ONE-SIDED UNCERTAINTY AND DELAY IN REPUTATIONAL BARGAINING

Dilip Abreu
Princeton University

David Pearce
New York University

and

Ennio Stacchetti
New York University

June 4, 2014

Abstract. A two-person infinite-horizon bargaining model where one of the players may have either of two discount factors, has a multiplicity of perfect Bayesian equilibria. Introducing the slightest possibility that either player may be one of a rich variety of stationary behavioral types singles out a particular solution and appears to support some axiomatic treatments in the early literature in their conclusion that there is a negligible delay to agreement. Perturbing the model with a slightly broader class of behavioral types that allows the informed player to delay making his initial demand still achieves powerful equilibrium refinement. But there is substantial delay to agreement, and predictions depend continuously on the ex ante probabilities of the patient and impatient types of the informed player, counter to what the literature suggests.

This research was supported in part by National Science Foundation Awards 0922125 and 0922167. We are grateful to Deepal Basak, Martin Osborne and several anonymous referees for providing valuable comments. Thanks also to Juan Dubra and Phil Reny for helpful discussions.

1. Introduction

Rubinstein (1982) delighted economists by establishing uniqueness of perfect equilibrium in an infinite horizon bargaining model. Once the surprise wore off, attention moved to another intriguing feature of the model: in the unique equilibrium, agreement is reached immediately. While this does not square well with some real-world phenomena (protracted haggling over prices, strikes in labor negotiations and so on), it was expected that introducing asymmetric information into the model would easily produce delay to agreement. If the purpose of holding out for a better deal is to signal the strength of one's bargaining position, then the existence of asymmetric information (without which there would be nothing to signal) might naturally be expected to go hand in hand with delay to agreement.

The asymmetric information bargaining literature did not unfold exactly as hoped. The early papers revealed a vast multiplicity of perfect Bayesian equilibria, even for one-sided asymmetric information (Rubinstein (1985)) or for only two periods in the case of bilateral informational asymmetry (Fudenberg and Tirole (1983)). More specific results relied on severely limited strategy spaces (Chatterjee and Samuelson (1987)), appeals to "reasonable" selections from the equilibrium correspondences (Sobel and Takahashi (1983), Cramton (1984), Chatterjee and Samuelson (1988)) or axiomatic restrictions of equilibrium (Rubinstein (1985) and Gul and Sonnenschein (1988)). The latter two papers study one-sided asymmetric information and produce solutions displaying virtually *no delay to agreement*. Gul and Sonnenschein's solutions have a further "Coasean" feature[1]: the uninformed player, facing an opponent drawn from a distribution of payoff types, does as badly as she would if she instead faced, with certainty, the strongest possible opponent from that distribution. (Both these results apply to situations where offers can be made frequently.)

This paper investigates the effects of introducing behavioral perturbations[2] into a bargaining model with one-sided asymmetric information. Will this achieve equilibrium selection, as in Fudenberg and Levine (1989), Abreu and Gul (2000) and Abreu and Pearce (2007)? And if so, will the predictions agree with the axiomatic treatment of Rubinstein (1985)? The model we perturb is the leading example of the class covered by Rubinstein

[1] Coase (1971) conjectured that when a durable goods monopolist faces buyers with a distribution of valuations, most sales will occur almost immediately, at a price near the infimum of that valuation distribution. Coase assumed that the monopolist was free to adjust prices frequently. In the 1980's a series of papers verified the conjecture with increasing conclusiveness and generality. See especially Stokey (1981), Bulow (1982), Fudenberg, Levine and Tirole (1985) and Gul, Sonnenschein and Wilson (1986), and for a critique of some of the axioms imposed, Ausubel and Deneckere (1989).

[2] A trio of papers (Kreps and Wilson (1982), Kreps, Milgrom, Roberts and Wilson (1982) and Milgrom and Roberts (1982)) opened the reputational literature by demonstrating the power that even small reputational perturbations could have in games with long or infinite horizons. Fudenberg and Levine (1989) show the success a patient long-run player can have against short-run opponents by choosing which of many reputational types to imitate. Much of the ensuing literature is surveyed authoritatively by Mailath and Samuelson (2006). See Wolitzky (2011) for a demonstration of the importance of "transparency" assumptions on announcements by behavioral types.

(1985). Player A is of known preferences, but she is unsure which of two discount rates player B uses to discount his payoff stream.

We begin in Section 2 by considering stationary behavioral types, who never waver from the demands they make at the beginning of the game. This is the class of types used by Myerson (1991) and Abreu and Gul (2000). We show that when ex ante probabilities of behavioral types are small, equilibrium is essentially unique: with high probability, play ends almost immediately, and the uninformed player's expected payoff is virtually what she would have received in a full information Rubinstein (1982) solution if her opponent were known to be the stronger (more patient) of the two possible rational types. This reinforces the message of Inderst (2005), who showed that in a durable goods monopoly problem, endowing the monopolist with an ex ante reputation for (possibly) being a behavioral type, does *not* overturn Coase's predictions unless that ex ante probability is substantial. See Kim (2009) for extensions of that work. Our results from Section 2 agree with Rubinstein (1985) regarding the negligible time to agreement, but give a patient type a higher payoff than does Rubinstein (1985) when that type has low ex ante probability.

Abreu and Pearce (2007) established that the stationary behavioral types are a "sufficiently rich" class[3] of perturbations to consider in stationary bargaining games (or more generally, in repeated games with contracts). There are reasons to doubt that this is true with asymmetric information. Cramton (1984) emphasizes the importance for an informed player of delaying his first offer, to signal strength (his Introduction opens with a dramatic illustration from military history). Accordingly, in Section 3 we expand the set of behavioral types for the informed player so that player B can use the tactic of delaying making a demand, without losing his reputation for being behavioral. This innovation turns out to be crucial. While equilibrium is still unique, it takes an entirely different form from Section 2. A hybrid equilibrium results from the patient player B trying to use delay to separate himself from the impatient version of B. For many parameter values, the uninformed player A does *substantially better* than in the Coasean solution, and there is considerable expected *delay to agreement*. Numerical results illustrate how the form of the solution varies with the exogenous parameters. A detailed summary of what form equilibria can take is available early in Section 3, in the subsection "Overview of Section 3".

Notably, the payoffs of the unique solution of Section 3 are continuous in the prior probability that the informed player B is the more patient of his two rational types. Contrast this to the durable goods monopoly problem, where only the lower support of the price distribution matters to the seller. The same discontinuity occurs in Gul and Sonnenschein (1986).

Section 4 addresses the question of existence of equilibrium. The characterization results of the earlier sections place strong restrictions on candidate equilibria. As a result, a partially constructive approach succeeds in guaranteeing the existence of equilibrium.

Section 5 concludes, emphasizing the conditions that give the uninformed player a

[3]More precisely, there is no advantage to being able to imitate any type outside this class, even if your opponent can.

more generous payoff than the Coasean prescription, and speculating about the implications for bargaining with two-sided asymmetric information.

Related Literature.

As discussed above, the received literature on one-sided asymmetric information about discount rates or valuations favors "Coasean" solutions, with efficient outcomes and no delay. Working in a model with fixed costs of waiting, rather than discounting, Bikhchandani (1992) offers axioms with a different flavor from those of Rubinstein (1985) that rule out all of his solutions, and can produce delay to agreement. Existence is guaranteed only if the three costs of waiting (those of the uninformed party and of the two types of her opponent) are sufficiently similar. Unlike in our work, equilibrium does not exhibit a "war of attrition" structure. Readers interested in the vast literature on bargaining games with asymmetric information (without reputational perturbations) might also look, for example, at Cho (1990), Watson (1998) and the excellent survey paper by Ausubel, Cramton and Deneckere (2002).

We do not draw any conclusions about a durable goods monopolist facing a continuum of buyers. First, the classic problem (Coase (1971)) concerns a distribution of buyer *valuations*, and we are instead focused on discount factors. Secondly, reputational perturbation of the buyers' side would be much more complicated, with a continuum of buyers, than in our model with just one informed player. As noted earlier, Inderst (2005) introduces a behavioral perturbation on the seller's side only, and concludes that unless the perturbation is done with significant probability, the Coasean conclusion persists. See also Kim (2009). There have been many non-reputational attempts to escape Coasean conclusions in durable goods monopoly. These include rental instead of purchase (Bulow (1982)), capacity constraints (Kahn (1986) and McAfee and Wiseman (2008)), non-Markovian strategies (Ausubel and Deneckere (1989) and Sobel (1991)), correlations between the costs of seller and buyers (Evans (1989), Vincent (1989), Deneckere and Liang (2006), Fuchs and Skrzypacz (2010) and Fuchs and Skrzypacz (2012), who allow for the arrival of new buyers) and stochastic, time-varying costs (Ortner (2012)).

Fanning (2013) obtains striking results in a reputational analysis of bargaining with a stochastic deadline. Like us, he considers the effects of allowing for behavioral types that are not restricted to the simplest strategies; he is able to derive the nonstationary type that is optimal (that is, the one that a rational player should imitate with probability close to 1).

2. Atemporal Types

Can reputational perturbations resolve the equilibrium multiplicity problem in bargaining models with asymmetric information? And is it enough to restrict perturbations to the simplest stationary types, as in Abreu and Gul (2000)? This Section takes a simple asymmetric information bargaining model of the kind studied in Rubinstein (1985) and allows for the presence of stationary behavioral types on each side. Equilibrium selection is achieved and the conclusions are Coasean in the sense explained in the Introduction.

The results here are a benchmark against which to compare what happens when slightly more complex behavioral types are introduced in Section 3.

Model.

Two players bargain over the division of a surplus. It is convenient to adopt a continuous/discrete time model in which a player can change his demand at any positive integer time, but can concede to an outstanding demand at any time $t \in [0, \infty)$. This modelling device, introduced in Abreu and Pearce (2007), allows us to do war of attrition calculations in continuous time, while avoiding the usual pathologies that arise in continuous time games. So that this modelling device will not introduce "openness" problems,[4] we need to divide each calendar time into three logically consecutive "dates", as follows. Time 0 is divided into three logically sequential dates $(0, -1)$, $(0, 0)$ and $(0, +1)$. Player A starts the game by making her opening demand $a \in [0, 1]$ at $(0, -1)$. Having observed A's demand, B makes a counterdemand $b \in [0, 1]$ at $(0, 0)$, and each player has the option of accepting the other's offer at $(0, +1)$. No discounting occurs in moving from date $(0, -1)$ to $(0, 0)$, or from $(0, 0)$ to $(0, +1)$. Each subsequent time $t \in \{1, 2, 3, \dots\}$ is divided into the date $(t, -1)$, when either player can change his or her demand, date $(t, 0)$ at which either player can make a new demand if the opponent's standing demand changed at $(t, -1)$, and finally the date $(t, +1)$, when either player can accept a newly changed demand. If at any time the players' demands a and b become compatible (that is, $a + b \leq 1$), each of the divisions of surplus $(a, 1 - a)$ or $(1 - b, b)$ is implemented with probability 1/2, and the game ends. Similarly, if A and B accept each other's offers as the game opens, the two demands are implemented with equal probability.

The players can be "rational" or "behavioral" independently. Rational player A's discount rate is $r^A > 0$ and rational player B's discount rate is either r_1^B or r_2^B, where $r_1^B > r_2^B > 0$ and $\beta_k = \mathbb{P}[r_k^B \mid \text{B is rational}]$, $k = 1, 2$. Each player's type (whether he is rational or behavioral and the value of his discount rate) is private information. Rational players are assumed to maximize the expected discounted value of their shares. If agreement is never reached, they both receive 0 payoffs. Behavioral types for A and B are represented by two finite[5] sets $\mathbb{A}, \mathbb{B} \subset (0, 1)$. A behavioral player A of type $a \in \mathbb{A}$ makes the initial demand a, never changes her initial demand, and accepts (immediately) a counterdemand b if and only if $1 - b \geq a$. Behavioral types $\mathbb{B}$ are similarly defined. Player $i \in \{A, B\}$ is behavioral with probability z^i and for each $a \in \mathbb{A}$ and $b \in \mathbb{B}$,

$$\pi^A(a) = \mathbb{P}[a \mid \text{A is behavioral}] \quad \text{and} \quad \pi^B(b) = \mathbb{P}[b \mid \text{B is behavioral}].$$

We denote this incomplete information game by $\Gamma(r, \beta, z)$, where $r = (r^A, r_1^B, r_2^B)$, $\beta = (\beta_1, \beta_2)$ and $z = (z^A, z^B)$. The parameters $\mathbb{A}$, $\mathbb{B}$, π^A and π^B are held fixed throughout.

[4]Suppose, for example, that at some time t, player A makes a demand to which B wants to acquiesce right away. In continuous time, there is no "first time after t", so $\mathcal{B}$ has no best response. This need for immediate acceptance is accommodated by adding time $(t, +1)$. Additionally, it is convenient to let B make an immediate counterdemand in response to a new demand by A at t.

[5]Finiteness of behavioral types is common in the literature, and simplifies the analysis greatly. There is no indication that relaxing it would change the asymptotic results.

Let $\underline{a} = \min \mathbb{A}$ and $\bar{a} = \max \mathbb{A}$, and similarly for $\mathbb{B}$. We assume that $\underline{a} + \bar{b} > 1$, $\bar{a} + \underline{b} > 1$, and that $\pi^A(a) > 0$ and $\pi^B(b) > 0$ for all $a \in \mathbb{A}$ and $b \in \mathbb{B}$.

Analysis.

Throughout the paper, by "equilibrium" we will mean a Nash equilibrium that satisfies sequential rationality along the equilibrium path.[6] Our results do not depend on any restrictions on behavior off the equilibrium path. Hereafter, we find it convenient to call $\mathcal{A}$ the rational player A and $\mathcal{B}_k$ the rational player B with discount rate r_k^B, $k = 1, 2$. We will use $\mathcal{B}$ to refer to a player B of either *rational* type.

After initial postures $(a, b) \in \mathbb{A} \times \mathbb{B}$ are adopted, each rational player randomly chooses a time to accept the opponent's demand (if the opponent has not accepted already). A rational player has the option of changing his initial demand at any $t \in \{1, 2, \dots\}$. But doing so would reveal that he is rational and in the continuation game, equilibrium implies that he should concede to the opponent's demand immediately. This is an implication of Lemma 1 below and its proof. Lemma 1 is another expression of Coasean dynamics (see Section 8.8 of Myerson (1991), Proposition 4 of Abreu and Gul (2000), and Lemma 1 of Abreu and Pearce (2007)). The proofs of Lemma 1 and most subsequent results are relegated to the Appendix.

Lemma 1. *Suppose that the players have made the initial demands (a, b) and that neither player has revealed rationality prior to time s. If revealing rationality at s is in the support of $\mathcal{A}$'s equilibrium strategy and if $\mathcal{A}$ reveals rationality at s while B does not, $\mathcal{A}$'s resulting equilibrium continuation payoff is $1 - b$ and B's is b. An analogous conclusion holds when $\mathcal{B}$ is the first to reveal rationality at s. Finally, if $\mathcal{A}$ reveals rationality at date $(0, -1)$ (by demanding $a \notin \mathbb{A}$), then $\mathcal{A}$'s continuation payoff is $(1 - z^B)(1 - \bar{b}) + z^B \sum \pi^B(b)(1 - b)$ and $\mathcal{B}$'s continuation payoff is $\bar{b}$. Moreover, in equilibrium, if $\mathcal{A}$ reveals rationality with positive probability at $s > 0$, then $\mathcal{B}$ does not, and conversely.*

Note that Lemma 1 includes the case when $s = 0$ and $\mathcal{B}$ is the first to reveal rationality at date $(0, 0)$ (by demanding $b \notin \mathbb{B}$).

An implication of Lemma 1 and its proof is that corresponding to any equilibrium in which $\mathcal{A}$ chooses $a \notin \mathbb{A}$ at date $(0, -1)$ there is an equivalent equilibrium in which $\mathcal{A}$ chooses instead $\bar{a} \in \mathbb{A}$ and concedes right away to any counterdemand $b \in \mathbb{B}$. A similar remark applies to a choice $b \notin \mathbb{B}$ by $\mathcal{B}$ at date $(0, 0)$. Henceforth we will restrict attention to equilibria where $\mathcal{A}$'s initial demand is in $\mathbb{A}$ and $\mathcal{B}$'s initial demand is in $\mathbb{B}$.

Let $\varphi^A(a)$ and $\varphi_k^B(b|a)$ denote respectively the equilibrium probabilities that $\mathcal{A}$ chooses an initial posture $a \in \mathbb{A}$ and that, after observing a, $\mathcal{B}_k$ chooses an initial pos-

[6] In a finite extensive form game, sequential rationality holds everywhere on the equilibrium path (see, for example, Myerson (1991), Lemma 4.4). To see that this need not be the case in an infinite game, consider a game in which Nature draws a number from the uniform distribution on $[0, 1]$, and after observing that number, the only strategic player chooses 1 (which gives him payoff 1) or 0 (which gives him payoff 0). Although it is natural to think that he would always choose 1, any strategy that chooses 1 on all but a set of measure 0 is also a Nash equilibrium.

ture $b \in \mathbb{B}$. A pair of choices $(a,b) \in \mathbb{A} \times \mathbb{B}$ with $a+b>1$ leads to the "subgame"[7] $\Gamma(r, \hat\beta_1(a,b), \hat\beta_2(a,b), \hat z^A(a), \hat z^B(a,b), a, b)$, where

$$\hat z^A(a) = \frac{z^A \pi^A(a)}{z^A\pi^A(a) + (1-z^A)\varphi^A(a)}$$
$$\hat z^B(a,b) = \frac{z^B\pi^B(b)}{z^B\pi^B(b) + (1-z^B)[\beta_1\varphi_1^B(b|a) + \beta_2\varphi_2^B(b|a)]}$$
$$\hat\beta_k(a,b) = \frac{(1-z^B)\beta_k\varphi_k^B(b|a)}{z^B\pi^B(b) + (1-z^B)[\beta_1\varphi_1^B(b|a) + \beta_2\varphi_2^B(b|a)]} \quad k=1,2.$$

are the posterior probabilities that player A is behavioral, and that player B is behavioral or $\mathcal{B}_k$, respectively. We will use the convention that a "hat" over a probabilistic variable denotes a posterior probability, where the arguments denote the conditioning information. Note also that for simplicity we will often omit the arguments (a,b) and simply write, for example, $\hat z^A$ and $\hat z^B$ instead of $\hat z^A(a)$ and $\hat z^B(a,b)$.

For any $(a,b) \in \mathbb{A} \times \mathbb{B}$ with $a+b>1$, the subgame $\Gamma(r, \hat\beta_1, \hat\beta_2, \hat z^A, \hat z^B, a, b)$ has a unique equilibrium outcome, similar to that obtained by Abreu and Gul (2000). We first solve the subgame for perturbations of arbitrary size, finding an expression (generalizing that of Abreu and Gul (2000)) that shows what causes a player to be "strong" or "weak", in the sense explained below. Simpler expressions obtain in the limit as perturbation probabilities approach 0. Once the asymptotic properties of the subgame are in hand, it is easy to see what demands players will make in equilibrium (with high probabilities).

Let $\mu^i = (1-\hat z^i) \times$ (probability that rational i concedes at time 0) be the (total) probability that $i \in \{A, B\}$ concedes at time 0. Then $\mu^A \cdot \mu^B = 0$ since when $\mu^B > 0$, for example, $\mathcal{A}$ strictly prefers to wait at time 0. If $0 < \mu^B \le \hat\beta_1$, then only $\mathcal{B}_1$ concedes immediately with positive probability, but if $\mu^B > \hat\beta_1$, then $\mathcal{B}_1$ concedes immediately with probability 1 and $\mathcal{B}_2$ concedes immediately with positive probability. Thus, if the players do not concede immediately, the relevant posteriors become

$$\hat z^i(0) = \frac{\hat z^i}{1-\mu^i} \quad i = A, B, \quad \hat\beta_2(0) = \min\left\{\frac{\hat\beta_2}{1-\mu^B}, 1-\hat z^B(0)\right\},$$

and $\hat\beta_1(0) = 1 - \hat z^B(0) - \hat\beta_2(0)$. Recall that $\hat z^i$ is the posterior probability that player i is behavioral at the start of the subgame. Below, $\hat z^i(\tau)$ denotes the posterior probability that player i is behavioral conditional on not conceding up to and including time τ. Consistent with this notation, $\hat z^i(0)$ denotes the posterior probability that player i is behavioral conditional on i not conceding immediately at the start of the subgame. Since (a,b) are held fixed here, the only new conditioning information for posteriors is τ. After the demands at time 0, there follows a war of attrition (hereafter WOA) divided into two time intervals

[7] In a simplifying abuse of terminology, we use the term subgame even when referring to continuation games that do not begin at a singleton information set.

$(0,\tau_1]$ and $(\tau_1,\tau_2]$. Player A concedes at a Poisson rate λ_1^A in $(0,\tau_1]$ and a Poisson rate λ_2^A in $(\tau_1,\tau_2]$, while B concedes at a Poisson rate λ^B in the whole interval $(0,\tau_2]$, where[8]

$$\lambda_k^A = \frac{r_k^B(1-a)}{a+b-1} \quad k=1,2, \qquad \lambda^B = \frac{r^A(1-b)}{a+b-1}\,.$$

If $\hat{\beta}_1(0)=0$ then $\tau_1=0$, and if $\hat{\beta}_2=0$ then $\tau_2=\tau_1$. If $\hat{\beta}_1(0)=0$ or $\hat{\beta}_2(0)=0$, the WOA is exactly that studied by Abreu and Gul (2000). When $\hat{\beta}_1>0$, λ_1^A keeps $\mathcal{B}_1$ indifferent between conceding and waiting, but at τ_1 A becomes convinced that she is not dealing with $\mathcal{B}_1$, and switches to the concession rate λ_2^A that keeps $\mathcal{B}_2$ indifferent between conceding and waiting. The concession rate λ^B keeps $\mathcal{A}$ indifferent in $(0,\tau_2]$. The players' reputations (that is, the posteriors that they are behavioral) grow exponentially over time:

$$\hat{z}^B(\tau)=\hat{z}^B(0)e^{\lambda^B\tau} \quad \text{and} \quad \hat{z}^A(\tau)=\begin{cases} \hat{z}^A(0)e^{\lambda_1^A\tau} & \tau\in(0,\tau_1] \\ \hat{z}^A(\tau_1)e^{\lambda_2^A\tau} & \tau\in(\tau_1,\tau_2]. \end{cases}$$

At time τ_2 both players' reputations reach 1 simultaneously. Given μ^A and μ^B, τ_1 and τ_2 are defined by

$$[\hat{z}^B(0)+\hat{\beta}_2(0)]e^{\lambda^B\tau_1}=1 \quad \text{and} \quad \hat{z}^B(0)e^{\lambda^B\tau_2}=1.$$

If at least one player is rational, the subgame ends in agreement randomly in the interval $[0,\tau_2]$ with probability 1; if both players are behavioral, the subgame never ends.

