frac{1-\alpha_1}{2-\alpha_1-\alpha_2}.$$

Moreover,

$$\lim_{n\to\infty} v_r^n = \lim_{n\to\infty} v_s^n = \frac{1}{2} - \frac{\left(\frac{1}{2}-\alpha_1\right)^2}{\alpha_2-\frac{1}{2}}.$$

3. *[**Convergence as rational types vanish**] Let* (z^n, r^n, s^n) *be a convergent sequence of pooling equilibria and* $\lim_{n\to\infty} z^n = 1$. *Then along any such sequence EITHER (1)*

$$\lim_{n\to\infty} r^n\left(\frac{1}{2}\right) = 1, \quad \textit{and} \quad \lim_{n\to\infty} s^n = \delta_1, \quad \textit{with} \quad \lim_{n\to\infty} v_r^n = \lim_{n\to\infty} v_s^n = 0,$$

OR (2)

$$\lim_{n\to\infty} r^n = \lim_{n\to\infty} s^n = \delta_{\frac{1}{2}}, \quad \textit{with} \quad \lim_{n\to\infty} v_r^n = \lim_{n\to\infty} v_s^n = \frac{1}{2}.$$

The first part of Proposition 1 establishes that for any fixed probability of facing a stubborn type $z \in (0,1)$, there exist pooling equilibria with two demands.[15] For instance, a pooling equilibrium, where both types randomize over $\alpha_1 = 49/100$ and $\alpha_2 = 99/100$ exists for any $z \leq 0.978$. Conversely, a pooling equilibrium, where both types randomize over $\alpha_1 = 1/3$ and $\alpha_2 = 4/5$ exists for any $z \leq 0.627$. More generally, for every $z \in (0,1)$ one can select (α_1, α_2) such that a two-demand pooling equilibrium exists.

To gain intuition for why players can be made indifferent over multiple demands, it is helpful to consider the tradeoffs involved in choosing a demand. First, players of either type face a tradeoff between the amount they receive if their opponent concedes and the speed at which the opponent concedes. Fixing the opponent's concession behavior, a player's payoff increases with their demand. However, the higher

[15] A complete characterization of all equilibria can be found in the working paper version on the author's website.

the demand, the slower the opponent concedes. This tradeoff makes intermediate demands particularly appealing and results in a rational payoff that is single-peaked in their own demand, as in AG.

Second, these tradeoffs vary between the two types. When demands are compatible, both types receive the same payoff. When demands are incompatible, however, the rational type's expected payoff is higher than that of the stubborn type. This is because, unlike the stubborn type, the rational player can choose to concede. A rational player is willing to wait as long as there is uncertainty regarding the opponent's type. But once the rational player assigns probability 1 to facing a stubborn opponent, the rational player strictly prefer to concede, whereas the stubborn type cannot concede. This results in the rational type achieving a higher (expected) payoff than the stubborn type when demands are incompatible.

The difference in payoffs when facing an incompatible demand depends on the level of the demands. Specifically, this difference is smaller for higher demands. Higher demands lead to a slower concession rate, prolonging the war of attrition. Consequently, the point at which the rational player strictly prefers to concede is pushed "far into the future." Due to discounting, the stubborn type's cost of being unable to concede becomes minimal when demands are high. Therefore, when demands are incompatible, the payoff difference between the two types is smaller the higher the demands. As a result, preferences do not satisfy the single-crossing property.

Finally, the appropriate off-path beliefs ensure that neither type has an incentive to deviate to out-of-equilibrium demands. Specifically, assigning sufficiently high probability to any deviation coming from the rational type ensures that both types find it optimal to stick to the equilibrium demands.[16] As in AG, off-path beliefs are

[16]Specifically, in a pooling equilibrium with support $\mathcal{C} = \{\alpha_1, \alpha_2\}$, requiring $\pi(\alpha) = \pi(\alpha_2)^{(1-\alpha)/(1-\alpha_2)}$ is sufficient if α is sufficiently low, i.e., if:

$$q_1 \frac{\alpha + 1 - \alpha_1}{2} + (1-q_1)(1-\alpha_2)(1-\mu_2^{\alpha}) \leq q_1 \frac{\alpha_2 + 1 - \alpha_1}{2} + (1-q_1)(1-\alpha_2)(1-\mu_2^{\alpha_2}).$$

If this condition is violated, we specify that $\pi(\alpha) = 0$, which deters the deviation. Note that in AG, any deviation is automatically assigned probability 1 as coming from the rational type, as stubborn central to disciplining deviations but play an additional role in my model. In their framework, any off-path demand is automatically attributed to the rational type, a reasonable specification given that committed types cannot deviate. This sustains a unique outcome by forcing play onto the set of commitment demands. In the present model, committed types also choose their initial demands: this means there are more deviations (by the stubborn type) that need to be deterred (a force towards unique predictions).[17] It is here that belief-based punishments play a crucial role: they are needed to "force" behavior by the stubborn type onto the equilibrium path. So, payoff multiplicity emerges if sufficiently high off-path demands are punished by assigning probability one to the rational type. If, instead, such dogmatic beliefs are ruled out and all deviations must be assigned positive probability of stemming from a committed type, then the only symmetric pooling equilibrium is the efficient outcome where all types demand 1/2. Thus, the Folk theorem result hinges on the scope for extreme off-path punishments, in contrast to AG where such punishments collapse behavior to a unique outcome.

The second part of Proposition 1 demonstrates the convergence of strategies and payoffs as the probability of facing a stubborn type vanishes. It shows that any feasible payoff can be sustained in equilibrium for either type, when the probability of facing a stubborn type is small enough. Here, inefficiency is measured by the distance between 1/2 and the lower demand α_1, as well as the distance between α_2 and 1. When α_1 is close to 0 (and hence, α_2 close to 1), a player's expected equilibrium payoff is close to 0. A demand α_2 close to 1 implies that a player almost certainly will face a demand of α_2 which induces a long war of attrition. If, on the other hand, α_1 is close to 1/2, a player's expected payoff is close to 1/2 (when players are equally patient). When demands are close to 1/2, the war of attrition

types have no choices to make.

[17]As discussed previously, there is also a second implication: conditional on being able to appropriately incentivize the stubborn type, there is greater flexibility on which demands can be assigned positive probability as coming from a behavioral type *on path*. This means it is easier to incentivize the rational type to be willing to make certain demands.

is short, minimizing inefficiency. By adjusting α_1 and α_2, one can generate in this fashion any payoff between 0 and 1/2. Interestingly, fixing α_1, a higher α_2 actually *increases* the limiting equilibrium payoff. Although this may seem counterintuitive – since a symmetric equilibrium with the highest payoff occurs when both types demand 1/2 with probability 1 – a rational player facing a demand of α_2 receives a payoff of $1 - \alpha_2 < 1/2$. Hence, conditional on meeting a demand of α_2, the rational type's payoff increases as α_2 decreases. However, there is a dominating effect: as α_2 increases, the likelihood of the rational type demanding α_2 decreases.

The third part of Proposition 1 considers the case where the probability of facing a stubborn type is high. In such cases, any pooling equilibrium, regardless of the number of demands in the support of the strategy, leads to one of two outcomes: either immediate agreement with compatible demands, or prolonged delays that ultimately result in a payoff of zero for both types as $z \to 1$. To illustrate the logic, return to the example of demands $\alpha_1 = 1/3$ and $\alpha_2 = 4/5$. When $z = 0.9$, players can still be made indifferent between these two demands. However, indifference alone is not sufficient for equilibrium: at such high probabilities of stubbornness, the cost of a prolonged standoff becomes substantial for stubborn types. As a result, a stubborn type has an incentive to deviate to a compatible demand (such as 1/5) to avoid the risk of deadlock with another stubborn opponent. This is why "moderate" demands can no longer be sustained in equilibrium as z becomes large, even if indifference across them can still be maintained.

Let me now illustrate the structure of the separating equilibria. Suppose the prior probability of the stubborn type is $z = 1/2$. Then the following is an equilibrium: the stubborn type demands 1/3 and the rational type demands 4/5.[18] In other words, the stubborn type makes a less aggressive demand than the rational type. After the initial demands are made, a player demanding 1/3 is believed to be stubborn with probability 1 and a player demanding 4/5 is believed to be rational with probability 1. If either player demands $\alpha \notin \{1/3, 4/5\}$, the player is believed to be rational

[18]There exists a separating equilibrium with these demands for $z \in [4/9, 2/3]$.

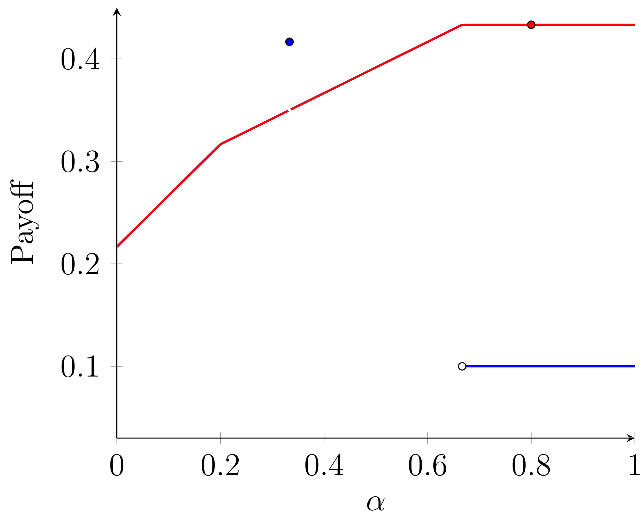

Figure 2: Expected payoff from making demand α for the rational (red) and stubborn (blue) type (separating equilibrium). Here $z = 1/2$, and the equilibrium demands are $1/3$ (for the stubborn type) and $4/5$ (for the rational type).

with probability 1. After the initial demands, the game unfolds as follows. If both players demand $1/3$, there is immediate agreement and both players receive $1/2$. If one player demands $1/3$, and the other $4/5$, the player demanding $4/5$ immediately concedes with probability 1. If both players demand $4/5$, then a war of attrition starts, giving both (rational) players a payoff of $1/5$.[19] Fixing the opponent's strategy, a rational player receives an expected payoff of 0.4333 from the equilibrium strategy, and the stubborn type receives 0.4167. Figure 2 illustrates the expected payoff from an equilibrium demand (as before shown as red and blue dots for the rational and stubborn type respectively) and contrasts it to the payoff a rational (stubborn) player could receive from deviating to any other demand. Note that so long as the demand is compatible with the stubborn type's demand (i.e., demands up to $2/3$ for the demands chosen in Figure 2), the payoff to a rational and a stubborn type from a

[19]Note that in the continuation game involving two rational players, there exists another equilibrium with immediate concession. However, any such continuation strategies would make it profitable for the stubborn type to imitate the rational type and hence, cannot be part of an equilibrium.

deviation is identical.