The analysis proceeds by backward induction. Lemma 2 summarizes the solution of the subgame we have been discussing, taking the demands made and the updated beliefs as parameters. The expression L in the lemma captures the relative "strengths" of the two players' positions: unless the situation is perfectly balanced, where $L=1$, one of the players is "weak", and needs to concede with positive probability at time zero. Player $\mathcal{A}$'s position is weakened by any of the following changes: an increase in her demand or a decrease in B's demand, an increase in her rate of interest or a decrease in either r_1^B or r_2^B, or a decrease in her reputation $\hat{z}^A$ or an increase in $\hat{z}^B$.

Lemma 2. *$\Gamma(r,\hat{\beta}_1,\hat{\beta}_2,\hat{z}^A,\hat{z}^B,a,b)$ with $(a,b)\in\mathbb{A}\times\mathbb{B}$ and $a+b>1$ has a unique equilibrium outcome. Let*

$$L=\frac{[\hat{z}^A]^{\lambda^B}}{[\hat{z}^B+\hat{\beta}_2]^{\lambda_1^A-\lambda_2^A}[\hat{z}^B]^{\lambda_2^A}}\,.$$

When $L\le 1$, $\mu^B=0$ and $\mu^A=1-L^{1/\lambda^B}$. When $L\ge 1$, $\mu^A=0$ and

$$\mu^B=\begin{cases} 1-\hat{z}^B/[\hat{z}^A]^{\lambda^B/\lambda_2^A} & \text{if} \quad 1-\hat{z}^B/[\hat{z}^A]^{\lambda^B/\lambda_2^A}\ge\hat{\beta}_1 \\ 1-1/L^{1/\lambda_1^A} & \text{otherwise.} \end{cases}$$

[8]These are *total* rates of concession. To attain the rate λ_2^A in the interval $(\tau_1,\tau_2]$, for example, $\mathcal{A}$ must concede at Poisson rate $\lambda_2^A/(1-\hat{z}^A(\tau))$, where $\hat{z}^A(\tau)$, defined below, is the posterior that A is behavioral given that he has not conceded by time τ. Similarly, in the same interval, $\mathcal{B}_2$ must concede at Poisson rate $\lambda^B/(1-\hat{z}^B(\tau))$.

Remark: Note that when $\hat{\beta}_2 = 0$ or $\hat{\beta}_1 = 0$, the WOA reduces to that studied by Abreu and Gul (2000) where there is only one type of rational player B. In this case,

$$L = \frac{[\hat{z}^A]^{\lambda^B}}{[\hat{z}^B]^{\lambda_1^A}} \quad \text{if } \hat{\beta}_2 = 0, \quad \text{and} \quad L = \frac{[\hat{z}^A]^{\lambda^B}}{[\hat{z}^B]^{\lambda_2^A}} \quad \text{if } \hat{\beta}_1 = 0.$$

Proof. Most of the results follow directly from the analysis in Abreu and Gul (2000). Here we only deduce the value of μ^B when $L > 1$. When $\mu^B \geq \hat{\beta}_1$, $\mathcal{B}_1$ concedes immediately with probability 1 (and $\mathcal{B}_2$ concedes immediately with nonnegative probability). Thus, if B does not concede immediately, A concludes that she is dealing with $\mathcal{B}_2$ or a behavioral type. Consequently, $\tau_1 = 0$ and A concedes to b at a constant Poisson rate λ_2^A in the interval $(0, \tau_2]$. Thus

$$\frac{\hat{z}^B}{1-\mu^B} e^{\lambda^B \tau_2} = 1 \quad \text{and} \quad \hat{z}^A e^{\lambda_2^A \tau_2} = 1.$$

These equations imply that $1 - \mu^B = \hat{z}^B / [\hat{z}^A]^{\lambda^B/\lambda_2^A}$.

When $\mu^B < \hat{\beta}_1$, $\mathcal{B}_1$ concedes immediately with probability less than 1, and $\tau_1 > 0$. In this case,

$$\left[\frac{\hat{\beta}_2 + \hat{z}^B}{1-\mu^B}\right] e^{\lambda^B \tau_1} = 1, \quad \frac{\hat{z}^B}{1-\mu^B} e^{\lambda^B \tau_2} = 1 \quad \text{and} \quad \hat{z}^A e^{\lambda_1^A \tau_1 + \lambda_2^A (\tau_2 - \tau_1)} = 1.$$

These equations imply that $1 - \mu^B = [1/L]^{1/\lambda_1^A}$. □

Lemma 3 concerns the limiting properties of equilibrium after demands have been made, as the initial reputations approach zero (in any manner not violating an arbitrarily loose bound). The striking result here is that player B's strength or weakness is affected neither by the interest rate r_1^B of his more impatient rational type, nor by the probability $\hat{\beta}_1$ of that type. Only r_2^B and $\hat{\beta}_2$ contribute to his strength (along with the impatience of player $\mathcal{A}$). This is explained by the fact that for $\hat{z}^A$ and $\hat{z}^B$ very small, *almost 100% of the war of attrition will be spent in the second phase* (see the paragraphs preceding Lemma 2 above), in which A faces the more patient type of $\mathcal{B}$. To understand why, consider the following example. Fix $\hat{\beta}_1$, the probability that B is the impatient rational type, at .9; the residual probability is divided between the probability (bounded above by .1) that B is the patient rational type, and $\hat{z}^B$, the probability B is behavioral. Absent any concessions at time 0, it takes a fixed amount of time τ_1 (dependent on $\hat{\beta}_1$, which we will not change) to finish the first stage of the war of attrition (given the rate λ^B at which B needs to concede to A, Bayes' Rule determines the time τ_1 at which nine tenths of the B population, that is, all the impatient ones, will have conceded). Now let $\hat{z}^B$ approach zero. The length of the entire war of attrition grows without bound, but the first stage is not increasing in length. Even though the impatient type was more abundant than the patient type at time zero, A spends almost 100% of the war of attrition fighting the patient type, when $\hat{z}^B$ is negligible. For this reason, λ_2^A appears in the statement of Lemma 3, whereas λ_1^A does not.

By Lemma 2, the subgame $\Gamma(r, \hat{\beta}_1, \hat{\beta}_2, \hat{z}^A, \hat{z}^B, a, b)$ has a unique equilibrium outcome. Hence, an equilibrium for $\Gamma(r, \beta, z)$ is fully specified by the probabilities $\varphi^A(a)$ and $\varphi_k^B(b|a)$, $k = 1, 2$ and $(a, b) \in \mathbb{A} \times \mathbb{B}$ with which behavioral types are mimicked.

Notice in Lemma 3 that the consequences of "weakness" in a player's position (see the discussion above Lemma 2) are exaggerated as reputational perturbations become very small[9]: the weak player concedes at time zero with probability approaching 1. The asymptotic results here do not depend on the relative ex ante probabilities that the players are behavioral, as long as the ratios of those probabilities do not approach 0 or ∞. This restriction is conveniently imposed by looking at vectors of perturbation probabilities that lie in a closed cone excluding the coordinate axes. *Throughout the paper we will phrase all of our asymptotic results (for slight reputational perturbations) in terms of truncated cones.* For any $R > 1$ and $\bar{z} > 0$, define the cone and truncated cone

$$K(R) = \{(z^A, z^B) \mid z^A > 0, z^B > 0 \text{ and } \max\{z^A/z^B, z^B/z^A\} \leq R\},$$
$$K(R, \bar{z}) = \{(z^A, z^B) \in K(R) \mid z^A \leq \bar{z} \text{ and } z^B \leq \bar{z}\}.$$

Lemma 3. *Let $R > 1$ and $\{z^\ell\} \subset K(R)$ be a sequence such that $z^\ell = (z^{A\ell}, z^{B\ell}) \downarrow (0, 0)$. For each ℓ, let φ^ℓ be an equilibrium of $\Gamma(r, \beta, z^\ell)$. Assume that $\varphi^\ell \to \varphi^\infty$ in $\mathbb{R}^{|\mathbb{A}|} \times [\mathbb{R}^{|\mathbb{A}| \times |\mathbb{B}|}]^2$. For a given $(a, b) \in \mathbb{A} \times \mathbb{B}$ with $a + b > 1$, consider the corresponding subgames $\Gamma(r, \hat{\beta}^\ell, \hat{z}^\ell, a, b)$. Let $(\hat{z}^{B\infty}, \hat{\beta}_1^\infty, \hat{\beta}_2^\infty)$ be the limit of $\{(\hat{z}^{B\ell}, \hat{\beta}_1^\ell, \hat{\beta}_2^\ell)\}$.*

(i) If $\varphi_1^{B\infty}(b|a) + \varphi_2^{B\infty}(b|a) > 0$ and $\lambda_2^A > \lambda^B$, then $\mu^{B\ell} \to 1$.
(ii) If $\varphi^{A\infty}(a) > 0$ and $\hat{z}^{B\infty} > 0$, then $\mu^{A\ell} \to 1$.
(iii) If $\varphi^{A\infty}(a) > 0$, $\hat{\beta}_2^{B\infty} > 0$ and $\lambda^B > \lambda_2^A$, then $\mu^{A\ell} \to 1$.

We now develop some notation that will be useful in the proof of Lemma 4 and subsequently. For each $a \in (0, 1)$ and $k = 1, 2$, let

$$b_k^*(a) = \max\left\{1 - a, 1 - \frac{r_k^B}{r^A}(1 - a)\right\}.$$

Assume A demands $a \in (0, 1)$. When $b_k^*(a) > 1 - a$, $b_k^*(a)$ is the "balanced counterdemand" that equalizes the Poisson rates of concessions when A faces only $\mathcal{B}_k$ (and behavioral types):

$$\lambda_k^A(a, b_k^*(a)) = \frac{r_k^B(1 - a)}{a + b_k^*(a) - 1} = \frac{r^A(1 - b_k^*(a))}{a + b_k^*(a) - 1} = \lambda^B(a, b_k^*(a)).$$

When the demand a is too modest, any counterdemand $b > 1 - a$ is excessive, that is, yields $\lambda_k^A > \lambda^B$. In this case we define $b_k^*(a) = 1 - a$. For $k = 1, 2$, we assume that

$$b_k^*(a) \notin \mathbb{B} \quad \text{and} \quad \underline{b} < b_k^*(a) \quad \text{for all } a \in \mathbb{A}.$$

[9]The same is true for the asymptotic characterizations in Abreu and Gul (2000). This effect first appears in Kambe (1999) who has no behavioral types ex ante, but rather a uniform probability that a player may get compulsive about any initial demand he might make. Compte and Jehiel (2002) investigate the role of outside options in the presence of behavioral types and some of their results also entail taking perturbation probabilities to zero.

This reflects the fact that generically, the exactly balanced response to a will not be in the finite type set. Let

$$\lfloor b_k^*(a) \rfloor = \max\left\{1 - a, \max\left\{b \in \mathbb{B} \mid b < b_k^*(a)\right\}\right\} \quad k = 1, 2.$$

Observe that when $\lfloor b_k^*(a) \rfloor > 1 - a$, $\lfloor b_k^*(a) \rfloor$ is the largest behavioral demand $b \in \mathbb{B}$ such that $\lambda^B(a, b) > \lambda_k^A(a, b)$. Furthermore, if $r_1^B > r_2^B$ (as assumed), there clearly exists $\underline{\Delta} > 0$ such that $\lambda_1^A(a, b_2^*(a) - \Delta) > \lambda^B(a, b_2^*(a) - \Delta)$ for all $\Delta \leq \underline{\Delta}$ and $a \in \mathbb{A}$. We will assume throughout that the grid of types $\mathbb{B}$ is fine enough that $\lfloor b_2^*(a) \rfloor > b_2^*(a) - \underline{\Delta}$ for all $a \in \mathbb{A}$.

Suppose that the reputational perturbations z^A and z^B are very slight. Once A has made an equilibrium demand $a \in \mathbb{A}$, with high probability $\mathcal{B}$ responds by demanding the highest amount $b \in \mathbb{B}$ such that b is less greedy than the balanced demand $b_2^*(a)$. Any demand higher than this, if made with noticeable probability in equilibrium, would leave $\mathcal{B}$ in a weak position, from which he would need to concede with probability near 1. Lemma 4 establishes the payoff consequences for each player. Note again the prominent role of $b_2^*(a)$: the interest rate r_1^B and the relative probabilities of $\mathcal{B}$'s patient and less patient types play no role at all in determining players' asymptotic strengths and payoffs.[10]

Lemma 4. *For any $R > 1$ and $\epsilon > 0$, there exists $\bar{z} > 0$ such that for all $z \in K(R, \bar{z})$, for any equilibrium (φ^A, φ^B) of $\Gamma(r, \beta, z)$, and for any $a \in \mathbb{A}$, the corresponding expected payoff for $\mathcal{A}$ satisfies*

$$v^A(a, z) \geq 1 - \lfloor b_2^*(a) \rfloor - \epsilon.$$

Moreover, if $\varphi^A(a) \geq \epsilon$, then

$$v_k^B(a, z) \geq \lfloor b_2^*(a) \rfloor - \epsilon \quad k = 1, 2.$$

For sufficiently small perturbation probabilities, Lemma 4 tells us that $\mathcal{A}$'s expected payoff from demanding a is arbitrarily close to $1 - \lfloor b_2^*(a) \rfloor$. As a corollary, almost all of $\mathcal{A}$'s weight must go on $\tilde{a}_2^*$, where

$$\tilde{a}_k^* \in \operatorname{argmin}_{a \in \mathbb{A}} \lfloor b_k^*(a) \rfloor \quad k = 1, 2.$$

For simplicity[11], we assume that the argmin is a singleton.

Corollary. *For any $R \in (0, \infty)$ and $\epsilon > 0$, there exists $\bar{z} > 0$ such that for all $z \in K(R, \bar{z})$, and for any equilibrium (φ^A, φ^B) of $\Gamma(r, \beta, z)$, $\varphi^A(a) < \epsilon$ for all $a \neq \tilde{a}_2^*$.*

With almost all of $\mathcal{A}$'s weight on $\tilde{a}_2^*$ and $\mathcal{B}$ almost always responding with $\lfloor b_2^*(a) \rfloor$, the expected payoffs to both players are clear.

[10]This is not true, of course, when z is distant from 0. In that case, although any counterdemand in the support of $\mathcal{B}_2$'s strategy is also in that of $\mathcal{B}_1$, and the war of attrition following such a demand still has two phases, the length of each phase is a non-negligible fraction of the total. In general, $\mathcal{B}_1$ may make some counterdemands that $\mathcal{B}_2$ does not.

[11]Our results can be rephrased throughout – at the cost of some clumsiness in the statements and proofs – for the general case.

Theorem 1. *For any* $R \in (0, \infty)$ *and* $\epsilon > 0$, *there exists* $\bar{z} > 0$ *such that for all* $z \in K(R, \bar{z})$, *and for any equilibrium* (φ^A, φ^B) *of* $\Gamma(r, \beta, z)$,

$$v^A(z) \geq 1 - \lfloor b_2^*(\tilde{a}_2^*) \rfloor - \epsilon \qquad \textit{and} \qquad v_k^B(z) \geq \lfloor b_2^*(\tilde{a}_2^*) \rfloor - \epsilon, \quad k = 1, 2.$$

Let a_2^* satisfy $r^A a_2^* = r_2^B(1 - a_2^*)$. The demand a_2^* is such that its balanced counterdemand proposes the same partition: $b_2^*(a_2^*) = 1 - a_2^*$. Note that if the grids of types $\mathbb{A}$ and $\mathbb{B}$ are fine then $\tilde{a}_2^* \approx a_2^*$ and $\lfloor b_2^*(\tilde{a}_2^*) \rfloor \approx 1 - a_2^*$, and more generally, $\lfloor b_2^*(a) \rfloor \approx b_2^*(a)$ for all $a \in \mathbb{A}$.

In summary, we see that with vanishing reputational perturbations of the simple type considered in Myerson (1991) and Abreu and Gul (2000), there is a unique division of surplus, and essentially no delay to agreement. Further, this division depends on the discount rate of the stronger (more patient) type of the informed player, but *not* on the discount rate of his weaker type, and *not* on the relative probabilities of those two types. The rest of the paper explores whether these results apply equally if richer perturbation types are admitted.

3. Temporal Types

In the model of Section 2, there is no scope for $\mathcal{B}$ to use initial silence to signal his patience: delaying making a demand would reveal his rationality, giving A the decisive reputational advantage. Here, we remedy the situation in the simplest possible way, allowing for reputational types of the informed player that wait a variety of lengths of time before making a demand. It turns out that in equilibrium, B can now signal that he is either behavioral, or the patient rational type. This can have dramatic implications for the payoffs achieved by the two players, showing that in this asymmetric information environment, simple atemporal types are not "canonical" in the sense of Footnote 3 (whereas they are in the symmetric information settings of Abreu and Pearce (2007)).

Model and Assumptions.

The temporal model differs from the atemporal model in that player B is now allowed to make his initial counterdemand with delay. Once player A makes her initial demand a, player B can accept it or wait[12] until some time $t \in [0, \infty)$ to make a counterdemand b. Similarly to the previous model, once the counterdemand b is made, the players can change their (counter)demands only at times $\{t+1, t+2, \dots\}$, but can concede to an outstanding demand at any time $\tau \in [t, \infty)$. It will turn out that in equilibrium, there is a positive probability that $\mathcal{A}$ will acquiesce to B's counterdemand at t immediately. Phenomena such as this are accommodated as in Section 2, by splitting each calendar time into three logically consecutive dates.

[12]We adopt this order of moves to match Rubinstein (1985), to which our perturbed model is most naturally compared, as well as Abreu and Gul (2000). Whereas we can assert, in Section 2, with a rich set of reputational types, that the order of initial demands makes little difference, we do not have analogous results here.

Formally, then, the set of dates is $\mathbb{R}_+ \times \{-1, 0, +1\}$. The game begins with a demand from A at date $(0, -1)$, that B may respond to with a counterdemand at date $(0, 0)$. If B does not respond at date $(0, 0)$, B's initial demand can be made at any $(t, -1)$ with $t > 0$. For any $t \geq 0$, if either player makes a new demand at $(t, -1)$, it can be responded to immediately at $(t, 0)$. An existing demand can be accepted at any date $(t, +1)$. No discounting applies between dates $(t, -1)$ and $(t, +1)$. Although she will not want to in equilibrium, player i can *replace* an existing demand with a new one at any date $(t+k, -1)$, $k = 1, 2, ...$, where $(t, -1)$ is the last date at which either player made a demand.

Rational players $\mathcal{A}$ and $\mathcal{B}_k$, $k = 1, 2$, have discount rates r^A and r_k^B, $k = 1, 2$, respectively, where $r_1^B > r_2^B$. The set of behavioral players for A is $\mathbb{A}$ as before (with the same interpretation), but behavioral types for B are now represented by the set $\mathbb{B} \times [0, \bar{T}]$, where $\bar{T}$ is a sufficiently distant time (as we discuss later). Again, $\mathbb{A}, \mathbb{B} \subset (0, 1)$. A behavioral player B of type (b, t) makes his initial counterdemand b at time t, never changes his demand, and concedes (immediately) to a demand a if and only if $1 - a \geq b$. A puzzle arises regarding one aspect of how behavioral types of B should be formulated. The idea is to provide $\mathcal{B}$ with types to imitate that are more aggressive than a rational player would be. Consider the eventuality in which $\mathcal{A}$ changes her demand before B has made his initial demand. Since that reveals her to be rational, $\mathcal{B}$ would immediately demand $\bar{b}$. Accordingly, we assume B's behavioral types are all "reactive" in exactly that sense also and immediately demand $\bar{b}$ (and only accept a demand $a \leq 1 - \bar{b}$).[13]

Let z^i be the probability that player $i \in \{A, B\}$ is behavioral, $\pi^A(a)$ be the conditional probability that A is type $a \in \mathbb{A}$ given that she is behavioral, and $\pi^B(b, t)$ be the conditional probability *density* that B is type $(b, t) \in \mathbb{B} \times [0, \bar{T}]$ given that he is behavioral. We assume that $\pi^B(b, t)$ is continuous in t for each $b \in \mathbb{B}$, and that there exists $\underline{\pi} > 0$ such that

$$\pi^A(a) \geq \underline{\pi} \quad \text{and} \quad \pi^B(b, t) \geq \underline{\pi} \quad \text{for all} \quad a \in \mathbb{A} \text{ and } (b, t) \in \mathbb{B} \times [0, \bar{T}].$$

We denote this game by $\Gamma(r, \beta, z)$.

As before, we assume that $\underline{a} + \bar{b} > 1$ and $\bar{a} + \underline{b} > 1$. Furthermore, we assume that

$$\bar{b} e^{-r_2^B \bar{T}} < 1 - \bar{a},$$

so that the more patient player $\mathcal{B}_2$ (and therefore $\mathcal{B}_1$ as well) would prefer to accept the demand $\bar{a}$ immediately to waiting until after $\bar{T}$ to make the counterdemand $\bar{b}$, even if $\bar{b}$ were then immediately accepted by A. This is what we meant earlier by $\bar{T}$ being sufficiently large. Finally, we assume that $\bar{T}$ is not an integer and that

$$(1 - \bar{a})^{1 + r^A / r_1^B} > (1 - \bar{b}) r^A / r_2^B.$$

[13]The assumption of reactive types is convenient. Otherwise at the very "end of the game" when virtually all normal types would have in the current equilibrium shown their hand, player $\mathcal{A}$ might find it in her interest to "skim through" (non-reactive) behavioral types who by this point have substantial posterior probability. But knowing this, normal types would find it worthwhile to wait for A to reveal her normalcy. While the full equilibrium analysis of the very end game would be much more complex, we conjecture that our asymptotic results would be unchanged but would entail more elaborate end game behavior after $\mathcal{B}$ had almost certainly moved.

Roughly speaking this requires[14] that $1-\bar{a}$ is sufficiently larger than $1-\bar{b}$. This assumption is used in the proof of Lemma 7.

In equilibrium, $\mathcal{A}$ chooses $a \in \mathbb{A}$ with probability $\varphi^A(a)$. For each demand $a \in \mathbb{A}$, player $\mathcal{B}_k$ either chooses to accept a immediately or to mimic a behavioral type $(b,t) \in \mathbb{B} \times [0,\bar{T}]$ with probability density $\varphi_k^B(b,t|a)$, $k=1,2$.[15] These densities are conditional upon $\mathcal{A}$ not changing her intitial demand before B moves. By Lemma 7 below, the latter eventuality never arises in equilibrium. Consequently, after the demand $a \in \mathbb{A}$ is made, A is behavioral with posterior probability

$$\hat{z}^A(a) = \frac{z^A \pi^A(a)}{z^A \pi^A(a) + (1-z^A)\varphi^A(a)}, \tag{1}$$

and after the counterdemand $(b,t) \in \mathbb{B} \times [0,\bar{T}]$, B is behavioral or $\mathcal{B}_k$, $k=1,2$, with posterior probabilities

$$\hat{z}^B(a,b,t) = \frac{z^B \pi^B(b,t)}{z^B \pi^B(b,t) + (1-z^B)[\beta_1 \varphi_1^B(b,t|a) + \beta_2 \varphi_2^B(b,t|a)]} \quad \text{and} \tag{2}$$

$$\hat{\beta}_k(a,b,t) = \frac{(1-z^B)\beta_k \varphi_k^B(b,t|a)}{z^B \pi^B(b,t) + (1-z^B)[\beta_1 \varphi_1^B(b,t|a) + \beta_2 \varphi_2^B(b,t|a)]} \quad k=1,2. \tag{3}$$

Overview of Section 3.

Readers of earlier versions have found this Section challenging, and suggested that it would be helpful to know from the beginning where the analysis was headed. This, then, is a reader's guide to the remainder of the Section.

What can an equilibrium of the temporal model look like? The analysis starts by considering the continuation game after player A makes her initial demand. The first three lemmas, culminating in Lemma 7, build a case for A remaining silent after that initial demand, until B makes his counterdemand. Lemma 1 of Section 2, which applies equally here, shows that if both players have spoken, then if either side reveals rationality, he or she might as well give up and acquiesce to the other side's demand. Readers can skip Lemmas 5 through 7 without much loss, but are urged to read carefully the subsequent material before Lemma 8: it lays out the basic strategic considerations of the subgame, and shows why there cannot be a separating equilibrium. Lemma 8 establishes some helpful facts about the "sneaking in curve" which plays a crucial role in the analysis of hybrid equilibria.

[14]Our present formulation of the assumption, while not essential to the conclusion, allows us to treat all a's symmetrically and streamlines an already involved argument. But our subsequent payoff characterizations require only that Lemma 7 hold for a's that are not "unreasonably large" (in terms of balancedness). For each such a, one would impose the inequality condition with a replacing $\bar{a}$ on the left-hand side. Large demands that violate this condition would be used with only vanishing probability by $\mathcal{A}$, because otherwise, they would yield payoffs lower than those attainable by demands that satisfy the inequality.

[15]Note that the probabilities $\varphi_k^B(b|a)$ of the previous section now correspond to densities $\varphi_k^B(b,t|a)$.

Recall that in the previous Section with atemporal types, both $\mathcal{B}_1$ and $\mathcal{B}_2$ put almost all their weight (asymptotically) on $\lfloor b_2^*(a) \rfloor$, which would be "slightly generous" in response to a in a game where there were no impatient type (that is, $\beta_1 = 0$). "Slightly generous" means "the least generous response that is more generous than balanced". This remains true for $\mathcal{B}_2$ in the temporal setting. But even asymptotically, $\mathcal{B}_1$ may put substantial weight on both $\lfloor b_2^*(a) \rfloor$ and $\lfloor b_1^*(a) \rfloor$. The latter occurs only very near $t = 0$. This is formalized in Lemma 10 and its corollaries. As a consequence, one can equivalently study a "reduced game" in which $\mathcal{B}$ is restricted not to use those counterdemands that get vanishing weight in equilibria of the full game.

A sequence of numbered observations preceding Lemma 11 establishes that in the reduced game, *any equilibrium must take one of two forms.* The first, which we call non-Coasean because there may be substantial delay to agreement (and it gives the uninformed player $\mathcal{A}$ more than she would get in a game where she surely faced her more patient opponent), involves an initial time interval (see Figure 2) in which only $\mathcal{B}_1$, the weak type of $\mathcal{B}$, is active, followed by an interval of pooling, and finally a third interval in which only the strong type $\mathcal{B}_2$ is active. (Being active at t means t is in the support of the set of times at which that type might speak.) In this kind of equilibrium, $\mathcal{B}_1$ will, with positive probability, either concede at time 0 to A's demand or, in a vanishing neighborhood of the origin, demand $\lfloor b_1^*(a) \rfloor$. Otherwise $\mathcal{B}_1$ demands $\lfloor b_2^*(a) \rfloor$somewhere in the first two of the three intervals just described. The strong type $\mathcal{B}_2$ is active in the second region (pooling with $\mathcal{B}_1$) and (alone) in the third region, in either case demanding $\lfloor b_2^*(a) \rfloor$. Asymptotically (in the ex-ante probabilities of behavioral types), the lengths, but not the probabilistic importance, of intervals two and three approach zero.

The second possible form an equilibrium could take, illustrated in figure 3, we call Coasean, because both rational types will accept almost immediately, and players get, asymptotically, what they would in the full information game where B was known to be the patient type. In this equilibrium, there is no initial region where $\mathcal{B}_1$ is active alone, and the pooling and "only $\mathcal{B}_2$ is active" regions are vanishingly short, as perturbation probabilities approach zero. Both $\mathcal{B}_1$ and $\mathcal{B}_2$ demand $\lfloor b_2^*(a) \rfloor$. Lemma 11 verifies who is active in each of these regions in the two kinds of equilibrium, and Lemmas 12 and 13 show which regions vanish asymptotically in each form of equilibrium. Note that although the lengths of these regions vanish, the probability of B's being active in those regions does not. Indeed, in the Coasean case, for example, both types of player B are active with probability 1 in an initial interval of vanishing length.

In a particular bargaining game, with fixed parameters, could equilibria of both types exist? No, and Lemmas 14 and 15 lay the groundwork for proving it. An equilibrium must always provide player $\mathcal{B}$ of each type with the right incentives to include each desired time in the support of his strategy. And to provide those payoffs, the equilibrium needs to have each type active at that time with the right density. Informally, think of this as the equilibrium having a "demand" for a certain amount of each type. If the supplies of each type differ from this because the ex ante probability that $\mathcal{B}$ is weak is too high, for example, the equilibrium "dumps" the surplus density of the weak type by having him make a positive probability concession at time zero. If, on the other hand, the probability of the weak type were sufficiently lower, there would be no way to make the equilibrium work, and equilibrium would have to be of the second (Coasean) type instead. Theorem 2 divides the parameter space into two regions, in the first of which only the first (non-Coasean) equilibrium can arise, and in the second of which only Coasean equilibrium can arise. Asymptotic payoffs are unique, for any set of exogenous parameters.

What can be said about $\mathcal{A}$'s demand? This is the subject of Theorems 4, 5 and 6. Since $\mathcal{A}$'s choice leads to complex subgames, this is a particularly messy problem, simplified somewhat by looking at a straightforward refinement (see details later). We prove that if the ex ante probability that B is the patient type is sufficiently high, $\mathcal{A}$ will make only the Coasean demand (what she would make were it common knowledge that she faced the patient type), and conversely if that probability is sufficiently low, she will demand strictly more than her Coasean payoff. Toward the end of Section 3 we present numerical illustrations of how different the graph of $\mathcal{A}$'s expected payoff as a function of her demand can look, depending on initial parameters. In some cases $\mathcal{A}$ can choose, by being more or less aggressive in her demand, a continuation equilibrium that is either Coasean in the subgame or non-Coasean. Sometimes she will forgo the non-Coasean option, because although it would have given her a concession with positive probability at time zero, it would also entail a substantial probability of delaying settlement. When $\mathcal{B}$'s two types differ greatly in patience and the patient type is not too likely, it becomes worthwhile for $\mathcal{A}$ to separate them by choosing a non-Coasean option. As in Section 2, proofs are mostly relegated to the Appendix.

Preliminary Analysis.

The analysis in the preceding Section leads one to suspect, correctly, that each side will eventually imitate a behavioral type, and a war of attrition (or an immediate probabilistic concession) ensues. We establish that (for small z) after Player A makes her initial demand she sticks with it until B makes a counterdemand. Once B makes a counterdemand subsequent behavior of both players is exactly as in the preceding section (with posterior probabilities of A and B being behavioral being defined in the natural way).

Lemma 5 points out that before player B has spoken, $\mathcal{A}$ is in a particularly delicate situation: if she reveals rationality, she expects B to act like the most aggressive behavioral type (because he is in a winning position no matter what types he imitates). The only exception to this expectation is if matters are even worse for $\mathcal{A}$, because the equilibrium expectation if B responds to $\mathcal{A}$'s revealing rationality by revealing rationality himself, gives $\mathcal{A}$ less than $1 - \bar{b}$ (that is, the equilibrium expectation is a particularly adverse selection for $\mathcal{A}$ from the set of equilibria of the full information subgame[16]). Lemma 7 verifies that consequently, $\mathcal{A}$ does not change her demand before B speaks.

Lemma 5. *If player $\mathcal{A}$ (who chooses $a \in \mathbb{A}$ at $t = 0$) reveals rationality before B makes a counterdemand, then $\mathcal{A}$'s continuation payoff is at most $1 - \bar{b}$.*

[16]Notice that demands in that subgame are simultaneous, so the uniqueness result of Rubinstein (1982) does not apply.

We seek to characterize equilibrium payoffs in the reduced game when $z = (z^A, z^B) \downarrow 0$. Sometimes we write $\varphi^A(a|z)$ and $\varphi_k^B(b, t|a, z)$ to make explicit the dependence on z.

Lemma 6 establishes that neither type of $\mathcal{B}$ gives more than vanishing weight to any counterdemand that is more aggressive than the balanced counterdemand for the more patient type of player B.

Lemma 6. *For any $R > 1$ and $\epsilon > 0$, there exists $\bar{z} > 0$ such that for all $z \in K(R, \bar{z})$, for any equilibrium (φ^A, φ^B) of $\Gamma(r, \beta, z)$, and for any $a \in \mathbb{A}$, if $b > \lfloor b_2^*(a) \rfloor$ then $\varphi_k^B(b, t|a, z) \leq \epsilon$ for all $t \in (0, \bar{T}]$ and $k = 1, 2$.*

Lemma 7. *For any $R > 1$ there exists $\bar{z} > 0$ such that for any $z \in K(R, \bar{z})$ and any equilibrium of $\Gamma(r, \beta, z)$, player $\mathcal{A}$ who chooses $a \in \mathbb{A}$ at $t = 0$, never reveals rationality before (or at the same time as) B makes a counterdemand.*

In all that follows, we assume that R, $\bar{z}$ and z are such that Lemma 7 is valid.

Consider a particular equilibrium of the game, and the subgame after A has made some demand $a \in \mathbb{A}$. In the subgame, there are expected discounted equilibrium payoffs v_1 and v_2 for the impatient and patient types of $\mathcal{B}$, respectively, discounted to time zero. Because $\mathcal{B}_2$ could adopt $\mathcal{B}_1$'s equilibrium strategy if he wanted to, $v_2 \geq v_1$. In the discussion below, we assume that $\mathcal{A}$ does not reveal rationality before B makes a counterdemand. This assumption was justified in Lemma 6 above. If $\mathcal{B}_k$ waits until some positive time t to make a demand $b \in \mathbb{B}$ in response to a, he must expect, if (b, t) is in the support of his equilibrium

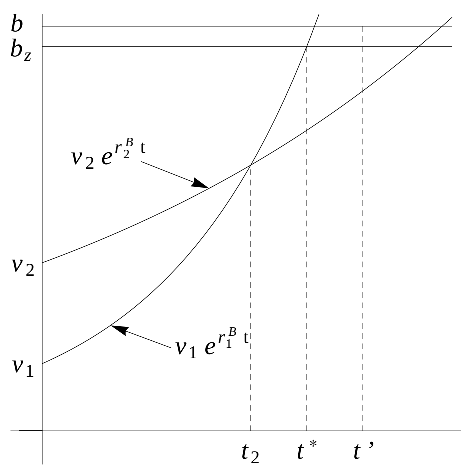

Figure 1

strategy in this subgame, to receive a payoff of $v_k e^{r_k^B t}$ (discounted to t) by doing so. Figure 1 shows the "indifference curves" for the respective types of player B, and some particular demand $b \in \mathbb{B}$ they might consider making. Note that if neither rational type of B ever demands b at time t, then if A observes (b, t), she concludes that she faces a behavioral type and concedes immediately. In this situation, the payoff $\mathcal{B}_k$ would receive if he deviated to making the demand (b, t), would be

$$b_z = (1 - \hat{z}^A)b + \hat{z}^A(1 - a)$$

(because if he is not conceded to, he waits an instant and concedes himself). Figure 1 also shows b_z. *We are assuming here that b and b_z lie above the point of intersection of the two indifference curves.*

In Figure 1, t^* labels the time at which $\mathcal{B}_1$'s indifference curve cuts b_z, and t_2 the time at which the two indifference curves intersect. (Eventually we shall introduce a time $t_1 < t_2$.) Consider times such as t' after t^* but before $\mathcal{B}_2$'s indifference curve cuts b_z. Player $\mathcal{B}_1$ will never demand b at time t' (even immediate acceptance by rational A would be insufficient to give him his equilibrium payoff). But (b, t') must be in the support of $\mathcal{B}_2$'s equilibrium strategy: if it were not, deviating to it would yield $\mathcal{B}_2$ a payoff of b_z, since it would be taken by A as evidence that B was behavioral.

We see that to the right of t^*, the equilibrium must provide a "payoff ramp" that keeps $\mathcal{B}_2$ on his level set. How is this accomplished? When $\mathcal{B}_2$ asks for b at t', there may be a concession by one side, followed by a WOA constructed so that each side is indifferent about conceding at any time. If $\mathcal{B}_2$'s payoff is ramping up to the right of t^*, it must be that A is conceding to B with increasing probability as B waits longer before speaking.

Given the single-crossing nature of the level sets, a natural question is: "Is there a fully separating equilibrium in the subgame, in which to the left of t_2, instantaneous concession probabilities by A rise at the rate that keeps type $\mathcal{B}_1$ indifferent, and to the right, at the rate that keeps type $\mathcal{B}_2$ indifferent?" Unfortunately, things are not that simple. Player $\mathcal{B}_2$ would have a profitable deviation: ask for b at some t'' slightly to the left of t_2, and get not only the concession payoff that $\mathcal{B}_1$ would receive there, but also the advantage of playing a WOA in which A's concession rate is calculated to keep the less patient $\mathcal{B}_1$ indifferent ($\mathcal{B}_2$ therefore receives surplus by playing this WOA, and this is a bonus to $\mathcal{B}_2$ beyond the payoff that $\mathcal{B}_1$ gets). The probability of getting this bonus when speaking at t'' is substantial and does not vanish as t'' approaches t_2, which is in conflict with the fact that by definition, the indifference curves of the two types cross at t_2. Here we say that $\mathcal{B}_2$ is "sneaking in" and playing the WOA (against an unsuspecting player A). We remark that $\mathcal{B}_1$ has no incentive to sneak in to the right of t_2: the slow WOA that $\mathcal{A}$ fights with $\mathcal{B}_2$ does not interest him.

Notice that (b, t'') must be in the equilibrium support of both types. (If it were in the support of $\mathcal{B}_2$ only, then $\mathcal{B}_2$ would get no bonus from the WOA, and his entire payoff would come from the concession A gives him, which type $\mathcal{B}_1$ could collect just as well as $\mathcal{B}_2$; then $\mathcal{B}_1$ and $\mathcal{B}_2$ would have the same payoffs at t'', a contradiction. We have already seen why it cannot be in the support of $\mathcal{B}_1$ only.) If A, upon hearing the demand b at t'', thinks it much more likely to have been said by $\mathcal{B}_1$ than type $\mathcal{B}_2$, then in the event that a WOA ensues (rather than a concession by A), it will have a long initial phase in which A concedes at a rate that would keep $\mathcal{B}_1$ indifferent; this is a major bonus for $\mathcal{B}_2$, and leaves him above his equilibrium indifference curve. On the other hand, if A thinks (b, t'') is sufficiently more likely to have been said by $\mathcal{B}_2$, the WOA will have a very brief first stage, which gives $\mathcal{B}_2$ such a small bonus that his expected payoff is below his equilibrium indifference curve. We show in the Densities section of the Appendix that there exist unique densities for $\mathcal{B}_1$ and $\mathcal{B}_2$ such that each type's payoff lies on the respective indifference curve.

Return for a moment to the fiction that player A believes, upon hearing the demand b any time before t^*, that there is no chance it was made by $\mathcal{B}_2$. We can plot the curve showing the expected utility $\mathcal{B}_2$ would get if he were to sneak in under this circumstance; we call this the *sneaking-in curve.* Denote the sneaking-in value at t by $\underline{v}_2(t)$. (We do not make explicit the dependence of the sneaking-in function on b; the relevant b will be clear from context.) Lemma 8 establishes that it is steeper than $\mathcal{B}_2$'s equilibrium indifference curve; Figures 2 and 3 below illustrate the two possible cases (in case 1, the sneaking-in curve intersects $\mathcal{B}_2$'s indifference curve to the right of the vertical axis, at the time we shall call t_1; in case 2, the sneaking-in curve cuts the vertical axis above v_2, in which case we say $t_1 = 0$).

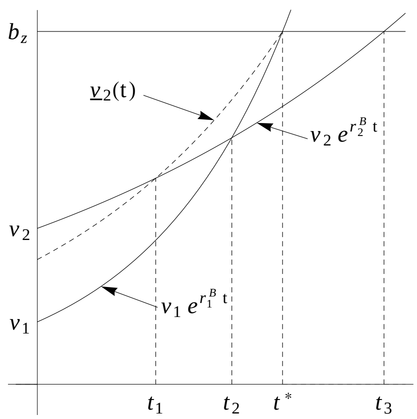

Figure 2: $v_1 < v_2$ and $v_2 > \underline{v}_2(0)$.

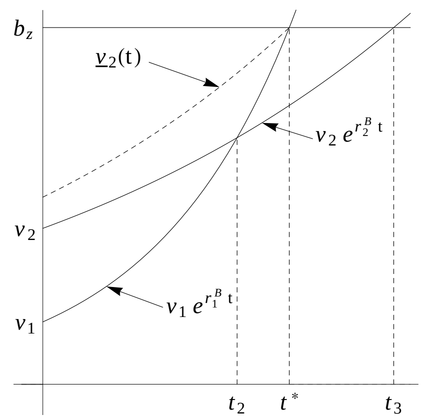

Figure 3: $v_1 < v_2$ and $v_2 \leq \underline{v}_2(0)$.

Lemma 8. *(i)* $v_1 e^{r_1^B t} < \underline{v}_2(t)$ *for* $t < t^*$ *and* $v_1 e^{r_1^B t^*} = \underline{v}_2(t^*)$*; and (ii)* $\underline{v}_2(t)$ *is steeper than* $v_2 e^{r_2^B t}$ *for all* t *such that* $v_2 e^{r_2^B t} \leq \underline{v}_2(t)$.

The analysis thus far has focused on the "horizontal" aspects of behavior in (t, b) space, that is, for a given demand b, at what times will each type of player $\mathcal{B}$ be active? Much more detailed distributional analysis will follow. But we turn at the moment to the complementary "vertical" question: which demands b will be used by $\mathcal{B}_1$ and $\mathcal{B}_2$, respectively? As z approaches 0, strong results emerge here.

Recall that $\lfloor b_k^*(a) \rfloor$ is the counterdemand by $\mathcal{B}_k$ that would be "slightly generous" in response to A's demand a, in a game in which $\mathcal{B}_k$ is the only type of rational B. As before, by "slightly generous" we mean the least generous response that is more generous than balanced. In the temporal setting, this is still true for $\mathcal{B}_2$. But even for a small z, $\mathcal{B}_1$ might demand either $\lfloor b_2^*(a) \rfloor$, or very near the origin, $\lfloor b_1^*(a) \rfloor$. All of this is established in Lemma 10 (apart from the case already covered by Lemma 6).

Lemma 9 is a key component of the proof of Lemma 10 and other subsequent results. Suppose z is small and consider s_0 such that $\mathcal{B}_j$ does not demand $\lfloor b_k^*(a) \rfloor$ after s_0. Then Lemma 9 asserts that $\mathcal{B}_k$'s equilibrium payoff ramp at s_0 is higher than or only slightly below $\lfloor b_k^*(a) \rfloor$. Obviously this implies that if the ramp is below $\lfloor b_k^*(a) \rfloor$ at s_0, it can remain below for at most a very short subsequent interval. Why might one have suspected this, even before reading Lemma 9? Suppose (for contradiction) that $\mathcal{B}_k$'s payoff had been below $\lfloor b_k^*(a) \rfloor$ for a substantial interval following s_0, meaning that for the early part of the interval, A's probability of conceding when B demands $\lfloor b_k^*(a) \rfloor$ is *not* close to 1. This in turn implies that in the war of attrition that ensues when A does not concede, A's initial nonconcession gives A only a relatively modest reputational boost. But, since B is demanding less than the balanced counteroffer to a, his reputation grows faster than A's. For small z, the only way for their reputations to arrive at 1 at the same time (given that A's initial concession is *not* with probability almost 1) is for B to be present at each point (except toward the end of the interval) with extremely high density (exploding as z approaches 0). But, the integral over the interval would then explode as z approaches 0, contradicting the possibility of the interval being of substantial length.

Lemma 9. *For any $R > 1$ and $\epsilon > 0$, there exists $\bar{z} > 0$ such that for all $z \in K(R, \bar{z})$, for any equilibrium (φ^A, φ^B) of $\Gamma(r, \beta, z)$, and for any $a \in \mathbb{A}$ with $\varphi^A(a|z) \geq \epsilon$, if for a given k and for $j \neq k$, $\varphi_j^B(\lfloor b_k^*(a) \rfloor, s|a, z) = 0$ for all $s \geq s_0$, then $v_k e^{r_k^B(s_0+\epsilon)} \geq \lfloor b_k^*(a) \rfloor$.*

Lemma 10. *For any $R > 1$ and $\epsilon > 0$, there exists $\bar{z} > 0$ such that for all $z \in K(R, \bar{z})$, for any equilibrium (φ^A, φ^B) of $\Gamma(r, \beta, z)$, and for any $a \in \mathbb{A}$ with $\varphi^A(a|z) \geq \epsilon$,*

(i) if $b < \lfloor b_k^*(a) \rfloor$ then $\varphi_k^B(b, t|a, z) = 0$ for all $t \in (0, \bar{T}]$, $k = 1, 2$.
(ii) if $\lfloor b_1^*(a) \rfloor < b < \lfloor b_2^*(a) \rfloor$ then $\varphi_1^B(b, t|a, z) \leq \epsilon$ for all $t \in (0, \bar{T}]$.
(iii) $\varphi_1^B(\lfloor b_1^*(a) \rfloor, t|a, z) = 0$ for all $t \in [\epsilon, \bar{T}]$.