As we will see, the feature that the stubborn type's demand is below the rational type's demand is true in any separating equilibrium: stubborn types are more conservative in their demands compared to rational types, because it is more costly to them if demands are incompatible.

Proposition 2. 1. *[**Existence of Separating Equilibria**] Fix any set of demands $\{\alpha_1, \dots, \alpha_K\}$ with $K \geq 2$. Then there exists a fully separating equilibrium if and only if $\alpha_1 < 1/2$, $\alpha_1 + \alpha_2 > 1$ and*

$$z \in \left[\frac{2(\alpha_1 - \sum_{k=2}^{K} r_k(1-\alpha_k))}{1 - 2\sum_{k=2}^{K} r_k(1-\alpha_i)}, 2\alpha_1 \right],$$

where $(r_2, \dots, r_K)$ are weights s.t. $\sum_{k=2}^{K} r_i = 1$. Moreover, in any such separating equilibrium, $s = \delta_{\alpha_1}$ and $r = \sum_{k=2}^{K} \delta_{\alpha_k} r_i$.

2. *[**Convergence as stubborn types vanish**] Let (z^n, r^n, s^n) be a convergent sequence of separating equilibria and $\lim_{n\to\infty} z^n = 0$. Then there exists $\alpha \in (0, 1/2)$ such that $\lim_{n\to\infty} r^n = \delta_{1-\alpha}$ and $\lim_{n\to\infty} s^n = \delta_\alpha$. Moreover,*

$$\lim_{n\to\infty} v_r^n = \lim_{n\to\infty} v_s^n = \alpha.$$

Conversely, for every $\alpha \in (0, 1/2)$, there exists a convergent sequence of separating equilibria (z^n, r^n, s^n) with $\lim_{n\to\infty} r^n = \delta_{1-\alpha}$ and $\lim_{n\to\infty} s^n = \delta_\alpha$. Moreover,

$$\lim_{n\to\infty} v_r^n = \lim_{n\to\infty} v_s^n = \alpha.$$

3. *[**Convergence as rational types vanish**] Let (z^n, r^n, s^n) be a convergent sequence of separating equilibria and $\lim_{n\to\infty} z^n = 1$. Then along any such sequence, $\lim_{n\to\infty} s^n = \delta_{1/2}$. Moreover, $\lim_{n\to\infty} r^n(1/2) = 0$. Along any such sequence,*

$$\lim_{n\to\infty} v_r^n = \lim_{n\to\infty} v_s^n = \frac{1}{2}.$$

The first part of Proposition 2 establishes existence of separating equilibria for any fixed ex ante probability of facing a stubborn type. This result is stronger than the first part of Proposition 1 as it provides both necessary and sufficient conditions for the existence of any separating equilibrium. It is straightforward to verify that for every $z \in (0, 1)$, one can select demands $\{\alpha_1, \alpha_2\}$ such that a separating equilibrium exists.[20] In any such equilibrium, if one player is stubborn, agreement is immediate. In other words, here, delay signals rationality rather than stubbornness: delay only occurs when two rational players face one another, which contrasts with the dynamics of pooling equilibria.[21]

The second part of Proposition 2 derives the limits of strategies and payoffs as the probability of facing a stubborn type vanishes. In the limit, the two types make exactly compatible demands, and in this way, any feasible payoff can be sustained in equilibrium for either type as the probability of facing a stubborn type vanishes.

Finally, the third part of Proposition 2 addresses convergence as the probability of facing a rational type approaches zero. Here, the stubborn type makes a demand of 1/2, resulting in a payoff of 1/2 for both types. The intuition behind this result is straightforward: as the probability of facing a stubborn type increases, so does the stubborn type's incentive to deviate from its separating demand to a complementary demand. This incentive diminishes as the stubborn type's separating demand increases, ultimately leading to the equilibrium outcome.

Taken together, Propositions 1 and 2 show that payoff multiplicity arises in sym-

[20]To see this, note that the lower bound converges to 0 as $\alpha_1 + \alpha_2 \to 1$ and the upper bound converges to 1 if $\alpha_2 \to 1$.

[21]Modeling the bargaining as a war of attrition entails some loss of generality here: if rational players could revise their demands, there exists another equilibrium of the continuation game which does not entail delay: if the incompatible demands revealing rationality are observed by both players, players revise their offers to 1/2. If one of the players does not revise their demand, probability 1 is placed on this deviation coming from a rational type and hence, not revising one's offer from the original incompatible offer does not lead to concession by the opponent – thereby ensuring the stubborn type has no incentive to deviate from his demand.

metric pooling and separating equilibria when the probability of facing a stubborn type is small, but breaks down as this probability becomes large. More generally, as the probability of the rational type vanishes, the game reduces to a Nash demand game among stubborn players. In any symmetric equilibrium – whether pooling, separating, or semi-separating – an argument similar to that in the proof of Proposition 2(c) implies that the limiting payoffs converge to $(1/2, 1/2)$ (efficient agreement) or $(0, 0)$ (complete breakdown).[22] By contrast, when asymmetric equilibria are considered, any division of the surplus between the two players can be sustained in the limit.

The preceding propositions do not exhaustively characterize equilibria, even under symmetry and finite support. In particular, there exist pooling equilibria supported on more than two demands for fixed z, as well as a variety of semi-separating equilibria. I restrict attention to the pooling and fully separating cases, as it suffices to establish payoff multiplicity. Allowing for asymmetric equilibria leads to a still richer set of outcomes. To see this, note that for any $\alpha \in (0, 1)$ there exists a pooling equilibrium in which player 1 demands α and player 2 demands $1-\alpha$, yielding asymmetric payoffs $(\alpha, 1-\alpha)$. Hence, an asymmetric Folk theorem holds: any division of the surplus between players can be sustained in equilibrium (for any $z \in (0, 1)$).

4 Type space

This section locates a boundary of the multiplicity result: once an AG type with $\alpha = 1/2$ is present (in addition to the stubborn type), the scope for payoff multiplicity disappears in the vanishing-behavioral limit, and equilibrium payoffs collapse to the efficient payoff 1/2 for all players. However, away from that vanishing-behavioral limit – or if the support of AG types does not include $\alpha = 1/2$ – equilibrium multiplicity

[22] Essentially, the stubborn type will always prefer to make a compatible demand in this case (unless making a compatible demand gives 0 payoff). Hence, within the class of symmetric equilibria, the stubborn type either demands 1/2 or 1. Hence, both types payoffs are either 1/2 or 0 in the limit.

can persist.

Specifically, consider a model with a (finite) set of AG types $\mathcal{C}^{\mathcal{AG}}$ with one simple change: each player is behavioral à la AG with probability z_{AG}, stubborn with probability z_S and rational otherwise. I denote by $\pi_{AG}(\alpha)$ the conditional probability that a player is an AG type α given that the player is an AG type. Hence, π_{AG} is a probability distribution on $\mathcal{C}^{\mathcal{AG}}$. Denote this game $B^{AG,S}$. For any comparison between $B^{AG,S}$ and B, I assume $z_{AG} + z_S = z$.

Proposition 3. 1. *Fix an equilibrium* (z, r, s) *with finite support* $\mathcal{C} = \{\alpha_1, \ldots, \alpha_K\}$ *in* B*. Then there exists an outcome equivalent equilibrium in the game* $B^{AG,S}$ *if and only if* $z_{AG}\pi_{AG}(\alpha_k) \leq zs_k$ *for all* $\alpha_k \in \mathcal{C}$*, and* $\pi_{AG}(\alpha) = 0$ *for any* $\alpha \notin \mathcal{C}$*.*

2. *Suppose* $1/2 \in \mathcal{C}^{\mathcal{AG}}$*. Let* $(z^n_{AG}, z^n_S, r^n, s^n)$ *be a convergent sequence of pooling equilibria and* $\lim_{n\to\infty}(z^n_{AG} + z^n_S) = 0$*. Then* $\lim_{n\to\infty} r^n = \delta_{1/2}$*. Moreover,* $\lim_{n\to\infty} v^n_r = \lim_{n\to\infty} v^n_s = 1/2$*.*

The first part of Proposition 3 implies that there is multiplicity away from the limit, even when there is an AG type demanding $\alpha = 1/2$: for an equilibrium in B there exists an outcome-equivalent equilibrium in $B^{AG,S}$ exactly when the joint distribution of behavioral types – AG types and stubborn types – can reproduce the on-path posteriors in the equilibrium in B. Intuitively, stubborn types act like "flexible" AG types whose demand can be allocated to match the baseline; feasibility therefore holds provided that, at every on-path demand α, the exogenous AG mass at α does not exceed the behavioral mass that the baseline equilibrium in B assigns to α. When an AG type with $\alpha = 1/2$ is present and z_{AG} is sufficiently small relative to z_S, this condition is met, so the posteriors can be replicated and the payoff multiplicity with $\alpha = 1/2$ persists away from the vanishing-behavioral limit. The second part of the proposition shows that this multiplicity collapses to the efficient payoff 1/2 whenever an AG type $\alpha = 1/2$ is present and the probability of facing a behavioral type (of either kind) vanishes.