The next two results follow from Lemmas 6 and 10.

Corollary. *Under the conditions of the previous lemma, for any $a \in \mathbb{A}$, if $\lfloor b_2^*(a) \rfloor > 1 - a$ then*

$$\int_0^{\bar{T}} \varphi_2^B(\lfloor b_2^*(a) \rfloor, t|a, z)dt \geq 1 - \bar{T}|\mathbb{B}|\epsilon.$$

Corollary. *Under the conditions of the previous lemma, for any $a \in \mathbb{A}$ such that $\lfloor b_2^*(a) \rfloor = 1 - a$,*

$$|v^A(a|z) - a| \leq \epsilon \quad \textit{and} \quad |v_k^B(a|z) - (1 - a)| \leq \epsilon \quad k = 1, 2.$$

The fact that almost all of $\mathcal{B}_2$'s weight is devoted to $\lfloor b_2^*(a) \rfloor$, and that any of $\mathcal{B}_1$'s that is not devoted to $\lfloor b_2^*(a) \rfloor$ almost always goes on $\lfloor b_1^*(a) \rfloor$ near the origin, suggests that we can work with a reduced game where both players have severely restricted strategy spaces, in the vertical dimension. We define and study that game now. We will show that as z approaches 0, equilibria in the reduced game and the true game are essentially unique, and coincide.

The Reduced Game.

Denote by $\tilde{\Gamma}(r, \beta, z)$ the "reduced game" in which after A makes an initial demand $a \in \mathbb{A}$, she cannot move until B makes a demand (see Lemma 7), and the set of B's behavioral types is $\{(\lfloor b_2^*(a) \rfloor, t) \mid t \in [0, \bar{T}]\}$. Note that the latter set depends on the choice a. As part of the specification of the reduced game, if $\lfloor b_2^*(a) \rfloor = 1 - a$, then the game ends with B accepting the demand a immediately. When $\lfloor b_2^*(a) \rfloor > 1-a$, the prior probability of type $(\lfloor b_2^*(a) \rfloor, t)$ is $z^B \pi^B(\lfloor b_2^*(a) \rfloor, t)$. Player B can only counterdemand $\lfloor b_1^*(a) \rfloor$ at time 0, and if he does, this demand is accepted with probability 1 immediately by A. After time 0, B can only counterdemand $\lfloor b_2^*(a) \rfloor$, and thereafter the players play a war of attrition.

We proceed to analyze $\tilde{\Gamma}(r, \beta, z)$. The first step is to analyze the subgame $\tilde{\Gamma}(r, z, a)$ that arises after A chooses some $a \in \mathbb{A}$ with probability $\varphi^A(a) > 0$ (in equilibrium). By definition, when $\lfloor b_2^*(a) \rfloor = 1-a$, we have immediate agreement. Hereafter we assume $\lfloor b_2^*(a) \rfloor > 1-a$.

Fix an equilibrium of this subgame and suppose $\mathcal{B}_k$'s payoff in the subgame is v_k. As $\mathcal{B}_k$ is limited in the reduced game (for $t > 0$) to the counterdemand $\lfloor b_2^*(a) \rfloor$, he chooses only the time at which he makes the counterdemand. Accordingly, we will simply write $\varphi_k^B(t)$ instead of $\varphi_k^B(\lfloor b_2^*(a) \rfloor, t|a)$. Since $r_1^B > r_2^B$, it follows immediately that $v_2 \geq v_1 \geq \lfloor b_1^*(a) \rfloor$. Since $\mathcal{B}_k$ must be indifferent among all times t for which $\varphi_k(t) > 0$, his continuation value at any such t (discounted to t) must be $v_k e^{r_k t}$.

In the reduced game, let $b_z = (1-\hat{z}^A)\lfloor b_2^*(a) \rfloor + \hat{z}^A(1-a)$. We argue via a series of observations that $b = \lfloor b_2^*(a) \rfloor$ and b_z lie above the point of intersection of the two indifference curves as depicted in Figures 1, 2 and 3. If B's counterdemand is accepted immediately by $\mathcal{A}$ with probability 1, then B's expected value is b_z, and strictly less otherwise. Let t^* solve $v_1 e^{r_1^B t} = b_z$, t_2 be the time when the two payoff curves $v_1 e^{r_1^B t}$ and $v_2 e^{r_2^B t}$ intersect, and t_3 solve $v_2 e^{r_2^B t} = b_z$. By assumption, $t^* < \bar{T}$ and $t_3 < \bar{T}$. Furthermore, $\varphi_1^B(t) = 0$ for $t \geq t^*$ and $\varphi_2^B(t) = 0$ for $t \geq t_3$. Also, if $\varphi_1^B(t) = \varphi_2^B(t) = 0$ for some $t \in [0, \bar{T}]$, $\mathcal{B}_k$'s payoff from counterdemanding $\lfloor b_2^*(a) \rfloor$ at time t (out of equilibrium) is b_z.

Observation 1: In any equilibrium, we must have

$$\Phi_k^B = \int_0^{\bar{T}} \varphi_k^B(t)dt = 1 \quad k = 1, 2,$$

unless $v_k = \lfloor b_1^*(a) \rfloor$, in which case we may have $\Phi_k^B < 1$ combined with $\mathcal{B}_k$ making the demand $\lfloor b_1^*(a) \rfloor$ at time $t = 0$ with probability $1 - \Phi_k^B$.

Observation 2: Assume $v_i e^{r_i^B t} < v_j e^{r_j^B t}$ and $v_i e^{r_i^B t} < b_z$. Then $\varphi_i^B(t) > 0$. Moreover, if $i = 2$, then $\varphi_1^B(t) = 0$.

When $v_1 = v_2$ we obtain the configuration depicted in Figure 4 below. In this case, by Observation 2 $\varphi_1^B(t) = 0$ for all $t \in (0, t^*)$. Hence, $\Phi_1^B = 0$, so $\mathcal{B}_1$ makes the demand $\lfloor b_1^*(a) \rfloor$ at time 0 with probability 1, and this demand is accepted by A immediately (as per the rules of the reduced game). Hence $v_1 = v_2 = \lfloor b_1^*(a) \rfloor$. For small z, this possibility is ruled out by Lemma 9.

Hereafter we assume that $v_2 > v_1$. Clearly $v_2 < b_z$, for if $v_2 = b_z$, then $v_1 = b_z$ also (since $v_2 = b_z$ is possible only if the counterdemand $\lfloor b_2^*(a) \rfloor$ at $t = 0$ is accepted immediately by $\mathcal{A}$ with probability 1). But then we are in the case dealt with above where we concluded that $v_1 = v_2 = \lfloor b_1^*(a) \rfloor < b_z$, a contradiction. The case $b_z > v_2 > v_1$ is depicted in Figures 2–3.

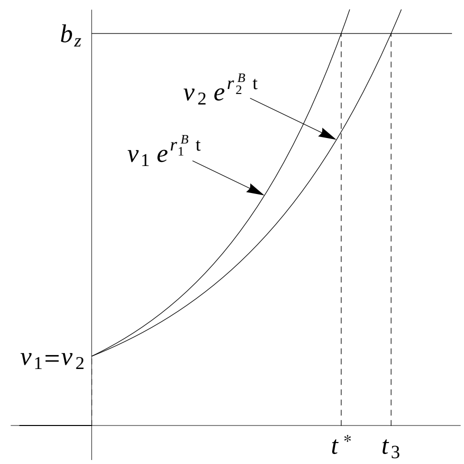

Figure 4: $v_1 = v_2$.

Observation 3: If $\varphi_k^B(t) > 0$ for $k = 1, 2$, then $v_1 e^{r_1^B t} < v_2 e^{r_2^B t} < \underline{v}_2(t)$.

Observation 4: $t_2 < t^* < t_3$.

By Lemma 8, $v_2 e^{r_2^B t}$ can intersect $\underline{v}_2(t)$ only from above, and consequently at most once. Hence any solution can have only one of the two configurations presented above in Figures 2 and 3. In the first configuration $v_2 e^{r_2^B t}$ intersects $\underline{v}_2(t)$ at some $t_1 \in (0, t_2)$. The second configuration corresponds to the case when $v_2 < \underline{v}_2(0)$ and $v_2 e^{r_2^B t}$ does not intersect $\underline{v}_2(t)$ in $(0, t_2)$. In this case we define $t_1 = 0$. In both configurations $v_2 > \lfloor b_1^*(a) \rfloor$, and Observation 1 then implies that $\Phi_2^B = 1$. Lemma 11 establishes which type is active in each region.

Lemma 11. *$\varphi_1^B(t) > 0$ and $\varphi_2^B(t) = 0$ for $t \in (0, t_1)$, $\varphi_1^B(t) > 0$ and $\varphi_2^B(t) > 0$ for $t \in (t_1, t_2)$, and $\varphi_1^B(t) = 0$ and $\varphi_2^B(t) > 0$ for $t \in (t_2, t_3)$.*

For a given z, suppose that $\mathcal{A}$ chooses a with some probability $\varphi^A(a|z) > 0$, and let $\hat{z}^A$ be defined as in (1). Then, for any (v_1, v_2) in the relevant range, not necessarily corresponding to equilibrium values in the subgame following the demand a, we may compute the densities with which $(\lfloor b_2^*(a) \rfloor, t)$ is chosen by $\mathcal{B}_1$ and $\mathcal{B}_2$ respectively so that the continuation values $v_k e^{r_k t}$, $k = 1, 2$ are achieved in the subgame following the demand of $\lfloor b_2^*(a) \rfloor$ at time t. These densities are denoted $\varphi_k^A(t|\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)$ and the corresponding integrals by $\Phi_k^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)$.[17]

Similarly, even for a v_1 not associated with equilibrium, for any a and corresponding b_z one can define t_1, t_2, t_3 and t^* as before. Lemma 12 says that if $\mathcal{B}_1$'s and $\mathcal{B}_2$'s indifference

[17] In the reduced game the set of behavioral types in the subgame following the demand a is constrained. But in the full game any counterdemand $b \in \mathbb{B}$ with $a + b > 1$ is possible, so in general, the densities will depend both on a and b.

curves through v_1 and v_2 cross just below b_z, then t_1, t_2, t_3 and t^* are all very close to one another.[18] This is always the case when z is close to 0: $\mathcal{B}_2$ is alone from t_2 to t_3, and as z approaches 0, the strength of counterdemand b against a becomes overwhelming, necessitating that the interval (t_2, t_3) be extremely short (see the discussion preceding Lemma 9). In this case, a Figure 3 equilibrium is "Coasean", in the sense that there is almost no delay to agreement, and that v_1 and v_2 are each close to b_z, that is, both types of $\mathcal{B}$ do virtually as well as $\mathcal{B}$ would if it were known that only the stronger of the two rational types, $\mathcal{B}_2$, were present (see Lemma 12). By contrast, a Figure 2 equilibrium is non-Coasean: there is noticeable delay before the pooling region is reached and an immediate demand/concession by $\mathcal{B}_1$ becomes a possibility. Recall that in the reduced game, an immediate counterdemand by B of $\lfloor b_1^* \rfloor$ is accepted by $\mathcal{A}$ immediately.[19] Furthermore, $\lfloor b_1^*(a) \rfloor = 1 - a$ for $a \le a_1^*$; for such values of a, $\mathcal{B}_1$ is effectively conceding to A's demand.

Lemma 12. *For all $a \in A$ and $\epsilon > 0$, there exists $\delta > 0$ such that for any $v_1 \in [1-a, b_z)$ and v_2 with $v_2 e^{r_2^B t^*} \in [b_z - \delta, b_z)$, $|t_i - t^*| \le \epsilon$ for $i = 1, 2, 3$. Furthermore, if $v_2 \le \underline{v_2}(0)$ then $t_1 = 0$ and $v_1 \ge b_z - \epsilon$.*

We noted earlier that for "target values" v_1 and v_2, there are unique densities φ_1 and φ_2 that will result in exactly those interim beliefs (of player A) that lead to the desired targets for the respective types. This takes only "local" considerations into account: φ_2, for example, is constructed according to what is needed point by point. Why should we expect it to integrate to 1 (as it should, being type 2's mixed strategy)? This sort of "global" consideration is crucial in determining whether equilibrium is Coasean (Figure 3) or non-Coasean (Figure 2). We investigate this more closely in what follows.

Consider a non-Coasean equilibrium. The support of $\mathcal{B}_2$'s strategy is $[t_1, t_3]$, and because $\mathcal{B}_1$ is alone (and hence weak) in $(0, t_1)$, even he must be extremely scarce there. This is established in Lemma 13 below. It also shows that in equilibrium, for small z, $v_2 e^{r_2^B t^*}$ is close to b_z. This property is invoked in Lemmas 14 and 15.

Lemma 13. *For any $a \in \mathbb{A}$, $R > 1$, $\epsilon > 0$ and $\delta > 0$, there exists $\bar{z} > 0$ such that if $z \in K(R, \bar{z})$, and $\varphi^A(a|z) \ge \epsilon$, then for all $v_1 \in [\lfloor b_1^*(a) \rfloor, b_z)$ and $v_2 \in [v_1, b_z)$,*

$$\int_0^{t_1} \varphi_1^B(t|\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)dt < \delta.$$

Furthermore, if $\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2) = 1$, then $v_2 e^{r_2^B t^*} \ge b_z - \delta$.

Return to our consideration of non-Coasean equilibrium. It follows from the preceding discussion that for z close to 0, we may focus on the integrals of φ_1 between t_1 and t^*, and of φ_2 between t_1 and t_3. The latter integral must be 1, whereas the former could be less than 1, with the residual (essentially) accounted for by an instantaneous

[18]Note that this does *not* mean t_1 is necessarily also close to (or coincident with) 0; in a Figure 2 equilibrium, it is typically not close to 0.

[19]The rationale for this aspect of the reduced game is provided by *(i)* and *(iii)* of Lemma 10.

demand/concession[20] with positive probability at $t = 0$. Crucially, the *ratio* of these two integrals is stable for all values of v_1 and v_2 in the range of interest, approaching a limit denoted $\beta_2\theta(a)/\beta_1$ as v_1 and v_2 cross close to b_z; this is the content of Lemma 14 below.

Think for a moment of the special case where $\beta_1 = \beta_2$ (the rational types of B are equally likely). If $\theta(a) > 1$, a non-Coasean equilibrium is impossible: it would require that $\mathcal{B}_1$'s mixed strategy integrate to strictly more than 1. More generally, for arbitrary β's, if the ratio of integrals $\beta_2\theta(a)/\beta_1 > 1$, non-Coasean equilibrium is impossible. Put informally, once the point-by-point needs for $\mathcal{B}_1$'s presence (relative to $\mathcal{B}_2$) are aggregated, they cannot be accommodated given the proportions of the two types actually available.

Conversely, Lemma 15 implies that when $\beta_2\theta(a)/\beta_1 < 1$, a Coasean equilibrium is impossible. To see why, note that $\beta_2\theta(a)/\beta_1 < 1$ would require (by Lemma 15) type 1 to be ending the game at time 0 with positive probability, meaning his utility is $\lfloor b_1^*(a)\rfloor < \lfloor b_2^*(a)\rfloor$. This contradicts the fact that in a Coasean equilibrium, $v_1 \uparrow b_z$ as $z \downarrow 0$.

To summarize, if $\beta_2\theta(a)/\beta_1 < 1$, any equilibrium must be non-Coasean (and $\mathcal{B}_1$'s utility will be $\lfloor b_1^*(a)\rfloor$), whereas if $\beta_2\theta(a)/\beta_1 > 1$, equilibrium must be Coasean, and both types get almost $\lfloor b_2^*(a)\rfloor$, and since this happens almost immediately, player A's utility is determined as well. Thus, if there are multiple Coasean equilibria, their payoffs must virtually agree. The same is true for non-Coasean equilibria: since all of them share the same value of v_1 (because in each case ending the game at the origin is in the support of $\mathcal{B}_1$'s equilibrium strategy), there is essentially only one value of v_2 such that the indifference curves through v_1 and v_2 intersect almost at b_z. That determines the time (the same across equilibria) to which all the t_k's converge, which is one of two things determining A's utility. The other is the probability $\mathcal{B}_1$ ends the game at the origin. By Lemma 1, this probability must also be the same across all non-Coasean equilibria. Thus, the reputational perturbations have, for any given parameters, asymptotically produced unique payoff predictions for each player and type.

The function θ introduced in Lemma 14 is central to our characterization results. It depends on $r^A, r_1^B, r_2^B, \mathbb{B}$ and, of course, $a \in \mathbb{A}$, but is *independent* of $z^A, z^B, \Pi, \beta_1, v_1$ and v_2, modulo the conditions stated below. The ratio in the first term of the LHS of the inequality in Lemma 14 compares the pooling probability of $\mathcal{B}$'s impatient type to the total probability of his patient type.

Lemma 14. *There exists a function $\theta : \mathbb{A} \to \mathbb{R}_+$ such that for all $\beta_1 \in (0,1)$ (and $\beta_2 = 1 - \beta_1$), $a \in \mathbb{A}$, $R > 1$, $z \in K(R)$, and $\epsilon > 0$, there exists $\delta > 0$ such that for all $v_1 \in [1-a, b_z)$ and v_2 with $v_2 e^{r_2^B t^*} \in [b_z - \delta, b_z)$ and $v_2 \geq \underline{v}_2(0)$,*

$$\left| \frac{\int_{t_1}^{t_2} \varphi_1^B(t|\hat{z}^A, z^B, a, \lfloor b_2^*(a)\rfloor, v_1, v_2)dt}{\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a)\rfloor, v_1, v_2)} - \frac{\beta_2}{\beta_1}\theta(a) \right| < \epsilon.$$

Let $\Phi_1^*(a) = \beta_2\theta(a)/\beta_1$. As per the earlier discussion, $\Phi_1^*(a)$ is (in the limit) the probability that the impatient player $\mathcal{B}_1$ pools with the patient player $\mathcal{B}_2$ by counterdemanding $\lfloor b_2^*(a)\rfloor$ near time t^*. Clearly $\Phi_1^*(a)$ is also a function of β_1 (and $\beta_2 = 1-\beta_1$), but

[20] See discussion preceding Lemma 12.

for simplicity this dependence is omitted. Note that θ is *not* a function of β_1. Note also the obvious fact that $\Phi_1^*(a)$ increases from 0 to ∞ as β_1 decreases from 1 to 0. Indeed we state the Lemma in terms of $\theta(a)$ instead of $\Phi_1^*(a)$ precisely to emphasize that, depending on the proportion of impatient types, $\Phi_1^*(a)$ could be greater than or less than 1.

Lemma 15. *Consider any $\beta_1 \in (0,1)$, $a \in \mathbb{A}$, $R > 1$, and $z \in K(R)$. Suppose $\Phi_1^*(a) < 1$. Then there exists $\delta > 0$ such that for all $v_1 \in [1-a, b_z)$ and v_2 with $v_2 e^{r_2^B t^*} \in [b_z - \delta, b_z)$ and $v_2 \leq \underline{v}_2(0)$,*

$$\frac{\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)}{\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)} < 1.$$

Recall the conditions of Lemma 10. We will argue that when $\varphi^A(a|z) \geq \epsilon$, then equilibrium in the subgame is "essentially" unique and consequently falls in one of the two cases described above. Furthermore, in case $v_2 > \underline{v}_2(0)$ (Figure 2), $v_1 = \lfloor b_1^*(a) \rfloor$ and $\mathcal{B}_1$ counterdemands $\lfloor b_1^*(a) \rfloor$ at time 0 with probability close to $1 - \Phi_1^*(a)$, where $\Phi_1^*(a)$ is the limit value of $\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)$ as $z \downarrow 0$. Let $\bar{t}^*$ be the limit of t^* and $\bar{v}_2^*$ be the limit of v_2 as $z \downarrow 0$. Then $\lfloor b_1^*(a) \rfloor = e^{-r_1^B \bar{t}^*} \lfloor b_2^*(a) \rfloor$ and $\bar{v}_2^* = e^{-r_2^B \bar{t}^*} \lfloor b_2^*(a) \rfloor$. Furthermore, $e^{-r^A \bar{t}^*} = [\lfloor b_1^*(a) \rfloor / \lfloor b_2^*(a) \rfloor]^{r^A / r_1^B} \equiv D$. In case $v_2 \leq \underline{v}_2(0)$ (Figure 3), $v_1 \approx v_2 \approx \lfloor b_2^*(a) \rfloor$ and $\mathcal{A}$ concedes to $\lfloor b_2^*(a) \rfloor$ with probability close to 1. The preceding discussion implicitly assumes that $a > a_2^*$. When $a \leq a_2^*$, $b_2^*(a) = 1 - a$, and the reduced game is specified to end immediately with $\mathcal{B}$ accepting A's demand a. Consequently, player $\mathcal{A}$'s asymptotic payoff is

$$v^A(a) = \begin{cases} (\beta_1 - \beta_2\theta(a))(1 - \lfloor b_1^*(a) \rfloor) + \beta_2(1+\theta(a))D(1 - \lfloor b_2^*(a) \rfloor) & \text{if } a > a_2^* \text{ and} \\ & \beta_2\theta(a)/\beta_1 < 1 \\ 1 - \lfloor b_2^*(a) \rfloor & \text{otherwise.} \end{cases}$$

In Theorem 2 below we revert to the true game $\Gamma(r, \beta, z)$. The main differences between $\Gamma(r, \beta, z)$ and $\tilde{\Gamma}(r, \beta, z)$ are that in $\Gamma(r, \beta, z)$, many types (b, t) with $b \neq \lfloor b_2^*(a) \rfloor$ will also be mimicked with positive density, but by Lemma 10 these densities will go to zero with z. Furthermore, $\lfloor b_1^*(a) \rfloor$ will only be mimicked (if at all) in a small initial interval. In particular:

(1) $\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2) \geq 1 - \epsilon$ (and does not exactly equal 1),
(2) if $\Phi_1^*(a) < 1$ then $|\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2) - \Phi_1^*(a)| \leq \epsilon$.
(3) if $\Phi_1^*(a) > 1$ then $\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2) \geq 1 - \epsilon$,

where $\epsilon \downarrow 0$ as $z \to (0,0)$. It is clear that the formulae for density functions do not depend upon whether we are in $\Gamma(r, \beta, z)$ or $\tilde{\Gamma}(r, \beta, z)$ and neither do the results regarding the ratios $[\Phi_1^B / \Phi_2^B]$.