Corollary 1. *Suppose $\mathcal{C}^{\mathcal{AG}} = \{\alpha_2\}$ with $\alpha_2 > 1/2$ and suppose $z_{AG} = z_S = z$. Then there exists $\bar{z} > 0$ such that for every $z < \bar{z}$ and every α_1 satisfying $\alpha_1 + \alpha_2 > 1$ and $\alpha_1 < 1/2$, there exists a pooling equilibrium with support $\{\alpha_1, \alpha_2\}$. Let (z^n, r^n, s^n) be a convergent sequence of pooling equilibria with support $\{\alpha_1, \alpha_2\}$ and $\lim_{n\to\infty} z^n = 0$. Then*

$$\lim_{n\to\infty} v_r^n = \lim_{n\to\infty} v_s^n = \frac{1}{2} - \frac{\left(\frac{1}{2} - \alpha_1\right)^2}{\alpha_2 - \frac{1}{2}} > 1 - \alpha_2.$$

The corollary establishes payoff multiplicity when there is a single AG type $\alpha > 1/2$ and the probability of facing a behavioral type (of either kind) is small.

Note that there exist equilibria in $B^{AG,S}$ where no outcome equivalent equilibrium exists in B: this is precisely because there are fewer incentive compatibility constraints to satisfy in $B^{AG,S}$ than in B.[23]

5 Varying abilities to concede

Many real-world negotiations, particularly in political and labor disputes, involve parties who are uncertain about their own ability to hold out. For instance, a labor union might be unsure about its members' commitment to continue striking, or a political negotiator may believe they can concede but remain uncertain about whether they can truly afford to do so. This uncertainty highlights the value of exploring a generalized version of the model presented in this paper. Specifically, I aim to capture the idea that players may have differing abilities to concede. Initially, a player may only know the probability with which they can concede, discovering their true capacity for concession only when they attempt it. This idea can be formalized in the following model.

Each player is either rational or stubborn. Ex ante, each player i privately observes a signal $\theta_i \in \Theta$, where $\Theta \subseteq [0,1]$ is finite. Conditional on $\theta_i = \theta$, player i is stubborn with probability θ (and rational with probability $1-\theta$). Signals are drawn i.i.d. from the discrete distribution $\bar{P}$ on Θ with masses $(\bar{p}_\theta)_{\theta\in\Theta}$, i.e. $\bar{P} = \sum_{\theta\in\Theta} \bar{p}_\theta \, \delta_\theta$, so that $\sum_{\theta\in\Theta} \bar{p}_\theta = 1$.

[23] Examples of such equilibria are readily available on request.

As before, the solution concept is weak PBE. Recall that PBE imposes no restriction on off-path beliefs: regardless of the mass points of $\bar{P}$, we can assign arbitrary beliefs to any off-path demand. Given the signal θ, a player i chooses a demand $\alpha^i \in [0,1]$ (as before players choose demands simultaneously). Once demands are made, players privately learn their type (rational or stubborn). The war of attrition ensues as before. The bargaining problem is denoted $\bar{B} = \{z, \Theta, \bar{P}, \rho\}$. The model as described in Section 2 is the special case, where $\bar{P} = P := \delta_1 z + \delta_0(1-z)$. As before, the unconditional probability of stubbornness is $z := \sum_{\theta \in \Theta} \bar{p}_\theta \, \theta$.[24]

A strategy for player i who has received signal θ is defined by a Borel probability measure $\sigma^{i,\theta}$ on $[0,1]$ and a collection of cumulative distributions $F^{i,\theta}_{\alpha_i,\alpha_j}$ on $\mathcal{R}_+ \cup \{\infty\}$ for all $\alpha_i + \alpha_j > 1$. $F^{i,\theta}_{\alpha_i,\alpha_j}(t)$ describes the probability that the rational player i who received signal θ concedes to player j by time t (inclusive), given his choice of α_i, when facing α_j. With some abuse of notation, I will denote the strategy for player i who has received signal θ by $(\sigma^{i,\theta}, F^{i,\theta}_{\alpha_i,\alpha_j})$.

We may additionally want to impose the (very reasonable) "no signaling what you don't know" (NSWYDK) condition. If we do so, the set of symmetric pooling equilibria depends on the mass points of $\bar{P}$. Specifically, if $\bar{p}_0 > 0$, everything stated in this section continues to hold. This is because a small positive mass on $\theta = 0$ ensures that players can be punished with beliefs for any off-path deviation by being identified as a rational $\theta = 0$ type in a manner consistent with NSWYDK. If instead $\bar{p}_0 = 0$, there is a unique symmetric pooling equilibrium satisfying the NSWYDK condition. In this equilibrium, all players demand $1/2$. If we impose NSWYDK and $\bar{p}_0 = 0$, this implies there is no off-equilibrium path demand which is believed to come from a rational type for sure. This makes certain deviations more attractive and therefore rules out multi-demand pooling equilibria.[25]

[24] Note that a player who only discovers their ability to concede upon attempting to do so may as well condition on the event that he is able to concede when choosing his concession strategy. Therefore, this model – where a player learns his type after demands are made – is strategically equivalent to the informal description provided above.

[25] To see this, note that in any pooling equilibrium with support $\{\alpha_1, ..., \alpha_K\}$, the payoff to the

Note that given a strategy profile, each on path demand induces a posterior probability distribution over types: each demand made can be viewed as a signal about a player's type. Throughout, I denote the (distribution over) posterior beliefs that a strategy profile induces by $\bar{\Pi}$.[26] Consider the two bargaining games, B with (Θ, P) and $\bar{B}$ with $(\Theta, \bar{P})$. Note that $\bar{P}$ is a garbling of P. Hence, ignoring incentive compatibility constraints, any probability distribution over types that can be generated with $\bar{P}$, can be generated with P. Given that payoffs are linear in θ, this implies that for any pooling equilibrium in $\bar{B}$ there exists an outcome-equivalent pooling equilibrium in B. The reverse is true only when $\bar{P}$ second-order stochastically dominates the probability distribution over types induced in the equilibrium of B, Π.[27]

Proposition 4. *Fix $\bar{B} = \{z, \Theta, \bar{P}, \rho\}$ and $B = \{z, \Theta, P, \rho\}$. Suppose $((r, F^r_{\alpha_i,\alpha_j}), s)$ is an equilibrium of B, which induces beliefs Π. Then there exists an outcome-equivalent equilibrium of $\bar{B}$ if, and only if, $\bar{P} \succsim_{SOSD} \Pi$.*

Proposition 4 highlights that many insights from the paper carry over to this model. However, it also clearly indicates where this similarity reaches its limits: if players in $\bar{B}$ separate (that is, different types choose different demands), the strategy profile in B that would generate the same information structure is generally not incentive-compatible. Clearly, if a player receiving signal $\theta = 0$ is willing to demand α, and a player receiving signal $\theta = 1$ is willing to demand α, then a player receiving signal $\theta \in (0, 1)$ is willing to demand α. The reverse however is not true: a player receiving signal θ may be willing to demand α, but a player receiving a fully informative signal does not. This implies, that when players in $\bar{B}$ separate by signal, there does not generally exist an outcome-equivalent equilibrium in B.[28]

stubborn type is given by $\sum_k q_k(1-\alpha_k)(1-\mu_k^{\alpha_K})$. By deviating to $\alpha > \alpha_K$, the stubborn type could guarantee a strictly higher payoff: $\sum_k q_k(1-\alpha_k)(1-\mu_k^{\alpha})$.

[26]In the simplest case, take $P = \delta_0 0.9 + \delta_1 0.1$, and revisit the pooling equilibrium with 1/3 and 4/5. Then $\bar{\Pi} = \Pi = \delta_{0.063}(0.9 * 0.381 + 0.1 * 0.23) + \delta_{0.121}(0.9 * 0.619 + 0.1 * 0.77)$.

[27]Examples that illustrate the key insight on equivalence and when it breaks down are available on request.

[28]Examples of such separating equilibria are readily available on request.

6 Conclusion

While this paper focuses on endogenizing behavioral types in a bargaining setting, the idea of endogenizing behavioral types applies more broadly. For instance, some agents may restrict attention to stationary strategies in a repeated game. Whatever drives their preference for this restriction does not mean that they do not choose optimally within the set of stationary strategies. There is a middle ground between rational and behavioral agents, and this paper is a first attempt to explore this territory in a well-known and tractable environment.

Appendix

For the proof of Proposition 1, it is helpful to establish some preliminary lemmas that impose some structure on possible equilibria. In a candidate pooling equilibrium, the payoff to the rational type of player j from demanding α_ℓ is:

$$\begin{aligned} v_r^j(\alpha_\ell) = &\sum_{k:\alpha_k \le 1-\alpha_\ell} q_k \frac{1-\alpha_k+\alpha_\ell}{2} \\ &+ \sum_{k:\alpha_k > 1-\alpha_\ell} q_k \left(\alpha_\ell - (\alpha_k + \alpha_\ell - 1) \min\left\{ \left(\frac{\mu_k}{\mu_\ell}\right)^{1-\alpha_k}, 1 \right\} \right). \end{aligned} \tag{1}$$

Similarly, we can write the payoff of a stubborn player j demanding α_ℓ in a candidate pooling equilibrium as:

$$v_s^j(\alpha_\ell) = v_r^j(\alpha_\ell) - \sum_{k:\ \alpha_k > 1-\alpha_\ell} q_k(1-\alpha_k)\mu_k^{\alpha_\ell} \max\left\{1, \left(\frac{\mu_\ell}{\mu_k}\right)^{\alpha_k+\alpha_\ell-1}\right\}. \tag{2}$$

Using (1),(2), given $z > 0$, a pooling equilibrium with support $\mathcal{C} = \{\alpha_1, \dots, \alpha_K\}$ requires, $\forall \alpha_\ell, \alpha_m \in \mathcal{C}$, and $j = 1, 2$,

$$v_r^j(\alpha_\ell) - v_r^j(\alpha_m) = 0, \tag{3}$$

$$v_s^j(\alpha_\ell) - v_s^j(\alpha_m) = 0, \tag{4}$$

$$\sum_{k=1}^{K} q_k = 1, \text{ and} \tag{5}$$

$$\sum_{k=1}^{K} q_k \mu_k^{1-\alpha_k} = z, \tag{6}$$

with $q_k, \mu_k \in [0, 1]$.

Lemma 1. *Fix any set of demands $\mathcal{C}$, where $\mathcal{C} = \{\alpha_1, \dots, \alpha_K\}$ is an arbitrary finite subset of $[0, 1]$. In any symmetric pooling equilibrium with support $\mathcal{C}$, $\mu_k > \mu_{k+1}$ for any $\alpha_k < 1 - \alpha_1$ and $\mu_k = \mu_{k+1}$ for $\alpha_k \ge 1 - \alpha_1$.*

Proof of Lemma 1. Note that for any two demands $\alpha_k, \alpha_{k+1} \in \mathcal{C}$, (3) must hold. Conditional on facing a compatible demand, the payoff to a rational player from demanding α_{k+1} is strictly higher than the payoff from demanding α_k (cf. (3)). Conditional on facing a demand which is compatible with α_k but not with α_{k+1}, the payoff to a rational player from demanding α_k is weakly higher than the payoff from demanding α_{k+1}. Hence, either (a) $\forall \alpha_m \in \mathcal{C}$, $\alpha_m + \alpha_k \geq 1$ and $\mu_m \geq \min\{\mu_{k+1}, \mu_k\}$, or (b) there must exist some $\alpha_n \in \mathcal{C}$ with $\alpha_k + \alpha_n > 1$ such that $\mu_{k+1} < \mu_n$ but $\mu_k > \mu_n$. Note that (a) implies $\mu_k = \mu_{k+1}$. Moreover, (b) implies $\mu_{k+1} < \mu_k$. The result follows. □

Lemma 2. *Fix any set of demands $\mathcal{C}$, where $\mathcal{C} = \{\alpha_1, \ldots, \alpha_K\}$ is an arbitrary finite subset of $[0,1]$ and $|\mathcal{C}| \geq 2$. In any symmetric pooling equilibrium with support $\mathcal{C}$, the following holds:*

1. *The lowest and highest demand in $\mathcal{C}$ are incompatible: $\alpha_1 + \alpha_K > 1$.*
2. *Consider any two demands, $\alpha_\ell, \alpha_m \in \mathcal{C}$, $\alpha_\ell > \alpha_m$. Then there exists $\alpha_n \in \mathcal{C}$ s.t. $\alpha_\ell + \alpha_n > 1$ and $\alpha_m + \alpha_n \leq 1$.*

Proof of Lemma 2. Part 1. Suppose $\alpha_1 + \alpha_K \leq 1$. Then $v_r(\alpha_1) = v_s(\alpha_1)$. This implies $v_r(\alpha_k) = v_s(\alpha_k)$ $\forall \alpha_k \in \mathcal{C}$. If $v_r(\alpha_k) = v_s(\alpha_k)$ $\forall \alpha_k \in \mathcal{C}$, then it must be that $\alpha_K \leq 1/2$. If $\alpha_K \leq 1/2$, then $\alpha_K = \alpha_1 = 1/2$, and hence, $|\mathcal{C}| = 1$. Hence, if $|\mathcal{C}| \geq 2$, $\alpha_1 + \alpha_K > 1$.

Part 2. Consider any two demands, $\alpha_\ell, \alpha_m \in \mathcal{C}$, $\alpha_\ell > \alpha_m$. Suppose $\forall \alpha_n \in \mathcal{C}$, either $\alpha_\ell + \alpha_n \leq 1$ or $\alpha_m + \alpha_n > 1$ (i.e., the set of compatible demands is constant between α_ℓ and α_m). Given that $\mu_\ell \leq \mu_m$, for all $\alpha_n \geq 1 - \alpha_m$:

$$\mu_n^{\alpha_\ell} \max\left\{1, \left(\frac{\mu_\ell}{\mu_n}\right)^{\alpha_n + \alpha_\ell - 1}\right\} < \mu_n^{\alpha_m} \max\left\{1, \left(\frac{\mu_m}{\mu_n}\right)^{\alpha_n + \alpha_m - 1}\right\}.$$

Hence, if (3) is satisfied for all $\alpha_\ell, \alpha_m \in \mathcal{C}$, then (4) cannot be satisfied.

Hence, for all $\alpha_\ell, \alpha_m \in \mathcal{C}$ with $\alpha_\ell > \alpha_m$, there exists $\alpha_n \in \mathcal{C}$ such that $\alpha_m \leq 1 - \alpha_n < \alpha_\ell$. (Note that it follows that the lowest pooling demand is compatible with all but the highest pooling demand.)

□

Proof of Proposition 1. The proof procedes in three steps.

Part 1: Existence of Two Demand Pooling Equilibria for fixed z. The next lemma implies existence of pooling equilibria with two demands for fixed z.

For $\alpha_1 < 1/2$, $\alpha_2 > 1 - \alpha_1$, let

$$\bar{z}(\alpha_1, \alpha_2) = \left(\frac{2(1-\alpha_1)}{n(1-\alpha_2)} \frac{1-m}{m} + m^{\frac{2\alpha_2 - 1}{1-\alpha_2}} \right)^{-\frac{1-\alpha_2}{\alpha_2 - \alpha_1}}$$

where

$$n = \frac{1 - 2\alpha_1}{\alpha_1 + \alpha_2 - 1}, \text{ and}$$
$$m = \min \left\{ \left(\frac{2(1-\alpha_1)(\alpha_1 + \alpha_2 - 1)}{(1-2\alpha_1)(2\alpha_2 - 1)} \right)^{\frac{1-\alpha_2}{\alpha_2}}, \frac{(\alpha_2 - \alpha_1)(\alpha_1 + \alpha_2 - 1)}{(1-\alpha_2)^2 + 2\alpha_1\alpha_2 - \alpha_1 - \alpha_1^2} \right\}.$$

The bound $\bar{z}(\alpha_1, \alpha_2)$ is well-defined for $\alpha_1 < 1/2$. The right-limit exists and equals one,

$$\lim_{\alpha_1 \uparrow 1/2} \bar{z} = 1,$$

so by continuity I define $\bar{z}(1/2, \alpha_2) := 1$.

Moreover, given z and $\mathcal{C} = \{\alpha_1, \alpha_2\}$, define the following system in (q_k, μ_k), $k = 1, 2$:

$$q_1 \left(\alpha_1 - \frac{1}{2} \right) + q_2 \left(\alpha_1 + \alpha_2 - 1 \right) \left(1 - \left(\frac{\mu_2}{\mu_1} \right)^{1-\alpha_2} \right) = 0, \text{ and} \tag{7}$$

$$q_1 \left(1 - \alpha_1 \right) \mu_1^{\alpha_2} - q_2 \left(1 - \alpha_2 \right) \mu_2^{1-\alpha_2} \left(\mu_1^{\alpha_1 + \alpha_2 - 1} - \mu_2^{2\alpha_2 - 1} \right) = 0, \tag{8}$$

$$\sum_{k=1}^{2} q_k \mu_k^{1-\alpha_k} = z, \tag{9}$$

$$q_1 + q_2 = 1, \text{ and} \tag{10}$$

$$\begin{aligned} &q_1 \left(1 - \alpha_1 \right) \left(1 - \mu_1^{\alpha_2} \right) + q_2 \left(1 - \alpha_2 \right) \left(1 - \mu_2^{\alpha_2} \right) \\ &\geq q_1 \mathbb{1}_{\alpha \leq 1-\alpha_1} \frac{1 - \alpha_1 + \alpha}{2} + q_2 \mathbb{1}_{\alpha < 1-\alpha_2} \frac{1 - \alpha_2 + \alpha}{2}, \ \forall \alpha \notin \mathcal{C}, \end{aligned} \tag{11}$$

$$q_1 \left(1 - \alpha_1 \right) + q_2 \left(1 - \alpha_2 \right) \geq \sum_{k=1,2} q_k \min\{ \frac{1 - \alpha_k + \alpha}{2}, 1 - \alpha_k \} \ \forall \alpha \notin \mathcal{C}, \tag{12}$$

where $0 < q_k < 1$, $0 < \mu_k < 1$.[29]

[29](7)–(10) corresponds to (3)–(6) when $K = 2$ (necessary conditions for existence). (11) and (12)

Lemma 3. *Fix any $\alpha_1 \leq 1/2$ and $1 > \alpha_2 > 1 - 1\alpha_1$. For every $z < \bar{z}$, there exist probabilities q_1 and q_2, and positive numbers μ_1 and μ_2 with $\mu_1, \mu_2 \leq 1$ that solve (7)–(12).*

Proof. There exists a pooling equilibrium with support $\{\alpha_1, \alpha_2\}$ only if the demands α_1 and α_2 along with probabilities q_1 and q_2, and positive numbers μ_1 and μ_2 with $\mu_1, \mu_2 \leq 1$ solve (7)– (12). Note that clearly (12) is satisfied for any $\alpha \notin \mathcal{C}$. Moreover, note that by Lemma 2, it must be that $\alpha_1 + \alpha_2 > 1, 0 < \alpha_1 \leq 1/2 < \alpha_2 \leq 1$. If $\alpha_2 = 1$, (8) can only be satisfied if $q_2 = 1$, hence, players would not pool over two demands. So, it is assumed that $\alpha_2 < 1$.