Theorem 2. *For any $R > 1$ and $\epsilon > 0$ there exists $\bar{z} > 0$ such that the following is true. Suppose $z \in K(R, \bar{z})$ and that φ is an equilibrium of $\Gamma(r, \beta, z)$. For a given $a \in \mathbb{A}$, assume*

$\varphi^A(a) \geq \epsilon$, and let v_k (resp. u^A) be the corresponding continuation value for $\mathcal{B}_k$ (resp. $\mathcal{A}$) induced by φ in the subgame after a is chosen.

(i) If $\Phi_1^*(a) < 1$, then $||b_1^*(a)| - v_1| \leq \epsilon$, $|\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^* \rfloor, v_1, v_2) - \Phi_1^*(a)| \leq \epsilon$, and $\mathcal{B}_1$ counterdemands $\lfloor b_1^*(a) \rfloor$ with probability $1 - \Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^* \rfloor, v_1, v_2)$. Furthermore, $|v_2 - \bar{v}_2^*| \leq \epsilon$.

(ii) If $\Phi_1^*(a) > 1$, then $|v_k - \lfloor b_2^*(a) \rfloor| \leq \epsilon$ for $k = 1, 2$, and $\mathcal{A}$ concedes to $\lfloor b_2^*(a) \rfloor$ with some probability in $[1 - \epsilon, 1]$ by time converging to 0.

(iii) Finally, $|u^A - v^A(a)| \leq \epsilon$.

Proof. Choose $\bar{z} > 0$ as required by Lemmas 9, 10 and 13.

(i) By Lemma 10, $\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2) \approx 1$. By Lemmas 10, 13 and 14, $\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2) \approx \Phi_1^*(a)$ and $v_2 e^{r_2^B t^*} \approx b_z$. If $v_2 \leq \underline{v}_2(0)$, then by Lemma 12 (which applies to Γ) $v_1, v_2 \approx \lfloor b_2^*(a) \rfloor$; hence $v_1 > \lfloor b_1^*(a) \rfloor$ and $\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_1^*(a) \rfloor, v_1, v_2)$ $= 0$, and by Lemma 10, $\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2) \approx 1$. Consequently if $\Phi_1^*(a) < 1$, the hypothesis $v_2 \leq \underline{v}_2(0)$ leads to a contradiction of Lemma 15. Hence $v_2 > \underline{v}_2(0)$. Lemma 6 and Lemma 10 then imply that

$$\int_0^{\epsilon} \varphi_1^B(t | \hat{z}^A, z^B, a, \lfloor b_1^*(a) \rfloor, v_1, v_2) \approx 1 - \Phi_1^*(a) > 0.$$

Now, Lemma 9 implies $v_1 \approx \lfloor b_1^*(a) \rfloor$. By Lemma 13, $v_2 \approx e^{-r_2^B \bar{t}^*} \lfloor b_2^*(a) \rfloor = \bar{v}_2^*$.

(ii) If $\Phi_1^*(a) > 1$ then by the above discussion, we must have $v_2 < \underline{v}_2(0)$ and the rest of the characterization follows.

(iii) The preceding discussion clarifies that if $\varphi^A(a) \geq \epsilon$, the payoffs to A in the corresponding subgame of Γ and $\tilde{\Gamma}$ converge. The conclusion follows directly. □

Perturbed Equilibrium: We showed that $\mathcal{A}$'s equilibrium payoff in the subgame following the choice of a is close to $v^A(a)$ under the assumption that $\varphi^A(a|z) \geq \epsilon$ for some $\epsilon > 0$ and z is sufficiently small. These payoff characterizations justify our focus on the "reduced game" and underlie the computation of the limit ratio of integrals denoted $\Phi_1^*(a)$. We unfortunately simply do not have the tools to make precise statements when these conditions are not satisfied. Obviously, for a small fixed ϵ and any z, and in particular small z, some choices made by $\mathcal{A}$ in equilibrium must satisfy $\varphi^A(a|z) \geq \epsilon$. It (almost) follows that (for small z), if $\varphi^A(a|z) > \epsilon$ then a maximizes v^A. But not quite. Suppose, for instance, $\mathbb{A} = \{a_1, a_2\}$ and $v^A(a_1) = .7$ and $v^A(a_2) = .5$. It is logically possible that in equilibrium $\lim_{z \to 0} \varphi^A(a_2|z) = 1$. How is this consistent with $v^A(a_2) < v^A(a_1)$? This could only happen if $\lim_{z \to 0} \varphi^A(a_1|z) = 0$ and the payoff to $\mathcal{A}$ in the subgame is *increasing* in $\varphi^A(a_1|z)$. It is hard to see how this could happen but we have not been able to prove this. To rule out this possibility it is most convenient to consider a natural refinement of equilibrium. This concept is very similar in spirit to a trembling hand perfect equilibrium with the particular feature that only player $\mathcal{A}$ trembles. As $\mathcal{B}$'s behavior is essentially unique (see Theorem 2), there is no need to consider trembles for B. Our refinement yields the conclusion that if $\varphi^{A\ell}(a|z) > \epsilon$, then a maximizes v^A.

Definition: For $0 \leq \epsilon < 1/|\mathbb{A}|$, an ϵ-perturbed equilibrium is a strategy profile in which $\mathcal{A}$ optimizes subject to the constraint that she chooses all $a \in \mathbb{A}$ with probability at least ϵ while $\mathcal{B}_k$'s behavior is fully optimal, $k = 1, 2$. A 0-perturbed equilibrium is just an equilibrium.

Note that Theorem 2 completely pins down buyer behavior after all initial choices $a \in \mathbb{A}$ that are taken with non-negligible probability (for small z). According to Theorem 3, for any $\epsilon > 0$ and for small enough z, if φ is an ϵ-perturbed equilibrium, then $\varphi^A(a) > \epsilon$ only if a maximizes v^A. Moreover, by Theorem 4, there exists an unperturbed equilibrium with precisely this property.

Theorem 3. *For any $R > 1$ and $\epsilon > 0$ there exists $\bar{z} > 0$ such that the following is true. Suppose $z \in K(R, \bar{z})$ and that φ is an ϵ-perturbed equilibrium of $\Gamma(r, \beta, z)$. For a given $a \in \mathbb{A}$, if $\varphi^A(a) > \epsilon$, then a maximizes $v^A(\cdot)$.*

Proof. By Theorem 2, there exists $\bar{z} > 0$ such that the expected payoff from choosing $a \in \mathbb{A}$ is approximately $v^A(a)$. It follows that if the constraint $\varphi^A(a) \geq \epsilon$ is not binding, a must indeed be a maximizer of $v^A(a)$. □

Theorem 4. *For any $R > 1$ and $\epsilon \in (0, 1/[1+|\mathbb{A}|]$ there exists $\bar{z} > 0$ such that the following is true. Suppose $z \in K(R, \bar{z})$ and that φ is an equilibrium of $\Gamma(r, \beta, z)$. If $\varphi^A(a) > \epsilon$ then a maximizes $v^A(\cdot)$.*

Proof. Let $\hat{\mathbb{A}} = \operatorname{argmax} v^A(a)$. Consider a "pseudo" equilibrium in which $\mathcal{A}$ optimizes subject to the constraint $\sum_{a \in \hat{\mathbb{A}}} \varphi^A(a) \geq |\hat{\mathbb{A}}|\epsilon$. Direct adaptation of the arguments of Theorem 7 imply that such an equilibrium exists. Since at least one maximizer of v^A is chosen with probability greater or equal to ϵ, by Theorem 2 (iii) $\mathcal{A}$'s payoff from such a choice is (approximately) at least max $v^A(a)$. Since this maximizer may be chosen with probability 1, $\mathcal{A}$'s payoff in a pseudo equilibrium is at least (approximately) max $v^A(a)$. Suppose $\varphi^A(a) > \epsilon$ for $a \notin \hat{\mathbb{A}}$. Again, by Theorem 2(iii), $\mathcal{A}$'s payoff from choosing such an a is strictly less than max v^A. It follows that in equilibrium $\varphi^A(a) \leq \epsilon$, a contradiction. Thus, a pseudo-equilibrium has the indicated properties. Furthermore, $\varphi^A(a) \leq \epsilon$ for all $a \notin \hat{\mathbb{A}}$ implies $\sum_{a \in \hat{\mathbb{A}}} \varphi^A(a) > |\hat{\mathbb{A}}|\epsilon$. Since this constraint is not binding, our pseudo-equilibrium is in fact a regular equilibrium (with the required properties).

It follows that: (i) if a is not a maximizer of v^A, then in a pseudo equilibrium, $\varphi^A(a) < \epsilon$; (ii) the constraint that a maximizer be chosen with probability at least ϵ is not binding. Consequently, φ is a regular equilibrium with the indicated properties. □

When is $\mathcal{A}$'s demand Coasean?

For some values of the parameters of the bargaining problem, $\mathcal{A}$ can achieve a non-Coasean outcome by making some demand a (that is, demand a would result in an equilibrium as in Figure 2, not Figure 3). That does not mean, however, that it is necessarily optimal to do so. The expected delay to agreement might leave $\mathcal{A}$ with a lower utility than her Coasean payoff (what she would get in a game in which the probability of $\mathcal{B}_1$ were 0). One might guess that if the weaker of the two types of $\mathcal{B}$ is sufficiently more likely (β_1 close enough to 1), $\mathcal{A}$ can demand more than her Coasean payoff and still get immediate agreement with high probability, dominating her option of making a Coasean demand. This is confirmed by Theorem 5, which also proves that if, instead, β_1 is sufficiently close to 0, $\mathcal{A}$ will demand only her Coasean payoff.

Theorem 5. *Let $\tilde{a}^*(\beta_1)$ maximize $v^A(\cdot;\beta_1)$. There exists $\epsilon > 0$ such that $\tilde{a}^*(\beta_1) = \tilde{a}_1^*$ for $\beta_1 \in [1-\epsilon, 1]$ and $\tilde{a}^*(\beta_1) = \tilde{a}_2^*$ for $\beta_1 \in [0, \epsilon]$.*

Proof. Inspection of the formula for v^A reveals that $v^A(a) = 1 - \lfloor b_1^*(a) \rfloor$ for β_1 large enough, and $v^A(a) = 1 - \lfloor b_2^*(a) \rfloor$ for β_1 small enough. The conclusion follows directly. □

Notice that one implication of Theorem 5 is that $\mathcal{A}$'s payoff does not jump downward as β_1 is decreased slightly from 1. Contrast this to the solution in Section 2, where only the upper bound of the support of B's discount rate affects $\mathcal{A}$'s payoff (as $z \to 0$).

What happens when β_1 is not extremely close to 0 or 1? We present some illustrative calculations. It is most convenient to do so in the limit case where the grid of demands is "very fine". Let $v_0^A(a)$ be the limit of $v^A(a)$ as the gridsizes of $\mathbb{A}$ and $\mathbb{B}$ converge to 0. Figures 5–8 plot the functions $v_0^A(a)$, $\Phi_1^*(a)$ and $1 - b_2^*(a)$, as well as the position of the "balanced demands" a_k^*, $k = 1, 2$ for various parameter configurations $(\beta, r^A, r_1^B, r_2^B)$. **Note**: in all cases, $\mathcal{A}$'s discount rate is normalized to 1.

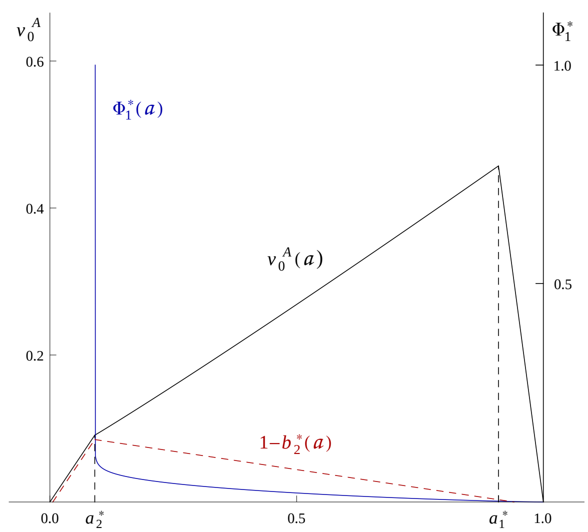

Figure 5: $(\beta_1, r_1^B, r_2^B) = (0.5, 10.0, 0.1)$

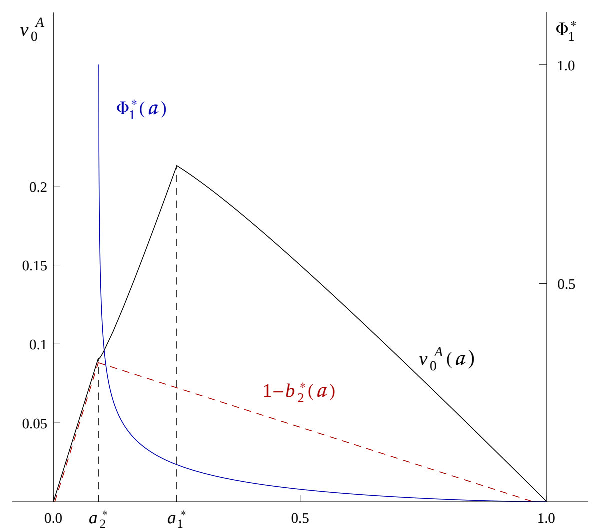

Figure 6: $(\beta_1, r_1^B, r_2^B) = (0.9, 0.333, 0.1)$

Figure 5 demonstrates that when the two types of $\mathcal{B}$ differ drastically, it is disastrous for $\mathcal{A}$ to treat them both as the stronger type, as the Coasean prescription would require. Accordingly, $\mathcal{A}$'s payoff to choosing a_2^* is far below her payoff to a_1^*. In Figure 6, $\mathcal{A}$ is relatively impatient and moreover the two types of $\mathcal{B}$ do not differ drastically. Nevertheless, increasing β_1 sufficiently makes $\mathcal{A}$'s optimal choice non-Coasean, in conformity with

Theorem 5. Figure 7 differs from Figure 6 only in that there is now a modest presence of type 1. This makes it expensive to separate them, so it is better for $\mathcal{A}$ simply to make the Coasean demand immediately. The payoff drop to the right of a_1^* occurs at the point where the unique equilibrium in the subgame switches from Coasean to non-Coasean. At that switch point and immediately to its right, the probability of B conceding at $t = 0$ is too low to compensate $\mathcal{A}$ for waiting (with high probability) until the vicinity of t^*.

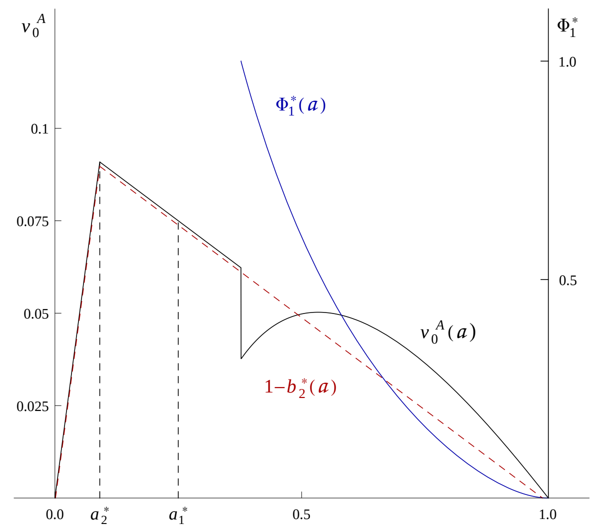

Figure 7: $(\beta_1, r_1^B, r_2^B) = (0.3, 0.333, 0.1)$

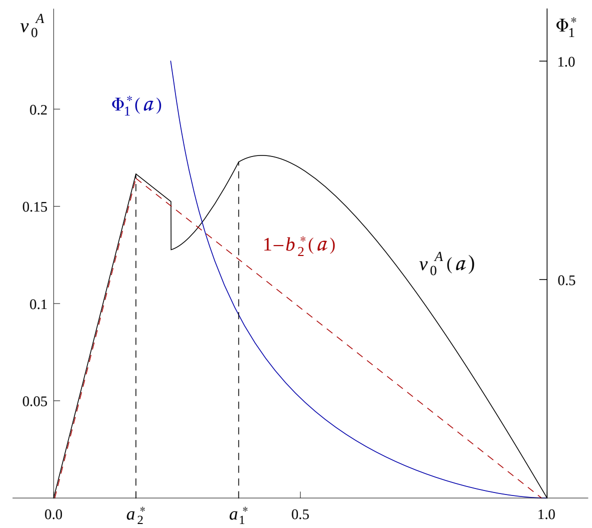

Figure 8: $(\beta_1, r_1^B, r_2^B) = (0.575, 0.6, 0.2)$

It is noteworthy that in all the examples discussed thus far, the optimal choice is either a_1^* or a_2^*. Theorem 6 clarifies that this is not an accident when the optimizer is less than or equal to a_1^*. However, there *are* cases where the optimizer is strictly greater than a_1^*. Figure 8 illustrates this. One might not expect to observe such cases, because increasing a beyond a_1^* *increases* the counterdemands of *both* types of $\mathcal{B}$. But it also increases the probability that type 1 makes his separating counterdemand at the origin (see the declining Φ_1^* graph in Figure 8, which is the complementary pooling probability). In addition, one can show that the increased demand induces a lower t^*, which is again good for $\mathcal{A}$.

Notice that in all the examples $\theta(a) = \Phi_1^*(a)\beta_1/\beta_2$ is *decreasing* in a. While we are not able to establish this property analytically, an extensive grid search in the region $\{(r_1^B, r_2^B) \mid 0.01 \le r_2^B < r_1^B \le 10\}$ (with gridsize equal to 0.01) did not yield a counterexample. When this property is satisfied then it indeed follows that the maximizer of v_0^A is either a_1^* or a_2^* or some $a > a_1^*$, as in the examples above.

Recall that for elementary reasons $\mathcal{A}$ will choose only $a \ge a_2^*$. Note also that equilibrium in the subgame following the choice of $a \in \mathbb{A}$ is Coasean if $\Phi_1^*(a) > 1$ (for small z). Suppose $\Phi_1^*(a_1^*) > 1$. Then θ decreasing implies $\Phi_1^*(a) > 1$ for all $a \in [a_2^*, a_1^*]$. Player $\mathcal{A}$'s (limit) payoff in a Coasean outcome is $1 - b_2^*(a)$. Since $b_2^*(a)$ is increasing in a, it follows that a_2^* is an optimal choice for $\mathcal{A}$. The role of θ decreasing when $\Phi_1^*(a_1^*) \le 1$ is detailed in the proof below.

Theorem 6. *Suppose* $\theta(a)$ *is decreasing in* a. *If* $a^* \in$ *argmax* $v_0^A(a)$ *and* $a^* \le a_1^*$, *then* $a^* = a_1^*$ *or* $a^* = a_2^*$.

Proof. Note that $v_0^A(a) = 1 - b_2^*(a) = a$ for $a \le a_2^*$. Hence a^* is never less than a_2^*. Assume first that $\Phi_1^*(a_1^*) \ge 1$. Then any demand $a \in [a_2^*, a_1^*]$ produces a Coasean outcome and $v_0^A(a) = 1 - b_2^*(a) = (1-a)r_2^B/r^A$, which is a decreasing function of a. Hence $a^* = a_2^*$.

Assume now that $\Phi_1^*(a_1^*) < 1$. Let $a_C = a_2^*$ if $\Phi_1^*(a_2^*) \le 1$ and let a_C be the unique solution of $\Phi_1^*(a_C) = 1$ otherwise. By the argument in the preceding paragraph, $v_0^A(a) < v_0^A(a_2^*)$ for $a \le a_C$ (when $a \ne a_2^*$). Now consider the region $a \in (a_C, a_1^*]$. Here $\Phi_1^*(a) < 1$ for all $a \in (a_C, a_1^*]$, and any such a demand produces a non-Coasean outcome. Let $\gamma(a) = \beta_1(1-\Phi_1^*(a)) = \beta_1 - \beta_2\theta(a)$; $\gamma(a)$ is the total probability that $\mathcal{B}$ counterdemands $b_1^*(a)$. Recall that $b_1^*(a) = 1 - a$ when $a \le a_1^*$. For any $\bar{\gamma} \in [0,1]$, define

$$\hat{v}^A(a, \bar{\gamma}) = \bar{\gamma}a + (1-\bar{\gamma})\left[\frac{1-a}{1 - r_2^B(1-a)/r^A}\right]^{r^A/r_1^B} \frac{r_2^B}{r^A}(1-a).$$

This function represents player $\mathcal{A}$'s asymptotic payoff when player $\mathcal{B}$ immediately accepts the demand a with probability $\bar{\gamma}$ and counterdemands $b_2^*(a)$ (with delay) with probability $1-\bar{\gamma}$. Note $\hat{v}^A(a,\bar{\gamma})$ is a convex function of a (one can easily check that $\hat{v}_{aa}^A(a,\bar{\gamma}) > 0$). Hence, $\hat{v}^A(a) \le \max\{\hat{v}^A(a_1^*), \hat{v}^A(a_2^*)\}$. By definition, $v_0^A(a) = \hat{v}^A(a, \gamma(a))$ for all $a \in (a_C, a_1^*]$. Since $\gamma(a) < \gamma(a_1^*)$ for all $a < a_1^*$, $v_0^A(a) < \hat{v}^A(a, \gamma(a_1^*))$ for all $a \in (a_C, a_1^*)$. Furthermore, $v_0^A(a_1^*) = \hat{v}^A(a_1^*, \gamma(a_1^*))$, and since $b_2^*(a_2^*) = 1 - a_2^*$, one can check that $v_0^A(a_2^*) = \hat{v}^A(a_2^*, \gamma(a_1^*)) = a_2^*$. It follows that $v_0^A(a) \le \max\{v^A(a_1^*), v^A(a_2^*)\}$. □

4. Existence

The proof of existence of ϵ-perturbed equilibrium may be useful in other related environments; it uses a novel mix of constructive and non-constructive elements.

For $0 \le \epsilon < 1/|\mathbb{A}|$, define the constrained unit simplex in $\mathbb{R}^{|\mathbb{A}|}$

$$\Sigma_\epsilon = \{\varphi^A \mid \varphi^A \ge \epsilon \quad \text{and} \quad \sum_{a \in \mathbb{A}} \varphi^A(a) = 1\}.$$

For each $a \in \mathbb{A}$, let

$$V(a) = \{(v_{a1}, v_{a2}) \mid 1 - a \le v_{a1} \le v_{a2} \le \bar{b}\} \quad \text{and} \quad V = \Pi_{a \in \mathbb{A}} V(a).$$

Note that V is a compact and convex subset of $\mathbb{R}^{2|\mathbb{A}|}$.

To prove existence, we construct an upper hemicontinuous correspondence $\Psi : \Sigma_\epsilon \times V \to \Sigma_\epsilon \times V$ such that each of its fixed points agrees with an equilibrium of $\Gamma(r, \beta, z)$.

For each $\varphi^A \in \Sigma_\epsilon$ and $a \in \mathbb{A}$, let the posterior $\hat{z}^A(a)$ be computed as in (1), and for each $b \in \mathbb{B}$, let $b_z(a) = (1 - \hat{z}^A(a))b + \hat{z}^A(a)(1-a)$.

Fix $(\hat{z}^A(a), a, v) \in (0,1) \times \mathbb{A} \times V$. Recall the definition of the intersection times (t^*, t_1, t_2, t_3) in Figures 2–4. These figures implicitly assumed that $a \le v_{a1} \le v_{a2}$, that

$v_{a1} < b_z(a)$, and that $v_{a2}e^{r_2^B t^*} < b_z(a)$, where $v_{a1}e^{r_1^B t^*} = b_z(a)$. We now define the densities $\varphi^B(t|\hat{z}^A(a), z^B, a, b, v_a)$ for all $(b,t) \in \mathbb{B} \times [0, \bar{T}]$ using Equations (4)–(5) and (10)–(11) in the Appendix as follows:

$$\varphi^B(t|\hat{z}^A(a), z^B, a, b, v_a) = \begin{cases} ((4), 0) & \text{if } v_{a1}e^{r_1^B t} < b_z(a),\ v_{a2}e^{r_2^B t} \geq \underline{v}_2(t|\hat{z}^A(a), v_{a1}) \\ ((10), (11)) & \text{if } v_{a1}e^{r_1^B t} < b_z(a) \text{ and} \\ & v_{a1}e^{r_1^B t} < v_{a2}e^{r_2^B t} < \underline{v}_2(t|\hat{z}^A(a), v_{a1}) \\ (0, (5)) & \text{if } v_{a2}e^{r_2^B t} \leq \min\{b_z(a), v_{a1}e^{r_1^B t}\} \\ (0, 0) & \text{if } v_{ak}e^{r_k^B t} \geq b_z(a),\ k = 1, 2. \end{cases}$$

Then, for $k = 1, 2$, define

$$\Phi_k^B(\hat{z}^A(a), z^B, a, b, v_a) = \int_0^{\bar{T}} \varphi_k^B(t|\hat{z}^A(a), z^B, a, b, v_a)dt$$
$$\bar{\Phi}_k^B(\varphi^A(a), v_a, a) = \sum_{b \in \mathbb{B}} \Phi_k^B(\hat{z}^A(a), z^B, a, b, v_a).$$

Let $\Gamma(r, \beta, \hat{z}^A(a), z^B, a)$ be the subgame where player A has demanded $a \in \mathbb{A}$ and player B believes that A is behavioral with probability $\hat{z}^A(a)$. An equilibrium for $\Gamma(r, \beta, \hat{z}^A(a), z^B, a)$ is a vector $v_a \in V(a)$ such that $\bar{\Phi}_k^B(\varphi^A(a), v_a, a) \leq 1$ and $[1 - \bar{\Phi}_k^B(\varphi^A(a), v_a, a)][v_{ak} - (1-a)] = 0$ for $k = 1, 2$.

Define $\Psi_{(\varphi^A, a)} : V(a) \to V(a)$ by

$$\Psi_{(\varphi^A, a)}(v_a) = P_{V(a)}(v_{a1}\bar{\Phi}_1^B(\varphi^A(a), v_a, a), v_{a2}\bar{\Phi}_2^B(\varphi^A(a), v_a, a)),$$

where $P_{V(a)}(w)$ is the projection of $w \in \mathbb{R}^2$ into $V(a)$. $\Psi_{(\varphi^A, a)}(v_a)$ is continuous in (φ^A, v_a).

Theorem 7. *Fix $R > 1$. Then there exists $\bar{z} > 0$ such that for all $z \in K(R, \bar{z})$ and $0 \leq \epsilon < 1/|\mathbb{A}|$, the game $\Gamma(r, \beta, z)$ has an ϵ-perturbed equilibrium.*

Proof. We first show that $v_a \in V(a)$ is an equilibrium of $\Gamma(r, \beta, \hat{z}^A(a), z^B)$ if and only if v_a is a fixed point of $\Psi_{(\varphi^A, a)}$.

A vector v_a in the interior of $V(a)$ is a fixed point of $\Psi_{(\varphi^A, a)}$ if and only if $\bar{\Phi}_1^B(\varphi^A(a), v_a, a) = \bar{\Phi}_2^B(\varphi^A(a), v_a, a) = 1$. In this case, v_a is clearly an equilibrium of $\Gamma(r, \beta, \hat{z}^A(a), z^B)$. The boundary of $V(a)$ is made up of three line segments. We now argue that $\Psi_{(\varphi^A, a)}$ has no fixed point on the upper or lower boundary of $V(a)$, and that a fixed point on the left boundary would correspond to an equilibrium where $\mathcal{B}_1$ concedes immediately to a with nonnegative probability. The left boundary is defined by $v_{a1} = 1 - a$ and $1 - a \leq v_{a2} < \bar{b}$ (we include here one of the endpoints that is at the intersection of the left boundary and the lower boundary, but not the other). A vector v_a on the left boundary is a fixed point of $\Psi_{(\varphi^A, a)}$ if and only if $\bar{\Phi}_k^B(\varphi^A(a), v_a, a) \leq 1$ for $k = 1, 2$, with equality for $k = 2$ unless $v_{a2} = 1 - a$. But then, since $v_{a1} = 1 - a$,

$[1-\bar{\Phi}_k^B(\varphi^A(a), v_a, a)][v_{ak}-(1-a)] = 0$ for $k=1,2$, and v_a is an equilibrium. Now, observe that when $v_a = (\bar{b}, \bar{b})$, $v_{ak}e^{r_k^B t} \geq b$ for all $(b,t) \in \mathbb{B}\times[0,\bar{T}]$ so $\bar{\Phi}^B(\varphi^A(a), v_a, a) = (0,0)$. In this case, $\Psi_{(\varphi^A,a)}(v_a) = P_{V(a)}(0) = (1-a, 1-a) \neq v_a$, so v_a is not a fixed point. The upper boundary (excluding end points) is defined by $v_{a2} = 1-\underline{b}$ and $a < v_{a1} < 1-\underline{b}$, and the lower boundary is defined by $a < v_{a1} = v_{a2} < 1-\underline{b}$. Assume v_a is in the upper boundary. Then for all $(b,t) \in \mathbb{B}\times[0,\bar{T}]$, $v_{a2}e^{r_2^B t} > b_z(a)$, and $\bar{\Phi}_2^B(\varphi^A(a), v_a, a) = 0$. Hence, $w = (v_{a1}\bar{\Phi}_1^B(\varphi^A(a), v_a, a), v_{a2}\bar{\Phi}_2^B(\varphi^A(a), v_a, a)) = (w_1, 0)$, where $w_1 \geq 0$, so $P_{V(a)}(w)$ lies in the lower boundary of $V(a)$ and $v_a \neq P_{V(a)}(w)$. Finally, assume v_a is in the lower boundary of $V(a)$. Then, for each $t \geq 0$, $v_{a2}e^{r_2^B t} \leq v_{a1}e^{r_1^B t}$ and $\bar{\Phi}_1^B(\varphi^A(a), v_a, a) = 0$. Hence, $w = (v_1\bar{\Phi}_1^B(\varphi^A(a), v_a, a), v_2\bar{\Phi}_2^B(\varphi^A(a), v_a, a)) = (0, w_2)$, where $w_2 \geq 0$, so $P_{V(a)}(w)$ lies in the left boundary of $V(a)$ and $v_a \neq P_{V(a)}(w)$. This establishes our claim.

For a given $(\varphi^A, v) \in \Sigma \times V$, we define $\mathcal{A}$'s payoff functions as follows

$$v_b^A(a|\varphi^A, v) = (1-b)\sum_{k=1}^{2}\beta_k \int_0^{\bar{T}} \varphi_k^B(t|\hat{z}^A(a), z^B, a, b, v_a)e^{-r^A t}dt$$

$$v^A(a|\varphi^A, v) = a\sum_{k=1}^{2}\beta_k[1-\bar{\Phi}_k^B(\varphi^A(a), v_a, a)]^+ + \sum_{b\in\mathbb{B}} v_b^A(a|\varphi^A, v),$$

where $[\xi]^+$ is ξ if $\xi > 0$ and 0 otherwise. Let

$$\Sigma_\epsilon^*(\varphi^A, v) = \{\phi \in \Sigma_\epsilon \mid \phi(a) = \epsilon \text{ for all } a \notin \operatorname{argmax} v^A(\cdot|\varphi^A, v)\}.$$

The correspondence $\Psi : \Sigma_\epsilon \times V \to \Sigma_\epsilon \times V$ is then defined by

$$\Psi(\varphi^A, v) = (\Sigma_\epsilon^*(\varphi^A, v), (\Psi_{(\varphi^A,a)}(v_a) : a \in \mathbb{A})).$$

We now argue that (φ^A, v) is a fixed point of Ψ if and only if it corresponds to an ϵ-perturbed equilibrium of $\Gamma(r,\beta,z)$. Suppose that (φ^A, v) is a fixed point of Ψ. Then v_a is an equilibrium of $\Gamma(r,\beta,\hat{z}^A(a), z^B, a)$ for each $a \in \mathbb{A}$. Also, by definition, $\varphi^A(a) > \epsilon$ implies that a is an optimal demand for $\mathcal{A}$ given that $\mathcal{B}_k$ counterdemands (b,t) with density $\varphi_k^B(t|\hat{z}^A(a), z^B, a, b, v_a)$ and accepts a immediately with probability $1-\bar{\Phi}_k^B(\varphi^A(a), v_a, a)$. Thus, (φ^A, φ^B) is an ϵ-perturbed equilibrium for $\Gamma(r,\beta,z)$. The converse is analogous.

Finally, since Ψ is upper hemicontinuous and convex-valued, by Kakutani's fixed point theorem, Ψ has a fixed point. □

5. Conclusion

The paper considers the effects of introducing behavioral types, with vanishingly low probabilities, into a bilateral bargaining model in which player 1 is uncertain which of two discount rates represents player 2's degree of impatience. Although for reasons of tractability we do not consider arbitrarily complex perturbations, behavioral types of the informed player may delay for various lengths of time before making their demands. This flexibility turns out to be crucial: Section 3 shows that for many parameter values, payoffs are non-Coasean[21] and there is substantial delay before agreement is reached. Contrast this to the benchmark model of Section 2, where reputational types make demands without delay, and solutions are always Coasean. In future work we would like to explore whether broader classes of behavioral types are advantageous for either player.

For any perturbations in the class we consider in Section 3, solutions always exist (see Section 4). With vanishingly slight perturbations (that is, when ex ante probabilities of behavioral types approach 0), all solutions[22] have approximately the same values for the uninformed player, and also for either rational type of her opponent. Thus, essential equilibrium selection is achieved.

The selected equilibria have some attractive properties. If β_1, the probability of the weaker type of player 2, is sufficiently low, player 1 does not bother trying to separate the types: she makes the Coasean demand, and there is virtually no delay. If, on the other hand, β_1 is close to 1, she is more aggressive, and a hybrid equilibrium results, with the weak type of 2 randomizing between immediately making a revealing counterdemand or waiting for quite some time before acceding to 1's demand. In this case, 1 does better than Coase would predict. Even if β_1 is moderate, 1 may opt for a subgame with a hybrid equilibrium. This is more likely if, for example, she is patient relative to her opponent, or if the two opponent types are radically different from one another (in the latter case, it is expensive to treat them both as though they were very patient). All of this is in stark contrast to the Coasean prediction favored by the literature, where the relative probabilities of the strong and weak types don't matter at all. We hope to show in future work that our non-Coasean results survive the introduction of more complex perturbation types.

Do these results suggest anything about reputational bargaining with two-sided asymmetric information? One might hope for a solution in which stronger types (having lower rates of discount, or, in a buyer/seller setting, lower valuations) wait longer before "peeling off", revealing their types by the time at which they break the silence. If a player faces an opponent having a broad range of fundamental types, it would be very costly to reveal rationality, if the Coasean outcome resulted; this would tend to make bargaining extremely slow. If instead, consistent with our results, the player whose type is revealed may still get a better-than-Coasean payoff, bargaining could be resolved in a realistic time frame. We plan to pursue these possibilities.

[21] Again, by Coasean we mean that the uninformed player does roughly as well as she would if instead she faced the stronger (more patient) opponent for sure.

[22] To be precise, this statement applies to all perturbed equilibria, as defined before Theorem 3 in Section 3.

5. Appendix

Atemporal Model.

Proof of Lemma 1

Suppose $\mathcal{A}$ reveals rationality in equilibrium at s and B has not. The proof of this part is similar to the proof of Lemma 1 of Abreu and Pearce (2007).

Step 1: There exists time $\bar{s} < \infty$ such that $\mathcal{A}$ accepts B's demand with probability 1 by $s + \bar{s}$ if B sticks with the demand b until $s + \bar{s}$. Let $\hat{s}$ satisfy

$$\frac{1}{2}(1-b) + \left[1 - \frac{1}{2}(1-b)\right] e^{-r_1^B \hat{s}} < 1 - b.$$

If $\mathcal{A}$ believes that B will reveal rationality with at most probability $(1-b)/2$ between s and $s+\hat{s}$, then $\mathcal{A}$'s expected payoff from waiting for B to reveal rationality until $s+\hat{s}$ is at most the LHS. Hence, $\mathcal{A}$ will wait until $s+\hat{s}$ only if $\mathcal{A}$ believes that B will reveal rationality with probability $\psi > (1-b)/2$ between s and $s+\hat{s}$. Conditional on player $\mathcal{A}$ not accepting player B's demand and on player B continuing to demand b until $s+\hat{s}$, a similar conclusion follows between $s+\hat{s}$ and $s+2\hat{s}$, and so on. The posterior probability $\hat{z}^B$ that player B is behavioral at s is strictly positive, and conditional on player B continuing with the demand b, the posterior probability that player B is behavioral at $s+n\hat{s}$ is $\hat{z}^B/(1-\psi)^n$. But $\mathcal{A}$ will accept B's demand if this posterior is greater or equal to 1. Hence $\mathcal{A}$ will wait until $s+n\hat{s}$ only if $n \leq \log(\hat{z}^B)/\log(1-\psi)$. Therefore, there exists $\bar{s} < \infty$ such that player $\mathcal{A}$ accepts player B's demand by $s+\bar{s}$ with probability 1, conditional on player B continuing to demand b between s and $s+\bar{s}$. Suppose that $\bar{s}$ is chosen such that the preceeding statement is false for any $s' < \bar{s}$.

Step 2: $\bar{s} = 0$. Suppose not. There exists $\epsilon > 0$ such that $be^{-r_k^B \epsilon} > 1-a$, $k = 1, 2$. It follows that conditional on sticking to the demand b until $s+\bar{s}-\epsilon$, player B will continue to stick with b with probability 1 until $s+\bar{s}$. Therefore $\mathcal{A}$ should accept B's demand with probability 1 strictly prior to $s+\bar{s}$, contradicting the definition of $\bar{s}$. All this implies that if $\mathcal{B}$ does not change his demand at *date* $(s, 0)$, then $\mathcal{A}$ will accept B's demand at date $(s, +1)$.

Note that an implication of the argument above (with the roles of A and B reversed) is that if in equilibrium $\mathcal{A}$ chooses $a \in \mathbb{A}$ at date $(0, -1)$ and $\mathcal{B}$ chooses $b \notin \mathbb{B}$ at $(0, 0)$, then $\mathcal{B}$ must concede with probability 1 to A at date $(0, +1)$. Similarly, if $\mathcal{A}$ reveals rationality in equilibrium at $(0, -1)$ by demanding $a \notin \mathbb{A}$ then, if $\mathcal{B}$ demands $\bar{b}$ at $(0, 0)$, $\mathcal{A}$ will concede with probability 1 to B at date $(0, +1)$. Hence, a lower bound on $\mathcal{B}$'s payoff is $\bar{b}$ and an *upper* bound on $\mathcal{A}$'s payoff is $(1-z^B)(1-\bar{b}) + z^B \sum \pi^B(b)(1-b)$. On the other hand, $\mathcal{A}$ could demand $\bar{a} \in \mathbb{A}$ at $(0, -1)$ and concede to any demand $b \in \mathbb{B}$ by B. If $\mathcal{B}$ demands $b \notin \mathbb{B}$, then by the preceding argument, $\mathcal{B}$ will concede to $\bar{a}$ at $(0, +1)$. It follows that a *lower* bound on $\mathcal{A}$'s equilibrium payoff is $(1-z^B)(1-\bar{b}) + z^B \sum \pi^B(b)(1-b)$. The required conclusion now follows.

Finally we show that if $\mathcal{A}$ reveals rationality with positive probability at time $s > 0$, then B does not. Let P (Q_k) be the conditional probability with which $\mathcal{A}$ ($\mathcal{B}_k$) reveals rationality at s. Suppose $P > 0$ and $Q_1 + Q_2 > 0$. We will argue that this yields a contradiction. The payoffs to exposing/revealing rationality ("E") and not revealing rationality ("$\bar{E}$") to $\mathcal{A}$ and $\mathcal{B}_k$ when $Q_k > 0$, are indicated below:

	E	$\bar{E}$
E	x, y	$1-b,\ b$
$\bar{E}$	$a,\ 1-a$	$x',\ y'$

The payoffs in the diagonal boxes follow from the first part. In the event $(\bar{E}, \bar{E})$, since the demands "on the table" can be accepted, $x' \geq 1-b$ and $y' \geq 1-a$. Hence, if E is a best response for $\mathcal{A}$ ($\mathcal{B}_k$), then $x \geq a$ ($y \geq b$). Since $a + b > 1$, this yields a contradiction. □

Proof of Lemma 3

(i) We show that under the assumed conditions (and using the notation of Lemma 1),

$$L^\ell \to \infty \quad \text{and} \quad \frac{\hat{z}^{B\ell}}{[\hat{z}^{A\ell}]^{\lambda^B/\lambda_2^A}} \to 0.$$

The conclusion then follows from Lemma 1. Note that

$$\hat{z}^{A\ell} \geq \frac{z^{A\ell}\pi^A(a)}{z^{A\ell}\pi^A(a) + 1 - z^{A\ell}},$$

where the lower bound is attained when $\varphi^{A\ell}(a) = 1$. Therefore

$$\frac{\hat{z}^{A\ell}}{\hat{z}^{B\ell}} \geq \frac{z^{A\ell}}{z^{B\ell}} \frac{\pi^A(a)}{\pi^B(b)} \frac{z^{B\ell}\pi^B(b) + (1 - z^{B\ell})[\beta_1\varphi_1^{B\ell}(b) + \beta_2\varphi_2^{B\ell}(b)]}{z^{A\ell}\pi^A(a) + 1 - z^{A\ell}},$$

and

$$\liminf_{\ell\to\infty} \frac{\hat{z}^{A\ell}}{\hat{z}^{B\ell}} \geq \frac{1}{R} \frac{\pi^A(a)}{\pi^B(b)} [\beta_1\varphi_1^{B\infty}(b) + \beta_2\varphi_2^{B\infty}(b)] \equiv c > 0.$$

Note that $\varphi_1^{B\infty}(b) + \varphi_2^{B\infty}(b) > 0$ also implies that $\hat{z}^{B\infty} = 0$. Now,

$$L^\ell = \left[\frac{\hat{z}^{A\ell}}{\hat{z}^{B\ell}}\right]^{\lambda^B} \frac{1}{[\hat{z}^{B\ell}]^{\lambda_2^A - \lambda^B}} \frac{1}{[\hat{z}^{B\ell} + \hat{\beta}_2^\ell]^{\lambda_1^A - \lambda_2^A}}$$

Since $\lambda_1^A > \lambda_2^A > \lambda^B$, $\lim\ L^\ell = \infty$. Furthermore,

$$\lim_{\ell\to\infty} \frac{\hat{z}^{B\ell}}{[\hat{z}^{A\ell}]^{\lambda^B/\lambda_2^A}} \leq \frac{1}{c} \lim_{\ell\to\infty} [\hat{z}^{A\ell}]^{1-\lambda^B/\lambda_2^A} = 0.$$

(ii) As remarked above, $\hat{z}^{B\infty} > 0$ implies that $\varphi_1^{B\infty}(b) = \varphi_2^{B\infty}(b) = 0$. Since $\varphi^{A\infty}(a) > 0$ and $\hat{z}^{B\infty} > 0$,

$$\lim_{\ell\to\infty} \hat{z}^{A\ell} = \lim_{\ell\to\infty} \frac{z^{A\ell}\pi^A(a)}{z^{A\ell}\pi^A(a) + (1 - z^{A\ell})\varphi^{A\ell}(a)} = 0$$

$$\lim_{\ell\to\infty} L^\ell = \lim_{\ell\to\infty} \frac{[\hat{z}^{A\ell}]^{\lambda^B}}{[\hat{z}^{B\ell} + \hat{\beta}_2^\ell]^{\lambda_1^A - \lambda_2^A}[\hat{z}^{B\ell}]^{\lambda_2^A}} = 0.$$

Consequently, $\lim L^\ell = 0$ and by Lemma 1, $\mu^{A\ell} \to 1$.