We can simplify (11) to

$$q_1 (1 - \alpha_1) (1 - \mu_1^{\alpha_2}) + q_2 (1 - \alpha_2) (1 - \mu_2^{\alpha_2}) \geq \max\{q_1 (1 - \alpha_1), q_1 \left(\frac{2 - \alpha_1 - \alpha_2}{2}\right) + q_2(1 - \alpha_2)\}. \tag{13}$$

In the following, we reduce the system (7)–(10) and (13) to one equation and one inequality, in one unknown μ_1. Then we derive an explicit upper bound on z such that a solution to this system exists. First, we use (10) to replace q_2 by $1 - q_1$ in (7), which we can then solve for q_1:

$$q_1 = \frac{2(\alpha_1 + \alpha_2 - 1)\left(1 - \left(\frac{\mu_2}{\mu_1}\right)^{1-\alpha_2}\right)}{(2\alpha_2 - 1) - 2(\alpha_1 + \alpha_2 - 1)\left(\frac{\mu_2}{\mu_1}\right)^{1-\alpha_2}}. \tag{14}$$

Using these expressions for q_1 and q_2, we can then solve (9) for μ_2:

$$\mu_2 = \left(\frac{2(\alpha_1 + \alpha_2 - 1)\mu_1^{1-\alpha_1} - z(2\alpha_2 - 1)}{2(\alpha_1 + \alpha_2 - 1)\left(\mu_1^{1-\alpha_1} - z\right) - (1 - 2\alpha_1)\mu_1^{1-\alpha_2}}\right)^{\frac{1}{1-\alpha_2}} \mu_1. \tag{15}$$

Then, replacing q_1, q_2 and μ_2 in (8), we can write the stubborn type's indifference as

ensure that neither type has any incentive to deviate to an out-of-equilibrium demand (necessary and sufficient conditions for existence).

a function of μ_1 only:

$$(1-\alpha_2)\left(2(-1+\alpha_1+\alpha_2)\mu_1^{1-\alpha_1}+z-2\alpha_2 z\right)$$
$$\cdot\left(1-\mu_1^{\alpha_2-\alpha_1}\left(\frac{2(-1+\alpha_1+\alpha_2)\mu_1^{1-\alpha_1}+(z-2\alpha_2 z)}{(-1+2\alpha_1)\mu_1^{1-\alpha_2}+2(-1+\alpha_1+\alpha_2)(\mu_1^{1-\alpha_1}-z)}\right)^{\frac{2\alpha_2-1}{1-\alpha_2}}\right) \tag{16}$$
$$+2(1-\alpha_1)(-1+\alpha_1+\alpha_2)\mu_1^{\alpha_2-\alpha_1}(\mu_1^{1-\alpha_2}-z)=0.$$

Finally, by replacing q_1, q_2 and μ_2, and further simplifying, using (16), we can write (13) as:

$$z \leq \frac{(\alpha_1+\alpha_2-1)\mu_1^{1-\alpha_2}\left(\alpha_2-\alpha_1+2(1-\alpha_2)\mu_1^{\alpha_2}\right)}{(\alpha_2-\alpha_1)\left(\alpha_1+\alpha_2-1\right)+(1-\alpha_2)(2\alpha_2-1)\mu_1^{\alpha_1}}. \tag{17}$$

Recall that $\mu_2 < \mu_1$ by Lemma 1. Hence, it follows from (15) that

$$\mu_1^{1-\alpha_2} < z. \tag{18}$$

Moreover, recall that $q_1 < 1$. Combining (15) and (14), this then requires

$$\mu_1^{1-\alpha_2} > \frac{2\left(\alpha_1+\alpha_2-1\right)\left(\mu_1^{1-\alpha_1}-z\right)}{1-2\alpha_1}. \tag{19}$$

In summary, we are left to find sufficient conditions on z such that (16), (17) can be satisfied subject to (18), (19) (and $\mu_1 \leq 1$). Formally, we will show that for every $z < \bar{z}$, (a) there exists $\mu_1 \in (0,1)$ solving (16) and (b) this solution satisfies (17)–(19).

Let us change variables from $(\alpha_1, \alpha_2, \mu_1)$ to

$$\gamma := \frac{1-2\alpha_1}{\alpha_1+\alpha_2-1}, \quad \Delta := \frac{\alpha_2-\alpha_1}{1-\alpha_2} \ \text{ and } \ y := \mu_1^{1-\alpha_2}.$$

Equations (16) and (17) simplify to:

$$F(y,z)=0, \tag{20}$$

$$z \leq H(y), \tag{21}$$

where

$$F(y,z) := 2x^{\Delta}(y-z)(1+\Delta) - \left(2z - 2y^{1+\Delta} + \gamma z\right)\left(1 - y^{\Delta}\left(1 - \frac{(y-z)\gamma}{2z - 2y^{1+\Delta} + \gamma y}\right)^{\frac{(2+\gamma)\Delta}{1+\gamma}}\right), \tag{22}$$

and

$$H(y) := y\left(\frac{2y^{\left(1+\frac{(2+\gamma)\Delta}{1+\gamma}\right)} + \Delta}{y^{1+\frac{\Delta}{1+\gamma}}(2+\gamma) + \Delta}\right). \tag{23}$$

Moreover, (18)–(19) simplifies to

$$z < y,\ y^{1+\Delta} - \frac{1}{2}\gamma y < z.$$

Note that combining the two inequalities implies,

$$I(y) < z, \tag{24}$$

where

$$I(y) := \frac{2y^{1+\Delta}}{2+\gamma} \tag{25}$$

Define

$$G(y,z) = \gamma(2+\gamma)\left(1 + \frac{(-y+z)\gamma}{2z + y(-2y^{\Delta}+\gamma)}\right)^{1+\frac{(2+\gamma)\Delta}{1+\gamma}}\Delta - 2(1+\gamma)(1+\Delta). \tag{26}$$

Claim 2. *If $G(y,z) < 0$, then*

$$-\frac{\partial F(y,z)/\partial y}{\partial F(y,z)/\partial z} > 0.$$

Note first that $F(y,z)$ is continuously differentiable in y and z and note further that $\partial F(y,z)/\partial z \neq 0$. Hence, by the implicit function theorem, there exists (a neighborhood of $(y,z) = (0,0)$ and) a unique function $J(y)$ defined over this neighborhood, whose graph $(y, J(y))$ is the set of all (y,z) such that $F(y,z) = 0$. Hence, $-(\partial F(y,z)/\partial y)/(\partial F(y,z)/\partial z) > 0$, is equivalent to $dJ(y)/dy > 0$.

Tedious algebra (using (20)) shows that

$$\frac{\partial F(y,z)/\partial y}{\partial F(y,z)/\partial z} \propto \frac{Num}{Den}, \tag{27}$$

where

$$Num := (-2+2v-\gamma)\Bigg(v^{1+\frac{(2+\gamma)\Delta}{1+\gamma}}\gamma + 2(1-v)(1+\Delta)\Bigg)\Bigg(v^{1+\frac{(2+\gamma)\Delta}{1+\gamma}}\gamma(2+\gamma)\Delta - 2(1+\gamma)(1+\Delta)\Bigg),$$

$$\begin{aligned} Den :=&(-2+2v-\gamma)\Bigg(v^{2+\frac{2(2+\gamma)\Delta}{1+\gamma}}(-1+2v)\gamma^2(2+\gamma)\Delta - 2v^{1+\frac{(2+\gamma)\Delta}{1+\gamma}}\gamma\Big((-2+2v-\gamma)(1+\gamma) \\ &+(-1+v)(2+\gamma)(-1+4v+\gamma)\Delta + 2(-1+v)(-1+2v)(2+\gamma)\Delta^2\Big) \\ &+4(-1+v)(1+\gamma)(1+\Delta)(2+\Delta)\Big(-(2+\gamma)(1+\Delta)+v\Big(2+\frac{(2+\gamma)\Delta}{1+\gamma}\Big)\Big)\Bigg), \end{aligned}$$

and

$$v := 1 + \frac{(z-y)\gamma}{2z + y(-2y^{\Delta}+\gamma)}.$$

Let us first consider the sign of Num. Note that $v<1$, and hence $(-2+2v-\gamma)<0$. Moreover, note that $v^{1+((2+\gamma)\Delta)/(1+\gamma)}\gamma + 2(1-v)(1+\Delta) > 0$. Hence, if $G(y,z)<0$, then $Num>0$. Next, consider the sign of Den. As before, $(-2+2v-\gamma)<0$. Moreover, it can easily be shown that if $G(y,z)<0$, then $Den/(-2+2v-\gamma)>0$.[30] Hence, if $G(y,z)<0$, then $Den<0$. Hence, if $G(y,z)<0$ is satisfied,

$$-\frac{\partial F(y,z)/\partial y}{\partial F(y,z)/\partial z} > 0.$$

Claim 3. *If $H(y)>z$, and $F(y,z)=0$, then $G(y,z)<0$.*

Suppose not. Suppose $H(y)>0$ (and hence, $z=0$) and $F(y,0)=0$. This implies $y=0$. Then $G(0,0)<0$. Suppose there exists a smallest $z>0$ such that $H(y)\geq z$, and $F(y,z)=0$ does not imply $G(y,z)<0$. For this smallest z, by continuity of all functions involved, we must have $z<H(y)$, $F(y,z)=0$ and $G(y,z)=0$. $F(y,z)=0$

[30]This can easily be seen when $G(y,z)=0$.

and $G(y,z)=0$ give explicit formulas for y and z, call them y^* and z^*. Tedious algebra shows $z^* > H(y^*)$, a contradiction.