(iii) As in (ii), $\hat{z}^{A\ell} \to 0$. Also

$$\hat{z}^{B\ell} \geq \frac{z^{B\ell}\pi^B(b)}{z^{B\ell}\pi^B(b) + 1 - z^{B\ell}}$$

and $\hat{z}^{B\ell} + \hat{\beta}_2^\ell \leq 1$ for each ℓ. Hence

$$\begin{aligned} L^\ell &= \left[\frac{\hat{z}^{A\ell}}{\hat{z}^{B\ell}}\right]^{\lambda_2^A} \frac{[\hat{z}^{A\ell}]^{\lambda^B - \lambda_2^A}}{[\hat{z}^{B\ell} + \hat{\beta}_2^\ell]^{\lambda_1^A - \lambda_2^A}} \\ &\leq \left[\frac{z^{A\ell}}{z^{B\ell}} \frac{\pi^A(a)}{\pi^B(b)} \frac{z^{B\ell}\pi^B(b) + 1 - z^{B\ell}}{z^{A\ell}\pi^A(a) + (1 - z^{A\ell})\varphi^{A\ell}(a)}\right]^{\lambda_2^A} [\hat{z}^{A\ell}]^{\lambda^B - \lambda_2^A}, \end{aligned}$$

and

$$\lim_{\ell\to\infty} L^\ell \leq \left[R \frac{\pi^A(a)}{\pi^B(b)} \frac{1}{\varphi^{A\infty}(a)}\right]^{\lambda_2^A} \times \lim_{\ell\to\infty} [\hat{z}^{A\ell}]^{\lambda^B - \lambda_2^A} = 0$$

since $\lambda^B > \lambda_2^A$. Consequently, by Lemma 1, $\mu^{A\ell} \to 1$. □

Proof of Lemma 4

Step 1: To simplify notation, let $\hat{b}_2 = \lfloor b_2^*(a) \rfloor$. For the first part, assume by way of contradiction that there exist a sequence $\{z^\ell\}$, $z^\ell = (z^{A\ell}, z^{B\ell}) \to 0$, $a \in \mathbb{A}$ and a corresponding sequence of equilibria $\{\varphi^\ell\}$ such that $v^{A\ell}(a, z^\ell) < 1 - \hat{b}_2 - \epsilon$ for all ℓ. Without loss of generality, we can also assume that $\varphi^\ell \to \varphi^\infty$ in $\mathbb{R}^{|\mathbb{A}|} \times [\mathbb{R}^{|\mathbb{A}|\times|\mathbb{B}|}]^2$. For each ℓ and $b \in \mathbb{B}$, consider the corresponding subgame $\Gamma(r, \hat{\beta}^\ell(a, b), \hat{z}^{A\ell}(a), \hat{z}^{B\ell}(a, b), a, b)$. Clearly $\mathcal{A}$ is guaranteed a payoff of at least $1 - b$ in this subgame (since $\mathcal{A}$ can always concede to b). Therefore, $\mathcal{A}$'s payoff is at least $1 - \hat{b}_2$ whenever $b \leq \hat{b}_2$. Suppose now that $b \in \mathbb{B}$ is such that $b > \hat{b}_2$ and $\varphi_k^{B\infty}(b|a) > 0$ for $k = 1$ or $k = 2$. Since $b > \hat{b}_2$, $\lambda_2^A(a, b) > \lambda^B(a, b)$. Then, Lemma 2 *(i)* implies that $\lim \mu^{B\ell}(a, b) = 1$ and $\mathcal{A}$'s total expected payoff in the subgame after the demands (a, b) are made is bounded below by $(1 - z^{B\ell})a + z^{B\ell}(1 - b) - \epsilon/4 \geq 1 - \hat{b}_2 - \epsilon/2$ for all ℓ sufficiently large since $a \geq 1 - \hat{b}_2$. Finally, if $b \in \mathbb{B}$ is such that $b > \hat{b}_2$ and $\varphi_k^{B\infty}(b|a) = 0$ for $k = 1, 2$, then $\mathcal{A}$'s expected payoff is only bounded below by 0, but the probability of reaching the subgame with demands (a,b) is $z^{B\ell}\pi^B(b)$. Thus, $\mathcal{A}$'s total expected payoff after making the demand a is bounded below by $1-\hat{b}_2-\epsilon$ for all ℓ sufficiently large, a contradiction.

Step 2: For the second part, assume again by contradiction that there exist a sequence $\{z^\ell\}$, $z^\ell=(z^{A\ell},z^{B\ell})\to 0$, $a\in\mathbb{A}$ and a corresponding sequence of equilibria $\{\varphi^\ell\}$ such that $\varphi^{A\ell}(a)\geq\epsilon$ and for either $k=1$ or $k=2$, $v_k^B(a,z^\ell)<\hat{b}_2-\epsilon$ for all ℓ. Without loss of generality, assume that $\varphi^\ell\to\varphi^\infty$.

For each $b\in\mathbb{B}$ consider the corresponding subgame $\Gamma(r,\hat{\beta}^\ell(a,b),\hat{z}^\ell(a,b),a,b)$. Assume that $\varphi_2^{B\infty}(b|a)>0$. Then $\hat{\beta}_2^\ell(a,b)\to\beta_2\varphi_2^{B\infty}(b|a)/[\beta_1\varphi_1^{B\infty}(b|a)+\beta_2\varphi_2^{B\infty}(b|a)]>0$. Furthermore, if $b=\hat{b}_2$, then $\lambda^B(a,b)>\lambda_2^A(a,b)$. Then, Lemma 2 *(iii)* implies that $\mu^{A\infty}(a,b)=1$ and consequently $v_2^B(a,z^\ell)\to\hat{b}_2$. As $\mathcal{B}_1$ could also choose to counterdemand $\hat{b}_2$ (possibly out of equilibrium), $\lim v_1^B(a,z^\ell)\geq\hat{b}_2$. But, since $\mathcal{B}_2$'s payoff must weakly exceed $\mathcal{B}_1$'s, $v_1^B(a,z^\ell)\leq v_2^B(a,z^\ell)$ for all ℓ, and it follows that $v_1^B(a,z^\ell)\to\hat{b}_2$ as well. To complete the proof, we establish that $\varphi_2^{B\infty}(b|a)=0$ for all $b\neq\hat{b}_2$.

Step 3: Consider any $b\in\mathbb{B}$ with $b<\hat{b}_2$ and suppose that $\varphi_2^{B\infty}(b|a)>0$. For the corresponding subgames $\Gamma(r,\hat{\beta}^\ell(a,b),\hat{z}^\ell(a,b),a,b)$, without loss of generality, assume that $(\hat{z}^{B\ell}(a,b),\hat{\beta}_1^\ell(a,b),\hat{\beta}_2^\ell(a,b))\to(\hat{z}^{B\infty}(a,b),\hat{\beta}_1^\infty(a,b),\hat{\beta}_2^\infty(a,b))$. Then $\hat{\beta}_2^\infty(a,b)>0$. Furthermore, $\lambda^B(a,b)>\lambda_2^A(a,b)$. Then, by Lemma 2(iii), $\mu^{A\infty}(a,b)=1$ and $v_2^B(a,z^\ell)\to b$. As in Step 2, we may also conclude that $v_1^B(a,z^\ell)\to b$. Now consider $\hat{b}_2$. If $\hat{z}^{B\infty}(a,\hat{b}_2)>0$, then by Lemma 2(ii), $\mu^{A\infty}(a,\hat{b}_2)=1$, which contradicts $v_2^B(a,\hat{z}^\ell)\to b<\hat{b}_2$. Hence $\hat{z}^{B\infty}(a,\hat{b}_2)=0$. If $\varphi_1^{B\infty}(\hat{b}_2|a)=0$ then $\hat{\beta}_2^\infty(a,\hat{b}_2)=1$. Note that $\hat{z}^{B\infty}(a,\hat{b}_2)+\hat{\beta}_1^\infty(a,\hat{b}_2)+\hat{\beta}_2^\infty(a,\hat{b}_2)=1$. Then, By Lemma 2 *(ii)*, $\mu^{A\infty}(a,\hat{b}_2)=1$, which yields a contradiction as before. Hence $\varphi_1^{B\infty}(\hat{b}_2|a)>0$ and $\varphi_1^{B\ell}(\hat{b}_2|a)>0$ for large ℓ. Therefore

$$\mu^{A\ell}(a,\hat{b}_2)\hat{b}_2+(1-\mu^{A\ell}(a,\hat{b}_2))(1-a)\approx b,$$

which implies that $\mu^{A\infty}(a,\hat{b}_2)<1$. Let τ_1^ℓ be the time until which A concedes at rate $\lambda_1^A(a,\hat{b}_2)$ in equilibrium φ^ℓ (see Lemma 1 for a definition), and let $E^\ell(\rho)=e^{-(\rho+\lambda_1^A(a,\hat{b}_2))\tau_1^\ell}$. Then

$$\int_0^{\tau_1^\ell}e^{-\rho s}\lambda_1^A(a,\hat{b}_2)e^{-\lambda_1^A(a,\hat{b}_2)s}ds=\frac{\lambda_1^A(a,\hat{b}_2)}{\rho+\lambda_1^A(a,\hat{b}_2)}(1-E^\ell(\rho)).$$

If $\mathcal{B}_2$ mimics $\hat{b}_2$, he obtains a payoff of

$$\tilde{v}_2^{B\ell}=\mu^{A\ell}(a,\hat{b}_2)\hat{b}_2+(1-\mu^{A\ell}(a,\hat{b}_2))\left[\frac{\lambda_1^A(a,\hat{b}_2)}{r_2^B+\lambda_1^A(a,\hat{b}_2)}(1-E^\ell(r_2^B))\hat{b}_2+E^\ell(r_2^B)(1-a)\right]$$

Recall that $[\hat{\beta}_2^{B\ell}+\hat{z}^{B\ell}]e^{\lambda^B(a,\hat{b}_2)\tau_1^\ell}=1$, so $E^\ell(\rho)=[\hat{\beta}_2^{B\ell}+\hat{z}^{B\ell}]^{(\rho+\lambda_1^A(a,\hat{b}_2))/\lambda^B(a,\hat{b}_2)}$. Then, $\varphi_1^{B\infty}(\hat{b}_2|a)>0$ implies that $\lim E^\ell(r_2^B)<1$. Since

$$\frac{\lambda_1^A(a,\hat{b}_2)}{r_1^B+\lambda_1^A(a,\hat{b}_2)}\hat{b}_2=1-a$$

and $\lambda_1^A/(r_1^B+\lambda_1^A) < \lambda_1^A/(r_2^B+\lambda_1^A)$, we have that $\tilde{v}_2^{B\infty} > b$, a contradiction.

Step 4: Finally, consider any $b \in \mathbb{B}$ with $b > \hat{b}_2$, and suppose that $\varphi_2^{B\infty}(b|a) > 0$. Now $\lambda_2^A(a,b) > \lambda^B(a,b)$ and by Lemma 2(i), $\mu^{B\infty}(a,b) = 1$. Thus, $v_2^B(a,\hat{z}^\ell) \to 1-a$, and hence $v_1^B(a,\hat{z}^\ell) \to 1-a$ also. Now consider $\hat{b}_2$. As in Step 3 we conclude that $\varphi_1^{B\infty}(\hat{b}_2|a) > 0$. Now consider $\mathcal{B}_2$'s payoff from mimicking $\hat{b}_2$. If $\hat{z}^{B\infty}(a,\hat{b}_2)+\hat{\beta}_2^\infty(a,\hat{b}_2) > 0$, then by Lemma 2 *(ii)* or *(iii)*, $\mu^{A\infty}(a,\hat{b}_2) = 1$, which contradicts $v_2^B(a,\hat{z}^\ell) \to 1-a$. Hence $\hat{\beta}_1^\infty(a,\hat{b}_2) = 1$. Furthermore, $\mu^{A\infty}(a,\hat{b}_2) = 0$. Now we can simply repeat the end of Step 3 (which merely uses $\mu^{A\infty}(a,\hat{b}_2) < 1$) to conclude that $\tilde{v}_2^{B\infty} > 1-a = \lim v_2^B(a,\hat{z}^\ell)$, a contradiction. □

Temporal Model.

Proof of Lemma 5

Suppose $\mathcal{A}$ reveals rationality at $(n,-1)$. Then, by assumption, all behavioral B's will demand $\bar{b}$ at $(n,0)$. Therefore, following the demand $\bar{b}$ at $(n,0)$, $\hat{z}^B > 0$ (even if $\mathcal{B}_k$'s equilibrium strategy for $k=1,2$, entails demanding $\bar{b}$ at $(n,0)$ with probability 1), and by assumption $\hat{z}^A = 0$. Consequently, the unique continuation equilibrium entails $\mathcal{A}$ accepting $\bar{b}$ immediately. Hence, behavioral B's payoff is $\bar{b}$ and $\mathcal{B}_k$'s payoff is at least $\bar{b}$. The conclusion follows directly. □

Proof of Lemma 6

As in the proof of Lemma 4, suppose by way of contradiction that *(i)* is false. Then there exist $a \in \mathbb{A}$, $b \in \mathbb{B}$ with $b > \lfloor b_2^*(a) \rfloor$, a sequence $\{z^\ell\} \subset K(R)$ such that $z^\ell = (z^{A\ell}, z^{B\ell}) \downarrow (0,0)$, a sequence $\{t^\ell\} \subset (0,\bar{T}]$, and a corresponding sequence of equilibria $\{\varphi^\ell\}$ such that $\varphi_1^{B\ell}(b,t^\ell|a) + \varphi_2^{B\ell}(b,t^\ell|a) \geq \epsilon$ for all ℓ. Without loss of generality (taking a subsequence if necessary), we can assume that $t^\ell \to t$ and $\varphi^{A\ell}(a) \to \varphi^{A\infty}(a) \geq \epsilon$. For each $k=1,2$, if $\{\varphi_k^{B\ell}(b,t^\ell|a,z^\ell)\}$ contains a bounded subsequence, we define $\varphi_k^{B\infty}(b,t|a)$ to be the limit of that subsequence, otherwise we define $\varphi_k^{B\infty}(b,t|a) = \infty$. But now the analysis of Lemma 3 *(i)* applies exactly and we conclude that for large enough ℓ (along a subsequence), B concedes with strictly positive probability $\mu^{B\ell}$ at the start of the WOA in the subgame $\Gamma(r,\hat{\beta}^\ell,\hat{z}^{A\ell},\hat{z}^{B\ell},a,b)$ at time t^ℓ. Indeed $\mu^{B\ell} \to 1$.

On the other hand, equilibrium payoffs for $\mathcal{B}_1$ and $\mathcal{B}_2$ in the subgame are bounded below by $1-a$ and hence must be strictly greater than $1-a$ after delay $t > 0$. It follows that in the above subgame, A must concede to B with strictly positive probability at the start of the WOA. This yields a contradiction since by Lemma 2, $\mu^{A\ell}\mu^{B\ell} = 0$. □

Proof of Lemma 7

Assume that $\mathcal{A}$ has demanded $a \in \mathbb{A}$ at time $t=0$. Recall that all behavioral types of B make demands by $\bar{T}$, and that we have assumed $\bar{T}$ is not an integer. Let $\lceil \bar{T} \rceil$ denote the smallest integer greater than $\bar{T}$. We first note that if by $t \in (\bar{T}, \lceil \bar{T} \rceil)$ $\mathcal{A}$ has not revealed rationality and B has not made a counterdemand, then B will be known to be rational and consequently $\mathcal{A}$'s equilibrium continuation payoff is at least a. But $\mathcal{B}$'s continuation payoff is at least $1-a$ since B can accept $\mathcal{A}$'s standing demand. Hence, in equilibrium $\mathcal{B}$ will accept $\mathcal{A}$'s demand immediately at t.

It follows that if the Lemma is false, $\mathcal{A}$ must reveal rationality prior to $\bar{T}$. That is, there must exist an equilibrium of $\Gamma(r,\beta,z)$ and a last integer date $\bar{t} \leq \bar{T}$ at which with positive probability $\mathcal{A}$ reveals rationality before (or at the same time as) B makes a counterdemand. Let $P > 0$ be the (conditional) probability with which $\mathcal{A}$ reveals rationality at $\bar{t}$ and $Q \geq 0$ the (conditional) probability with which $\mathcal{B}$ makes a counterdemand at $\bar{t}$. Denoting $\mathcal{A}$ exposing/revealing rationality by E and B counterdemanding by C, we will argue that the payoffs to $\mathcal{A}$ and B at $\bar{t}$ are as in the table below:

	C	$\bar{C}$
E	x, y	$(1-\bar{b})^-$, $\bar{b}^+$
$\bar{E}$	a^+, $(1-a)^-$	U, $(1-a)^+$

where, for example, a^+ denotes a number greater or equal to a.

By the proof of Lemma 5, payoffs in the "box" $E\bar{C}$ are as indicated. We turn now to the box $\bar{E}\bar{C}$. Note that by the definition of $\bar{t}$, $\mathcal{A}$ does not reveal rationality (strictly) after $\bar{t}$. We now argue that in this case $\mathcal{A}$'s continuation payoff U is strictly larger than $1-\bar{b}$. Clearly $\mathcal{B}_k$'s continuation equilibrium payoff is at least $1-a$, for $k=1,2$. Let s_k satisfy $(1-a)e^{r_k^B s_k} = \lfloor b_2^*(a) \rfloor$. $\mathcal{B}_k$ will never counterdemand $b \leq \lfloor b_2^*(a) \rfloor$ at $t > s_k$, $k=1,2$. Define $W = [\beta_1 e^{-r^A s_1} + \beta_2 e^{-r^A s_2}](1-\lfloor b_2^*(a) \rfloor)$. Let $\epsilon > 0$ satisfy $(1-\bar{T}|\mathbb{B}|\epsilon)W > 1-\bar{b}$ and let $\bar{z}$ be as defined in Lemma 6, for given $R > 1$ and ϵ. By Lemma 6, all $b > \lfloor b_2^*(a) \rfloor$ are mimicked with negligible density (less than ϵ). It follows that if $\mathcal{A}$ adopts a strategy of simply accepting B's counterdemand then a lower bound on $\mathcal{A}$'s payoff is $(1-\bar{T}|\mathbb{B}|\epsilon)W$. If $\lfloor b_2^*(a) \rfloor = 1-a$, then $s_k = 0$, $k=1,2$, and $W=a$. Now, suppose $\lfloor b_2^*(a) \rfloor > 1-a$. Since

$$e^{-r^A s_k} = \left[\frac{1-a}{\lfloor b_2^*(a) \rfloor}\right]^{r^A/r_k^B} \qquad \text{and} \qquad \lfloor b_2^*(a) \rfloor \leq b_2^*(a) = 1 - \frac{r_2^B}{r^A}(1-a),$$

we conclude that

$$W \geq \left[\beta_1 \left[\frac{1-a}{\lfloor b_2^*(a) \rfloor}\right]^{r^A/r_1^B} + \beta_2 \left[\frac{1-a}{\lfloor b_2^*(a) \rfloor}\right]^{r^A/r_2^B}\right] \frac{r_2^B}{r^A}(1-a).$$

Since $(1-a)/\lfloor b_2^*(a) \rfloor > 1-a$,

$$W > \frac{r_2^B}{r^A}\left[\beta_1 (1-a)^{1+r^A/r_1^B} + \beta_2 (1-a)^{1+r^A/r_2^B}\right] > \frac{r_2^B}{r^A}(1-a)^{1+r^A/r_1^B} > 1-\bar{b},$$

where the last inequality follows from our assumption about $1-\bar{a}$ and $1-\bar{b}$.

The preceding discussion of payoffs in column $\bar{C}$ also implies that if $Q=0$ then $\mathcal{A}$ is strictly better off not revealing rationality at $\bar{t}$.

It follows that if the Lemma is false, we must have $Q > 0$. Then the posterior probability that B is behavioral (if B counterdemands at $\bar{t}$) is zero. If in addition $\mathcal{A}$ does not reveal rationality, then $\mathcal{A}$'s payoff is greater or equal than a. This justifies payoffs in the box $\bar{E}C$.

We now generate a contradiction by showing that $x + y \geq a^+ + b^+ > 1$, which is clearly infeasible. By hypothesis, $P > 0$ and $Q > 0$. However, if $y < b^+$ then $\mathcal{B}$ must set $Q = 0$, and since $U > (1 - \bar{b})$ if $x < a^+$, $\mathcal{A}$ must set $P = 0$. Hence $x + y \geq a^+ + b^+$, which generates the required contradiction. This completes the proof. □

Proof of Lemma 8

Let $\mu_1(t)$ and τ_1^S be defined by

$$v_1 e^{r_1^B t} = \mu_1(t) b + (1 - \mu_1(t))(1 - a) \quad \text{and} \quad \frac{\hat{z}^A}{1 - \mu_1(t)} e^{\lambda_1^A \tau_1^S} = 1.$$

That is, $\mu_1(t)$ is the immediate probability of concession by A that delivers the appropriate continuation value for $\mathcal{B}_1$ and τ_1^S is the length of the corresponding WOA between $\mathcal{A}$ and $\mathcal{B}_1$ (alone). Note that when $\varphi_1(b, t|a) > 0$, whether $\varphi_2(b, t|a) > 0$ or $\varphi_2(b, t|a) = 0$, $\mu_1(t)$ is always uniquely defined by the first equation above, a fact we will use later. Let

$$E = e^{-(\lambda_1^A + r_2^B)\tau_1^S} = \left[\frac{\hat{z}^A}{1 - \mu_1(t)} \right]^{(\lambda_1^A + r_2^B)/\lambda_1^A}.$$

For $t < t^*$, $(1 - \mu_1(t)) > \hat{z}^A$ so $\tau_1^S > 0$ and $E < 1$. Then

$$\begin{aligned} \underline{v}_2(t) &= \mu_1(t) b + (1 - \mu_1(t)) C \quad \text{where} \\ C &= b \int_0^{\tau_1^S} e^{-r_2^B s} \lambda_1^A e^{-\lambda_1^A s} ds + (1 - a)\, e^{-(\lambda_1^A + r_2^B)\tau_1^S} \\ &= b \frac{\lambda_1^A}{\lambda_1^A + r_2^B} (1 - E) + (1 - a) E. \end{aligned}$$

As in Step 3 of Lemma 4

$$C = (1-a) + b(1-E) \left[\frac{\lambda_1^A}{\lambda_1^A + r_2^B} - \frac{\lambda_1^A}{\lambda_1^A + r_1^B} \right] = (1-a) + (1-a)(1-E) \frac{r_1^B - r_2^B}{\lambda_1^A + r_2^B} > (1-a).$$

It follows that $\underline{v}_2(t) > v_1 e^{r_1^B t}$. However, at t^*, $1 - \mu_1(t^*) = \hat{z}^A$ so $\tau_1^S = 0$, and in this case $C = 1 - a$ and $\underline{v}_2(t^*) = v_1 e^{r_1^B t^*}$, as required. This establishes (i).