The last two claims imply that if we derive an upper bound on y, we have an upper bound on z (recall, $z<y$).

Claim 4. *For every $z \leq \bar{z}$, there exists $y \in (0,1)$ solving $F(y,z)=0$. This solution satisfies $I(y) < z \leq H(y)$.*

Recall from the previous proof that

$$v = 1 + \frac{(z-y)\gamma}{2z + y(-2y^{\Delta}+\gamma)}.$$

Then $I(y) < z \leq H(y)$ becomes

$$0 < v \leq \frac{\Delta}{y^{1+\frac{\Delta}{1+\gamma}}\gamma + \Delta},$$

and (20) (when multiplied by $-(2+2v-\gamma)/(2+\gamma)$) becomes $\tilde{F}(v)=0$, where

$$\tilde{F}(v) := -v\gamma + y^{\Delta}\left(v^{1+\frac{(2+\gamma)\Delta}{1+\gamma}}\gamma - 2(-1+v)(1+\Delta)\right).$$

Note that $\tilde{F}(0) = 2y^{\Delta}(1+\Delta) > 0$. Hence, by the intermediate value theorem, it suffices to find

$$0 < v \leq \frac{\Delta}{y^{1+\frac{\Delta}{1+\gamma}}\gamma + \Delta},$$

such that $\tilde{F}(v) < 0$, i.e., v such that

$$y < \tilde{F}_1(v) := \left(\frac{v\gamma}{v^{1+\frac{(2+\gamma)\Delta}{1+\gamma}}\gamma + 2(1-v)(1+\Delta)}\right)^{\frac{1}{\Delta}}.$$

Hence, we can choose y as

$$\max_{v \leq \frac{\Delta}{\gamma+\Delta}} \tilde{F}_1(v).$$

(Note that $v \leq \Delta/(\gamma+\Delta)$ implies $v \leq \Delta/(y^{1+\Delta/(1+\gamma)}\gamma + \Delta)$ given that $y \leq 1$.) This is achieved by

$$v^* = \min\left\{\frac{\Delta}{\gamma+\Delta}, \left(2\frac{1+\gamma+\Delta+\gamma\Delta}{\gamma(2+\gamma)\Delta}\right)^{\frac{1+\gamma}{1+\gamma+2\Delta+\gamma\Delta}}\right\}.$$

Hence, the upper bound on y and hence, z is given by

$$\bar{z} = \tilde{F}_1(v^*).$$

Hence, to conclude,

$$\bar{z} = \left(\frac{v^* \gamma}{(v^*)^{1+\frac{(2+\gamma)\Delta}{1+\gamma}} \gamma + 2(1 - v^*)(1 + \Delta)} \right)^{\frac{1}{\Delta}},$$

where

$$v^* = \min \left\{ \frac{\Delta}{\gamma + \Delta}, \left(2 \frac{1 + \gamma + \Delta + \gamma\Delta}{\gamma(2 + \gamma)\Delta} \right)^{\frac{1+\gamma}{1+\gamma+2\Delta+\gamma\Delta}} \right\}.$$

Equivalently,

$$\bar{z}(\alpha_1, \alpha_2) = \left(\frac{2(1 - \alpha_1)(\alpha_1 + \alpha_2 - 1)}{(1 - 2\alpha_1)(1 - \alpha_2)} \frac{1 - m}{m} + m^{\frac{2\alpha_2 - 1}{1 - \alpha_2}} \right)^{-\frac{1-\alpha_2}{\alpha_2 - \alpha_1}},$$

where

$$m = \min \left\{ \left(\frac{2(1 - \alpha_1)(\alpha_1 + \alpha_2 - 1)}{(1 - 2\alpha_1)(2\alpha_2 - 1)} \right)^{\frac{1-\alpha_2}{\alpha_2}}, \frac{(\alpha_2 - \alpha_1)(\alpha_1 + \alpha_2 - 1)}{(1 - \alpha_2)^2 + 2\alpha_1\alpha_2 - \alpha_1 - \alpha_1^2} \right\}.$$

□

Part 2. Convergence as $z \to 0$.

Proof. The proof of Part 2 of Proposition 1 has the following steps. First, in any sequence of equilibria, $\mu_k \to 0$ for $k = 1, 2$ (Claim 5). Second, an equilibrium with support $\{\alpha_1, \alpha_2\}$ exists in the limit (Claim 6).

NB. Given that we have established that $\bar{z} > 0$ for all α_1, α_2 satisfying $\alpha_1 \leq 1/2$, $\alpha_1 + \alpha_2 > 1$, $\alpha_2 < 1$, we can drop (13) from now on.

Claim 5. *For (7)–(10) to be satisfied,* $\lim_{z \to 0} \mu_k = 0$ *for* $k = 1, 2$.[31]

[31] Here and in what follows, $z \to 0$ is shorthand for $\lim_{n \to \infty} z_n = 0$ (and likewise for other variables).

By (9) and (10), either $\lim_{z\to 0} q_k = 0$ or $\lim_{z\to 0} \mu_k = 0$ for $k = 1, 2$. Moreover, if $\lim_{z\to 0} q_k = 0$, then $\lim_{z\to 0} \mu_\ell = 0$. Recall that by Lemma 1, $\mu_2 < \mu_1$, $\forall z > 0$. Hence, by (9), it follows that $\lim_{z\to 0} \mu_2 = 0$. If $\lim_{z\to 0} \mu_2 = 0$, then (7) can only be satisfied if $\lim_{z\to 0} \mu_1 = 0$: if $\lim_{z\to 0} q_1 = 0$, then it must be that $\lim_{z\to 0} l_{2,1} = 1$, and hence, $\lim_{z\to 0} \mu_1 = 0$. Therefore, $\lim_{z\to 0} \mu_k = 0$ for $k = 1, 2$.

NB. Recall that by Lemma 1, in order for (7) to be satisfied it must be that $\mu_{k+1} \leq \mu_k$, $\forall k$, $\forall z > 0$. Hence, all ratios μ_ℓ/μ_k and μ_ℓ/μ_{k+1} in (7) and (8) are bounded above by 1. Hence, without loss, assume that these ratios converge. Call the ratios $l_{\ell,k}$ and $l_{\ell,k+1}$.

Claim 6. *There exists $\bar{z} > 0$ such that, for $z < \bar{z}$, the system (7)–(10) has a solution with*

$$\lim_{z\to 0} r_1 = \lim_{z\to 0} q_1 = \frac{2(\alpha_1 + \alpha_2 - 1)}{2\alpha_2 - 1}, \text{ and} \tag{28}$$

$$\lim_{z\to 0} s_1 = \frac{1 - \alpha_2}{2 - \alpha_1 - \alpha_2}. \tag{29}$$

We first reduce the system (7)–(10) to two equations. Then we use Taylor approximations to derive (28) and (29). As before, using (10), we can replace q_2 by $1 - q_1$ in (7). Similarly, we can then solve (7) for q_1 as a function of μ_1 and μ_2 only and replace q_1 and q_2 in (8) and (9). We can write the stubborn type's indifference condition, (8), as:

$$\frac{(1 - 2\alpha_1)(1 - \alpha_2)\left(\mu_2^{1-\alpha_2}\mu_1^{\alpha_1 - 1} - l_{2,1}^{\alpha_2}\right) + 2(1 - \alpha_1)(\alpha_1 + \alpha_2 - 1)\left(l_{2,1}^{1-\alpha_2} - 1\right)}{\mu_1^{-\alpha_2}\left(2(\alpha_1 + \alpha_2 - 1) l_{2,1}^{1-\alpha_2} - (2\alpha_2 - 1)\right)} = 0. \tag{30}$$

We can then show that

$$\lim_{z\to 0} \frac{\mu_2^{1-\alpha_2}}{\mu_1^{1-\alpha_1}} = \frac{2(1 - \alpha_1)(\alpha_1 + \alpha_2 - 1)}{(1 - 2\alpha_1)(1 - \alpha_2)}. \tag{31}$$

More precisely,

$$\mu_1 = \left(\frac{(1 - 2\alpha_1)(1 - \alpha_2)}{2(1 - \alpha_1)(\alpha_1 + \alpha_2 - 1)}\right)^{\frac{1}{1-\alpha_1}} \mu_2^{\frac{1-\alpha_2}{1-\alpha_1}} + \mathcal{O}\left(\mu_2^{\frac{1-\alpha_2}{1-\alpha_1}(1+\alpha_2-\alpha_1)}\right). \tag{32}$$

To derive (31) and (32), note that for (30) to be satisfied either

$$\lim_{z\to 0} l_{2,1} = K_0, \quad \text{or} \quad \lim_{z\to 0} \frac{\mu_2^{1-\alpha_2}}{\mu_1^{1-\alpha_1}} = K,$$

where K_0 is some positive constant. If $\lim_{z\to 0} l_{2,1} = K_0$, then $\lim_{z\to 0} \mu_2^{1-\alpha_2}/\mu_1^{1-\alpha_1} \to \infty$, and hence, (30) cannot be satisfied. If $\lim_{z\to 0} \mu_2^{1-\alpha_2}/\mu_1^{1-\alpha_1} = K_0$, then $\lim_{z\to 0} l_{2,1} = 0$. Hence, we can solve (30) for K_0:

$$K_0 = \frac{2\left(1-\alpha_1\right)\left(\alpha_1+\alpha_2-1\right)}{\left(1-2\alpha_1\right)\left(1-\alpha_2\right)}, \tag{33}$$

and (31) follows. Using Taylor approximation, we can then derive (32). Using (32), we can rewrite (9) and (14) as

$$q_1 = \frac{2\left(\alpha_1+\alpha_2-1\right)}{2\alpha_2-1} - k_1 \mu_2^{\frac{(1+\alpha_2)(1-\alpha_1)-(1-\alpha_2)^2}{1-\alpha_1}} + \mathcal{O}\left(\mu_2^{\frac{2\left(2\alpha_2-\alpha_1-\alpha_2^2\right)}{1-\alpha_1}}\right), \tag{34}$$

$$z = \frac{\left(1-2\alpha_1\right)\left(2-\alpha_1-\alpha_2\right)}{\left(1-\alpha_1\right)\left(2\alpha_2-1\right)} \mu_2^{1-\alpha_2} + \mathcal{O}\left(\mu_2^{\frac{1-2\alpha_1+\alpha_2(2-\alpha_2)}{1-\alpha_1}}\right), \tag{35}$$

where

$$k_1 = \left(\frac{2\left(\alpha_1+\alpha_2-1\right)}{2\alpha_2-1}\right)^2 \left(\frac{1-2\alpha_1}{2\left(\alpha_1+\alpha_2-1\right)}\right)^{\frac{\alpha_2-\alpha_1}{1-\alpha_1}} \left(\frac{1-\alpha_1}{1-\alpha_2}\right)^{\frac{1-\alpha_2}{1-\alpha_1}}.$$

To derive (34), note that we can write $l_{2,1}^{1-\alpha_2}$ as

$$l_{2,1}^{1-\alpha_2} = \left(\frac{\left(1-2\alpha_1\right)\left(1-\alpha_2\right)}{2\left(1-\alpha_1\right)\left(\alpha_1+\alpha_2-1\right)}\right)^{-\frac{1-\alpha_2}{1-\alpha_1}} \mu_2^{\frac{(1+\alpha_2)(1-\alpha_1)-(1-\alpha_2)^2}{1-\alpha_1}} + \mathcal{O}\left(\mu_2^{\frac{2\left(2\alpha_2-\alpha_1-\alpha_2^2\right)}{1-\alpha_1}}\right).$$

Using (35), and recalling that $s_1 = (\mu_1^{1-\alpha_1} q_1)/z$, we can now write s_1 as a function of μ_2 only:

$$s_1 = \frac{1-\alpha_2}{2-\alpha_1-\alpha_2} - k_2 \mu_2^{\frac{(1+\alpha_2)(1-\alpha_1)-(1-\alpha_2)^2}{1-\alpha_1}} + \mathcal{O}\left(\mu_2^{\frac{1-2\alpha_1+2\alpha_2-\alpha_2^2}{1-\alpha_1}-1+\alpha_2}\right), \tag{36}$$

where

$$k_2 = \left(\frac{\left(1-2\alpha_1\right)\left(1-\alpha_2\right)}{2\left(\alpha_1+\alpha_2-1\right)\left(1-\alpha_1\right)}\right)^{\frac{\alpha_2-\alpha_1}{1-\alpha_1}} \left(\frac{2\left(\alpha_1+\alpha_2-1\right)\left(1-\alpha_1\right)}{\left(2\alpha_2-1\right)\left(2-\alpha_1-\alpha_2\right)}\right).$$

Hence,

$$\lim_{z\to 0} r_1 = \frac{2(\alpha_1+\alpha_2-1)}{2\alpha_2-1}, \text{ and} \lim_{z\to 0} s_1 = \frac{1-\alpha_2}{2-\alpha_1-\alpha_2}. \tag{37}$$

Finally, I now show that the system (7)–(10) can be solved locally around $z = 0$, with $s_1 \in (0,1)$, $r_1 \in (0,1)$.

As before, I replace q_2 by $1 - q_1$ in equations (9), (7) and (8) (using (10)). In analogue to before, I then solve (7) for q_1 as a function of μ_1 and μ_2 only:

$$q_1 = \frac{2(\alpha_1+\alpha_2-1)\left(1-\left(\frac{\mu_2}{\mu_1}\right)^{1+\alpha_2}\right)}{(2\alpha_2-1)-2(\alpha_1+\alpha_2-1)\left(\frac{\mu_2}{\mu_1}\right)^{1+\alpha_2}}. \tag{38}$$

Using this, I can then use (9) to solve for μ_2 as a function of z and μ_1:

$$\mu_2 = \mu_1 \left(\frac{2(\alpha_1+\alpha_2-1)\mu_1^{1-\alpha_1} - (2\alpha_2-1)z+}{2(\alpha_1+\alpha_2-1)\left(\mu_1^{1-\alpha_1}-z\right)-(1-2\alpha_1)\mu_1^{1-\alpha_2}}\right)^{\frac{1}{1-\alpha_2}}. \tag{39}$$

Hence, I can express (8) as a function of μ_1 and z only. Let me introduce two auxiliary variables, p and u, where

$$p = z^{\frac{\alpha_1-\alpha_2(1-\alpha_1)+2\alpha_2^2}{(1-\alpha_1)(1-\alpha_2)}}, \text{ and} \tag{40}$$

$$u = \mu_1^{1-\alpha_1} z^{-1} - \frac{(1-\alpha_2)(2\alpha_2-1)}{2(2-\alpha_1-\alpha_2)(\alpha_1+\alpha_2-1)}. \tag{41}$$

Given (40) and (41), one can derive:

$$\left.\frac{dp}{du}\right|_{(p,u)=(0,0)} = \frac{(2-\alpha_1-\alpha_2)}{1-\alpha_1}\left(\frac{2(2-\alpha_1-\alpha_2)(\alpha_1+\alpha_2-1)}{(1-\alpha_2)(2\alpha_2-1)}\right)^{\frac{\alpha_2-\alpha_1}{1-\alpha_1}} > 0. \tag{42}$$

I can rewrite (8) as a function of p and u, using (40) and (41). Denote this new function $\Delta^s_{p,u}$. Taking derivatives w.r.t. p and u, evaluating these derivatives at $p = u = 0$ (the solution for $z = 0$), and rearranging, I get (42), which is clearly finite and positive:

$$\begin{aligned}\left.\frac{dp}{du}\right|_{(p,u)=(0,0)} &= -\left.\frac{\partial\Delta^s_{p,u}/\partial u}{\partial\Delta^s_{p,u}/\partial p}\right|_{(p,u)=(0,0)} \\ &= \frac{(2-\alpha_1-\alpha_2)}{1-\alpha_1}\left(\frac{2(2-\alpha_1-\alpha_2)(\alpha_1+\alpha_2-1)}{(1-\alpha_2)(2\alpha_2-1)}\right)^{\frac{\alpha_2-\alpha_1}{1-\alpha_1}}.\end{aligned} \tag{43}$$

Hence, by the Implicit Function Theorem, the system (7)–(10) can be solved locally around $z = 0$ when $K = 2$, with $r_1 \in (0, 1)$, and $s_1 \in (0, 1)$. $\square$

Part 3: Convergence as $z \to 1$.

Proof. Consider a pooling equilibrium with support $\{\alpha_1, \ldots, \alpha_K\}$. Recall that any such equilibrium has to satisfy (3)–(6). By (6), either $\lim_{z\to 1} q_k = 0$ or $\lim_{z\to 1} \mu_k = 1$. The stubborn type's payoff from α_K in a candidate equilibrium is:

$$v_s(\alpha_K) = \sum_{k=1}^{K} q_k(1 - \alpha_k)(1 - \mu_k^{\alpha_K}).$$

This implies that $\lim_{z\to 1} v_s(\alpha_K) = 0$.

Consider instead, the payoff from making a demand of $1 - \alpha_K$, where α_K is the highest demand being made in equilibrium:

$$v_s(1 - \alpha_K) = \sum_{k=1}^{K} q_k \left(\frac{2 - \alpha_K - \alpha_k}{2}\right).$$

Unless $q_K = 1$ and $\alpha_K = 1$, $v_s(1 - \alpha_K) > 0$ for any $z \in [0, 1]$. This implies that if $\alpha_K < 1$, there exists $\bar{z} < 1$ such that for any $z > \bar{z}$, the stubborn type prefers to deviate to $1 - \alpha_K$. This implies $\lim_{z\to 1} s = \delta_1$. $\square$

$\square$

Proof of Proposition 2. The proof procedes in three steps.

Part 1: Existence of Separating Equilibria for fixed z.

Proof. We start with some preliminary observations. Note that for any z, in any equilibrium, any separating demand by the stubborn type must be smaller than the lowest demand assigned positive probability by the rational type. Moreover, the separating demand can be no higher than 1/2. Otherwise, the rational type would have an incentive to deviate from this lowest demand to the separating demand by the stubborn type.

Note further that for any z, in any separating equilibrium, there can be at most one separating demand by the stubborn type. If there were multiple, the payoff to the stubborn type from the higher separating demand would be strictly higher.

Finally, note that the lowest separating demand by the rational type must be incompatible with the separating demand by the stubborn type. Otherwise, the stubborn type would prefer to deviate to the rational type's lowest separating demand. Hence, $\alpha_1 + \alpha_2 > 1$. Consider the following strategy profile. Suppose the rational type randomizes over demands $\alpha_2, \ldots, \alpha_K$, i.e., $r = \sum_{k=2}^{K} \delta_{\alpha_k} \tilde{r}_k$ with $\sum_{k=2}^{K} \tilde{r}_k = 1$, and suppose $s = \delta_{\alpha_1}$. If both players demand α_1, there is immediate agreement. If both players demand a "rational equilibrium demand" (i.e., demands $\alpha^i, \alpha^j \in \operatorname{supp} r$), there is no immediate concession, and rational player i demanding α^i and facing α^j concedes at rate $\lambda^i = (\rho(1-\alpha^i))/(\alpha^i + \alpha^j - 1)$. If player i demands α_1 and player j demands $\alpha_j \in \operatorname{supp} r$, player j immediately concedes to i with probability 1. If a player faces an incompatible demand $\alpha \notin \operatorname{supp} s \cup \operatorname{supp} r$, he does not concede. After the initial demands are made, a player demanding α_1 is believed to be stubborn with probability 1. A player demanding $\alpha \neq \alpha_1$ is believed to be stubborn with probability 0. This results in the following payoffs:

$$v_s(\alpha_1) = z\frac{1}{2} + (1-z)\alpha_1, \text{ and}$$
$$v_r(\alpha_\ell) = z(1-\alpha_1) + (1-z)\sum_{k=2}^{K} \tilde{r}_k(1-\alpha_k), \ \forall \alpha_k \in \operatorname{supp} r.$$

Then the rational type has no incentive to deviate to the stubborn type's demand if and only if

$$v_r(\alpha_K) \geq z\frac{1}{2} + (1-z)\alpha_1,$$

or equivalently, if and only if

$$z \geq \underline{z} = \frac{2\left(\alpha_1 - \sum_{k=2}^{K} \tilde{r}_k(1-\alpha_k)\right)}{1 - 2\sum_{k=2}^{K} \tilde{r}_k(1-\alpha_k)} = \frac{2\left(\alpha_1 + \sum_{k=2}^{K} \tilde{r}_k \alpha_k - 1\right)}{2\sum_{k=2}^{K} \tilde{r}_k \alpha_k - 1}.$$

Moreover, note that the stubborn type has no incentive to deviate to $1-\alpha_1$ iff:

$$v_s(\alpha_1) \geq z(1-\alpha_1).$$

Or equivalently, if and only if $z \leq \bar{z} = 2\alpha_1$. Clearly, neither the rational nor the stubborn type has any incentive to deviate to any other demand (given the strategy profile and system of beliefs given above). Hence, there exists a separating equilibrium if and only if $\alpha_1 < 1/2$, $\alpha_1 + \alpha_2 > 1$, and

$$z \in [\underline{z}, \bar{z}] .$$

□

Part 2: Convergence as $z \to 0$.

Proof. Note that $\underline{z}$ is increasing in the excess surplus demanded ($\alpha_1 + \sum_{k=2}^{K} \tilde{r}_k \alpha_k - 1$ and $\underline{z} = 0$ if and only if $\alpha_1 + \alpha_2 = 1$ and $\tilde{r}_1 = 0$ and $\tilde{r}_2 = 1$. Hence, $\lim_{z \to 0} r = \delta_{1-\alpha}$, and $\lim_{z \to 0} s = \delta_{\alpha}$. □

Part 3: Convergence as $z \to 1$.

Proof. Note $\bar{z}$ is increasing in α_1 and $\bar{z} = 1$ if and only if $\alpha_1 = 1/2$. Hence, $\lim_{z \to 1} s = \delta_{1/2}$. □

□

Proof of Proposition 3. Part (i): First, note that we may specify the probability that a stubborn type picks α_k in $B^{AG,S}$, denoted $\pi_S(\alpha_k)$, such that $z_{AG}\pi_{AG}(\alpha_k) + z_S \pi_S(\alpha_k) = z s_k$ for every k. This ensures that the posterior probability that a player is rational is the same in both games, for every α_k. Hence, continuation payoffs are the same. It follows that any equilibrium in B is an equilibrium in $B^{AG,S}$ (there are fewer deviations to deter).[32]

Part (ii) follows from Abreu and Gul (2000): given any distribution over behavioral types, if the type $\alpha = 1/2$ is present, the rational type makes the offer $\alpha = 1/2$ with probability 1 in the limit as $z \to 0$. □

[32]Note that this also implies that there may be equilibria in $B^{AG,S}$ that are not equilibria in B: given that some behavioral types (the AG types) have no ability to deviate, there are fewer deviations to consider.

Proof of Corollary 1. Claim 5 and 6 (Proposition 1) implies that for z sufficiently small $s(\alpha_2) > 1/2$. Hence, if $z_{AG} = z_S$, Part 1 of Proposition 3 implies the corollary. □

Proof of Proposition 4. Consider the two bargaining games, B with (Θ, P) and $\bar{B}$ with $(\Theta, \bar{P})$. Note that $\bar{P}$ is less informative about types than P. Formally, $\bar{P}$ is a garbling of P, i.e., there exists a Markov matrix Γ such that $\bar{P} = \Gamma P$. Hence, ignoring incentive compatibility constraints (i.e., ignoring (3) and (4)), any probability distribution over types that can be generated with $\bar{P}$, can be generated with P. Suppose then that $((r, F^r_{\alpha_i,\alpha_j}), s)$ is an equilibrium of B, which induces beliefs Π. Then there exists a strategy profile in $\bar{B}$ that induces beliefs Π if and only if $\bar{P} \succsim_{SOSD} \Pi$.

If $(\sigma^\theta, F^\theta_{\alpha_i,\alpha_j})_\theta$ is an equilibrium of $\bar{B}$ with support $\mathcal{C}$, then the payoff to a type $\theta \in (0,1)$ from demanding $\alpha \in \mathcal{C}$ is simply a weighted average of the payoff a type 0 and the payoff a type 1 would receive:

$$v_\theta(\alpha) = \theta v_1(\alpha) + (1-\theta) v_0(\alpha).$$

Hence, the payoff is linear in θ. Therefore, if $\theta, \theta' \in \Theta$ are indifferent over a set of demands, so are $0, 1 \in \Theta$, and vice-versa. This implies that (3) and (4) are satisfied in B. Further, if a deviation outside of the support is unprofitable for both types 0 and 1, then it is also unprofitable for a type $\theta \in (0,1)$, keeping the beliefs off-path fixed. □

References

[1] Abreu, Dilip, and Faruk Gul (2000): “Bargaining and Reputation,” *Econometrica*, 68(1), 85–117.

[2] Abreu, Dilip, and David Pearce (2007): “Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts,” *Econometrica*, 75(3), 653–710.

[3] Abreu, Dilip, David Pearce, and Ennio Stacchetti (2015): “One-sided Uncertainty and Delay in Reputational Bargaining,” *Theoretical Economics*, 10(3), 719–773.

[4] Abreu, Dilip, and Rajiv Sethi (2003): “Evolutionary Stability in a Reputational Model of Bargaining,” *Games and Economic Behavior*, 44(2), 195–216.

[5] Atakan, Alp, and Mehmet Ekmekci (2014): “Bargaining and Reputation in Search Markets,” *Review of Economic Studies*, 81(1), 1–29.

[6] Ausubel, Lawrence, and Raymond Deneckere (1992): "Durable Goods Monopoly with Incomplete Information," *Review of Economic Studies*, 59(4), 795–812.

[7] Ausubel, Lawrence, and Raymond Deneckere (1993): "Efficient Sequential Bargaining," *Review of Economic Studies*, 60(2), 435–461.

[8] Backus, Matthew, Thomas Blake, Bradley Larsen, and Steven Tadelis (2020): "Sequential Bargaining in the Field: Evidence from Millions of Online Bargaining Interactions," *Quarterly Journal of Economics*, 135(3), 1319–1361.

[9] Binder, Sarah (1999): "The Dynamics of Legislative Gridlock, 1947–96," *American Political Science Review*, 93(3), 519–533.

[10] Card, David (1990): "Strikes and Wages: A Test of an Asymmetric Information Model," *Quarterly Journal of Economics*, 105(3), 625–659.

[11] Compte, Olivier, and Philippe Jehiel (2002): "On the Role of Outside Options in Bargaining with Obstinate Parties," *Econometrica*, 70(4), 1477–1517.

[12] Fanning, Jack (2016): "Reputational Bargaining and Deadlines," *Econometrica*, 84(3), 1131–1179.

[13] Fanning, Jack (2018): "No Compromise: Uncertain Costs in Reputational Bargaining," *Journal of Economic Theory*, 175, 518–555.

[14] Fanning, Jack (2021): "Mediation in Reputational Bargaining," *American Economic Review*, 111(8), 2444–2472.

[15] Fanning, Jack, and Alexander Wolitzky (2022): "Reputational Bargaining." In *Bargaining: Current Research and Future Directions*, edited by Emin Karagözoglu and Kyle B. Hyndman, Palgrave Macmillan.

[16] Fearon, James (1994): "Domestic Political Audiences and the Escalation of International Disputes," *American Political Science Review*, 88(3), 577–592.

[17] Fudenberg, Drew, and Jean Tirole (1991): *Game Theory*. Cambridge, MA: The MIT Press.

[18] Kambe, Shinsuke (1994): "Bargaining with Imperfect Commitment," *Games and Economic Behavior*, 28(2), 217–237.

[19] Kim, Kyungmin (2009): "The Coase Conjecture with Incomplete Information on the Monopolist's Commitment," *Theoretical Economics*, 4(1), 17–44.

[20] Myerson, Roger, and Mark Satterthwaite (1983): "Efficient Mechanisms for Bilateral Trading," *Journal of Economic Theory*, 29(2), 265–281.

[21] Myerson, Roger (1991): *Game Theory: Analysis of Conflict*. Cambridge, MA: Harvard University Press.

[22] Ozyurt, Selçuk (2014): "Audience Costs and Reputation in Crisis Bargaining," *Games and Economic Behavior*, 88, 250–259.

[23] Ozyurt, Selçuk (2015a): "Search for a Bargain: Power of Strategic Commitment," *AEJ: Microeconomics*, 7(3), 320–353.

[24] Ozyurt, Selçuk (2015b): "Bargaining, Reputation and Competition," *Journal of Economic Behavior and Organization*, 119, 1–17.

[25] Wolitzky, Alexander (2012): "Reputational Bargaining with Minimal Knowledge of Rationality," *Econometrica*, 80(5), 2047–2087.