We now establish (ii). Let

$$\Omega = -r_2^B \underline{v}_2(t) + \underline{v}_2'(t) = -r_2^B [\mu_1(t)(b - C) + C] + \mu_1'(t)(b - C) + (1 - \mu_1(t)) \frac{dC}{dt}.$$

We first show that $\Omega > 0$. Since

$$\mu_1(t) = \frac{v_1 e^{r_1^B t} - (1-a)}{a+b-1} \quad \Longrightarrow \quad \mu_1'(t) = r_1^B \mu_1(t) + \lambda_1^A,$$
$$\frac{dE}{dt} = \frac{E}{1-\mu_1(t)} \left[\frac{\lambda_1^A + r_2^B}{\lambda_1^A}\right] \mu_1'(t) \quad \text{and} \quad \frac{dC}{dt} = -(1-a)\left[\frac{r_1^B - r_2^B}{\lambda_1^A + r_2^B}\right] \frac{dE}{dt},$$

we obtain that

$$\Omega = (r_1^B - r_2^B)\left[\mu_1(t)(b-C) - \frac{(1-a)E}{\lambda_1^A}(r_1^B \mu_1(t) + \lambda_1^A)\right] + \lambda_1^A(b-C) - r_2^B C.$$

Using the expression for C we deduced above, one can check that $\lambda_1^A(b-C) - r_2^B C = (1-a)(r_1^B - r_2^B)E$. Thus

$$\begin{aligned}\Omega &= (r_1^B - r_2^B)\left[\mu_1(t)(b-C) - \frac{(1-a)E}{\lambda_1^A} r_1^B \mu_1(t)\right] \\ &= (r_1^B - r_2^B)\frac{\mu_1(t)(1-a)}{\lambda_1^A}\left[r_1^B - \left(\frac{\lambda_1^A}{\lambda_1^A + r_2^B}(r_1^B - r_2^B)(1-E) + r_1^B E\right)\right] > 0\end{aligned}$$

since $[\lambda_1^A/(\lambda_1^A + r_2^B)](r_1^B - r_2^B)(1-E) + r_1^B E < (r_1^B - r_2^B)(1-E) + r_1^B E < r_1^B$. Assume that $v_2 e^{r_2^B t} \leq \underline{v}_2(t)$. Then

$$\underline{v}_2'(t) > r_2^B \underline{v}_2(t) \geq r_2^B \left[v_2 e^{r_2^B t}\right] = \frac{d}{dt}\left[v_2 e^{r_2^B t}\right]. \qquad \square$$

Proof of Lemma 9

As in the proof of Lemma 4, suppose by way of contradiction that the Lemma is false. Then there exists a sequence $\{z^\ell\} \subset K(R)$ such that $z^\ell = (z^{A\ell}, z^{B\ell}) \downarrow (0,0)$, a corresponding sequence of equilibria $\{\varphi^\ell\}$, and $a \in \mathbb{A}$ such that $\varphi_j^B(\lfloor b_k^*(a)\rfloor, s|a, z^\ell) = 0$ for all $s \geq s_0$ and $v_k e^{r_k^B(s_0+\epsilon)} < \lfloor b_k^*(a)\rfloor$. Moreover, for some $t^\ell \in [s_0, s_0+\epsilon]$ it must be the case that $\varphi_k^{B\ell}(\lfloor b_k^*(a)\rfloor, t^\ell|a, z^\ell) \leq 1/\epsilon$ (since $\int \varphi_k^{B\ell} \leq 1$). The argument for $k=2$ is virtually identical to Lemma 3 *(ii)*. We present instead the very similar argument for $k=1$. Now,

$$\hat{z}^{B\ell} \geq \frac{z^{B\ell}\pi^B(b, t^\ell)}{z^{B\ell}\pi^B(b,t^\ell) + (1-z^{B\ell})\beta_1/\epsilon}$$

since $\varphi_1^{B\ell}(\lfloor b_1^*(a)\rfloor, t^\ell|a, z^\ell) \leq 1/\epsilon$ (and $\varphi_2^{B\ell}(\lfloor b_1^*(a)\rfloor, t^\ell|a, z^\ell) = 0$ by assumption). Furthermore, since $\hat{\beta}_2^\ell = 0$, $L^\ell = [\hat{z}^{A\ell}]^{\lambda^B}/[\hat{z}^{B\ell}]^{\lambda_1^A}$, and

$$L^\ell \leq \left[\frac{z^{A\ell}\pi^A(a)}{z^{B\ell}\pi^B(b,t^\ell)} \times \frac{z^{B\ell}\pi^B(b,t^\ell) + (1-z^{B\ell})\beta_1/\epsilon}{z^{A\ell}\pi^A(a) + (1-z^{A\ell})\varphi^{A\ell}(a)}\right]^{\lambda_1^A} [\hat{z}^{A\ell}]^{\lambda^B - \lambda_1^A}.$$

Since $\pi^B(b, t^\ell) \geq \underline{\pi}$,

$$\lim_{\ell\to\infty} L^\ell \leq \left[R\,\frac{\pi^A(a)}{\underline{\pi}}\,\frac{\beta_1/\epsilon}{\varphi^{A\infty}(a)}\right]^{\lambda_1^A} \times \lim_{\ell\to\infty}[\hat{z}^{A\ell}]^{\lambda^B-\lambda_1^A}.$$

Now, $\lambda^B > \lambda_1^A$ and

$$\hat{z}^{A\ell} \leq \frac{z^{A\ell}\pi^A(a)}{z^{A\ell}\pi^A(a) + (1 - z^{A\ell})\epsilon}$$

since $\varphi^{A\ell}(a) \geq \epsilon$. It follows that $\lim\,[\hat{z}^{A\ell}]^{\lambda^B-\lambda_1^A} = 0$. Consequently, by Lemma 2, $\mu^{A\ell} \to 1$. But $\mu^{A\ell} \to 1$ implies $v_1^\ell e^{r^{B_1}(s_0+\epsilon)} \geq v_1^\ell e^{r_1^B t^\ell} \to \lfloor b_1^*(a)\rfloor$, a contradiction □

Proof of Lemma 10

(i) This builds on the derivation of the sneaking in function $\underline{v}_2$ of Lemma 8. Consider $b < \lfloor b_2^*(a)\rfloor$. We first argue that $\varphi_2^B(b, t|a, z) = 0$ for all $t \in [0, \bar{T}]$. Suppose not. Then there exist sequences as before with $\varphi_2^{B\ell}(b, t^\ell|a, z^\ell) > 0$. Since $\mathcal{B}_2$ can concede to a at time zero, it must be that $b > 1 - a$. Since the payoff from mimicking b at t^ℓ is strictly less than $b_z^\ell = \hat{z}^{A\ell}(1 - a) + (1 - \hat{z}^{A\ell})b$, it must be that $v_2^\ell e^{r_2^B t^\ell} < b_z^\ell$. Let $t^{*\ell}$ satisfy $v_1^\ell e^{r_1^B t^{*\ell}} = b_z^\ell$. We wish to first argue that $v_2^\ell e^{r_2^B s} = b_z^\ell$ for some $s > t^{*\ell}$. If $t^\ell \geq t^{*\ell}$ this is obvious (since $v_2^\ell e^{r_2^B t^\ell} < b_z^\ell$). Now suppose that $t^\ell < t^{*\ell}$.

If $\varphi_1^{B\ell}(b, t^\ell|a, z^\ell) = 0$, then $v_1^\ell e^{r_1^B t^\ell} \geq v_2^\ell e^{r_2^B t^\ell}$, since $\mathcal{B}_1$ always has the option of first counterdemanding b at t^ℓ. On the other hand, if $\varphi_1^{B\ell}(b, t^\ell|a, z^\ell) > 0$, then $v_2^\ell e^{r_2^B t^\ell} \leq \underline{v}_2(t^\ell)$. In either case, $v_2^\ell e^{r_2^B t^\ell} \leq \underline{v}_2(t^\ell)$. Since $v_2^\ell e^{r_2^B s}$ is flatter than $\underline{v}_2(s)$ and $v_1^\ell e^{r_1^B t^{*\ell}} = \underline{v}_2(t^{*\ell})$ (see Lemma 8), $v_2^\ell e^{r_2^B s_0} = b$ for some $s_0 > t^{*\ell}$. For all $t' \geq t^{*\ell}$, $v_1 e^{r_1^B t'} \geq b_z$, and consequently $\varphi_1^{B\ell}(\lfloor b_2^*(a)\rfloor, t'|a, z^\ell) = 0$. But this contradicts Lemma 9, as depicted in Figure 9 below, where ϵ' is defined by $be^{r_2^B\epsilon'} = (\lfloor b_2^*(a)\rfloor + b)/2$ (and here and below ϵ' plays the role of the ϵ in the statement of Lemma 9).

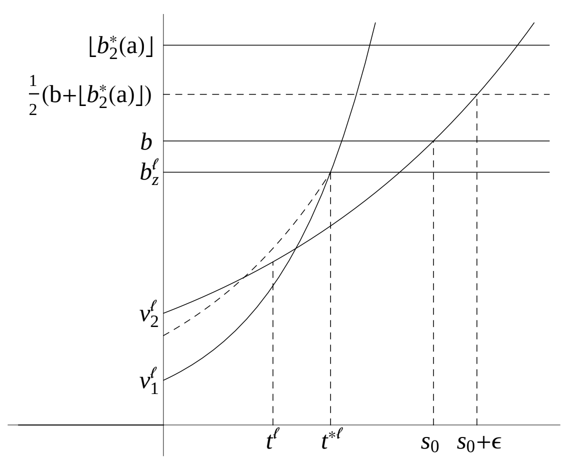

Figure 9

Now consider $b < \lfloor b_1^*(a) \rfloor$. We wish to argue that $\varphi_1^B(b,t|a,z) = 0$. Suppose not. Then, there exist sequences (analogous to the earlier sequences) such that $\varphi_1^{B\ell}(b,t^\ell|a,z^\ell) > 0$ for all ℓ. But this implies $v_1^\ell < b$. Let ϵ' be defined by $be^{r_1^B \epsilon'} = (\lfloor b_1^*(a) \rfloor + b)/2$. Since $\lfloor b_1^*(a) \rfloor < \lfloor b_2^*(a) \rfloor$, for ℓ large enough $\varphi_2^{B\ell}(b,s|a,z^\ell) = 0$ for all $s \in [0,\epsilon']$ by our earlier conclusion. Again we have contradicted Lemma 9 for $s_0 = 0$ and ϵ' as defined above.

(ii) We know from *(i)* that $\varphi_2^B(b,t|a,z) = 0$. If the result is not true there exist sequences as in *(i)* such that $\varphi_1^{B\infty}(b,t|a) > 0$ and $\varphi_2^{B\ell}(b,t^\ell|a,z^\ell) = 0 = \hat{\beta}_2^\ell$ along the sequence. Observe that $\lambda_1^A > \lambda^B$. Now an argument almost identical to the proof of Lemma 3 *(i)* [replace $\hat{z}_2^\ell + \hat{\beta}_2^\ell$ by $\hat{z}_2^\ell$ and subsequently λ_2^A by λ_1^A] implies $\mu^{B\ell} \to 1$. In particular, $\mu^{B\ell} > 0$. This contradicts the requirement that $\mu^{A\ell} > 0$.

(iii) Let $\bar{z}_0$ be chosen to satisfy *(i)*-*(ii)* for the given ϵ. By *(i)*, for $z \le \bar{z}_0$ and for all $t \ge 0$, $\varphi_2^B(\lfloor b_1^*(a) \rfloor, t|a,z) = 0$. By Lemma 9 there exists $\bar{z} \le \bar{z}_0$ such that $v_1 e^{r_1^B \epsilon} \ge \lfloor b_1^*(a) \rfloor$. Then *(iii)* follows for all $z \le \bar{z}$, and the proof is complete. □

Proofs of Observations 3 and 4

Proof of Observation 3: Consider t such that in equilibrium $\varphi_k^B(t) > 0$ $k = 1,2$. Then, in the WOA following B's counterdemand at t, A will concede at t with probability $\mu_1(t)$ as defined in the proof of Lemma 8, and thereafter at a rate of concession λ_1^A in $(t, t+\tau_1]$ and at rate λ_2^A in $(t+\tau_1, t+\tau_2)$. Player B concedes at rate λ^B. The times τ_1 and τ_2 satisfy:

$$(\hat{z}^B + \hat{\beta}_2)e^{\lambda^B \tau_1} = 1, \quad \hat{z}^B e^{\lambda^B \tau_2} = 1 \quad \text{and} \quad \frac{\hat{z}^A}{1-\mu_1(t)} e^{\lambda_1^A \tau_1} e^{\lambda_2^A(\tau_2 - \tau_1)} = 1.$$

In the proof of Lemma 8, we also let τ_1^S be such that

$$\frac{\hat{z}^A}{1-\mu_1(t)} e^{\lambda_1^A \tau_1^S} = 1.$$

Clearly, $\tau_2 > \tau_1$, so $\tau_1^S > \tau_1$. Note that $\mathcal{B}_2$'s payoff (in equilibrium) is given by $v_2(t) = \mu_1 \lfloor b_2^*(a) \rfloor + (1-\mu_1)C$, where we simply replace τ_1^S by τ_1 in E. It follows that $v_1 e^{r_1^B t} < v_2 e^{r_2^B t} < \underline{v}_2(t)$. □

Proof of Observation 4: Since $v_1 < v_2$, $v_2 > \lfloor b_1^*(a) \rfloor$. Hence, by Observation 1, $\Phi_2^B = 1$ and $\varphi_2^B(t) > 0$ for some $t > 0$; at such a t, $v_2 e^{r_2^B t} < b_z$. If $t \ge t^*$ then $v_2 e^{r_2^B t} < b_z \le v_1 e^{r_1^B t}$. Since $v_2 > v_1$, all this implies that $t_2 < t^*$. If $t < t^*$ and $v_1 e^{r_1^B t} > v_2 e^{r_2^B t}$, we may similarly conclude that $t_2 < t^*$. If $t < t^*$ and $v_1 e^{r_1^B t} \le v_2 e^{r_2^B t}$, then by Observation 3 $v_1 e^{r_1^B t} \le v_2 e^{r_2^B t} \le \underline{v}_2(t)$. Since $\mathcal{B}_2$'s indifference curve cannot intersect $\underline{v}_2(\cdot)$ from below (Lemma 8),

$$v_2 e^{r_2^B s} - v_1 e^{r_1^B s} \le \underline{v}_2(s) - v_1 e^{r_1^B s} \qquad \text{for all } s \in [t, t^*],$$

and since $\underline{v}_2(t^*) = v_1 e^{r_1^B t^*}$, we must have that $v_1 e^{r_1^B t_2} = v_2 e^{r_2^B t_2}$ for some $t_2 < t^*$. Finally, $t_2 < t^*$ implies that $t^* < t_3$ since $\mathcal{B}_2$'s indifference curve is flatter than $\mathcal{B}_1$'s. □

Proof of Lemma 11

If $\mathcal{B}_1$ counterdemands $\lfloor b_2^*(a) \rfloor$ at any $t > t^*$, then his payoff (discounted to 0) is less than $b_z e^{-r_1 t} < v_1$. Hence $\varphi_1(t) = 0$ for all $t > t^*$. For the same reason, $\varphi_2(t) = 0$ for all $t > t_3$ Let $t < t_2$, so $v_1 e^{r_1 t} < v_2 e^{r_2 t} \leq b_z$ It follows that either $\varphi_1(t) > 0$ or $\varphi_2(t) > 0$. Assume that $\varphi_1(t) = 0$ and $\varphi_2(t) > 0$. After making the counterdemand $\lfloor b_2^*(a) \rfloor$ at time t, $\mathcal{B}_2$ can get his expected value $v_2 e^{r_2 t}$ in the ensuing WOA by waiting to see if $\mathcal{A}$ concedes to $\lfloor b_2^*(a) \rfloor$ right away, and conceding to a immediately if $\mathcal{A}$ does not. But $\mathcal{B}_1$ can obtain the same expected payoff by mimicking $\mathcal{B}_2$: counterdemanding $\lfloor b_2^*(a) \rfloor$ at time t and following the same strategy after that. This is a contradiction. Hence, $\varphi_1(t) > 0$ for all $t \in (0, t_2)$. Finally suppose that $t \in (t_2, t_3)$. Again, either $\varphi_1(t) > 0$ or $\varphi_2(t) > 0$. Assume now that $\varphi_1(t) > 0$. $\mathcal{B}_1$ can get his expected value $v_1 e^{r_1 t}$ in the ensuing WOA by waiting until some time $t + \tau > t$ to concede if $\mathcal{A}$ does not concede first. But, if $\mathcal{B}_2$ mimicks $\mathcal{B}_1$, he expects a strictly higher payoff as $r_1 > r_2$, which is a contradiction since by assumption $v_1 e^{r_1 t} > v_2 e^{r_2 t}$. Hence $\varphi_1(t) = 0$ and $\varphi_2(t) > 0$ for all $t \in (t_2, t_3)$. □

Before we prove Lemmas 12–15, we explicitly construct the densities $\varphi_k^B(b, t|a)$, $k = 1, 2$, and corresponding linear approximations that we will later use to compute various limits as reputations converge to 0.

Densities.

Assume that A has demanded $a \in \mathbb{A}$, and that in equilibrium $\varphi^A(a) > 0$, so the posterior $\hat{z}^A$ that A is behavioral is given by (1). For any $b \in \mathbb{B}$ with $a + b > 1$, we now derive the densities with which $\mathcal{B}_k$'s choose b at various times.

By Lemma 2 of the atemporal types model, the WOA $\Gamma(r, \hat{\beta}, \hat{z}^A, \hat{z}^B, a, b)$ has a unique equilibrium and hence a unique equilibrium value v_k^B for each $\mathcal{B}_k$ with $\hat{\beta}_k > 0$, $k = 1, 2$. Hereafter, fix $(\hat{z}^A, z^B, a, b)$ once and for all, where $a + b > 1$ (possibly, $b = \lfloor b_2^*(a) \rfloor = \lfloor b_2^*(a) \rfloor$). Consider the WOA that arises after the counterdemand b at time t, when $\mathcal{A}$ believes that $\mathcal{B}_k$ counterdemands (b, t) with probability $\varphi_k^B(t)$, $k = 1, 2$.[23] By equations (2)–(3), $(\hat{z}^B(t), \hat{\beta}_1(t), \hat{\beta}_2(t))$ are functions of $(\varphi_1^B(t), \varphi_2^B(t))$,[24] and thus $(\varphi_1^B(t), \varphi_2^B(t))$ leads to a unique equilibrium value $v_k^B(t)$ for each $\mathcal{B}_k$ with $\varphi_k^B(t) > 0$, $k = 1, 2$, in the corresponding WOA. In any equilibrium, it must be the case that $v_k^B(t) = v_k e^{r_k t}$ for some fixed v_k, $k = 1, 2$. Given (v_1, v_2), the equilibrium value functions can be inverted to construct the functions $(\varphi_1^B(t), \varphi_2^B(t))$ so that for each t, the corresponding WOA delivers the equilibrium value $v_k e^{r_k t}$ for each $k = 1, 2$ with $\varphi_k^B(t) > 0$.

Fix (v_1, v_2) so that $1 - a \leq v_1 \leq v_2 < b_z$.[25] We now solve for $(\varphi_1^B(t), \varphi_2^B(t))$ in each

[23] To simplify notation here, since we fix b and focus only on the time dimension, we write $\varphi_k^B(t)$ instead of $\varphi_k^B(t|\hat{z}^A, z^B, a, b, v_1, v_2)$.

[24] They are also functions of (a, b), but since these variables have been fixed, we omit them here.

[25] In equilibrium, $v_1 \geq 1 - a$ always. If $v_1 \geq b_z$, then $\varphi_k^B(t) \equiv 0$ for $k = 1, 2$. If $1 - a \leq v_1 < b_z \leq v_2$, then $\varphi_2^B(t) \equiv 0$ and only equation (4) below is relevant.

one of the separating and pooling intervals. Consider the WOA after the counterdemand (b,t) with $t>0$. Let $\mu_k(t)$ be such that

$$v_k e^{r_k t} = \mu_k(t) b + (1-\mu_k(t))(1-a), \quad k=1,2.$$

When $\varphi_1^B(t)>0$, $\mu_1(t)$ is the required probability of immediate concession by A to deliver $\mathcal{B}_1$ his corresponding expected payoff. Similarly, when $\varphi_1^B(t)=0$ and $\varphi_2^B(t)>0$, $\mu_2(t)$ is the required probability of immediate concession by A to deliver $\mathcal{B}_2$ his corresponding expected payoff. When $\varphi_k^B(t)>0$ it must be that $1-a<v_k e^{r_k t}<b$, hence

$$\mu_k(t) = \frac{v_k e^{r_k t} - (1-a)}{a+b-1} \in (0,1),\ k=1,2.$$

It is also useful to define the function $\bar\mu_k(t) = 1-\mu_k(t)$, $k=1,2$. Since $v_k e^{r_k^B t} \in (1-a,b)$, $\bar\mu_k(t)$ is the distance from $v_k e^{r_k^B t}$ to b relative to the total distance from $1-a$ to b.

Recall the various pooling and separating intervals from Lemma 11.

Separating Interval for $\mathcal{B}_1$: Let t be such that $\varphi_1^B(t)>0$ and $\varphi_2^B(t)=0$. Then, By Lemma 2, the length of the WOA τ_1 must be such that

$$1 = \frac{\hat z^A}{\bar\mu_1(t)} e^{\lambda_1^A \tau_1} = \frac{z^B \pi^B(b,t)}{z^B\pi^B(b,t) + (1-z^B)\beta_1\varphi_1^B(t)} e^{\lambda^B \tau_1}.$$

Let $Z(z^B,b,t) = z^B\pi^B(b,t)/(1-z^B)$. It follows that

$$\varphi_1^B(t) = \frac{Z(z^B,b,t)}{\beta_1}\left[\left[\frac{\bar\mu_1(t)}{\hat z^A}\right]^{\lambda^B/\lambda_1^A} - 1\right]. \tag{4}$$

Separating Interval for $\mathcal{B}_2$: Let t be such that $\varphi_1^B(t)=0$ and $\varphi_2^B(t)>0$. By a similar argument, we now obtain that

$$\varphi_2^B(t) = \frac{Z(z^B,b,t)}{\beta_2}\left[\left[\frac{\bar\mu_2(t)}{\hat z^A}\right]^{\lambda^B/\lambda_2^A} - 1\right]. \tag{5}$$

Pooling Interval: Let t be such that $\varphi_1^B(t)>0$ and $\varphi_2^B(t)>0$. Here again, $\mu_1(t)$ must be the probability that A concedes immediately. By Lemma 2, there exist $0<\tau_1<\tau_2$ such that

$$[\hat z^B(t)+\hat\beta_2(t)]e^{\lambda^B\tau_1}=1, \quad \hat z^B(t)e^{\lambda^B\tau_2}=1, \quad \text{and} \quad \frac{\hat z^A}{\bar\mu_1(t)} e^{\lambda_1^A\tau_1+\lambda_2^A(\tau_2-\tau_1)}=1.$$

Therefore

$$\bar\mu_1(t) = \hat z^A\left[\frac{1}{\hat z^B(t)+\hat\beta_2(t)}\right]^{\lambda_1^A/\lambda^B}\left[\frac{\hat z^B(t)+\hat\beta_2(t)}{\hat z^B(t)}\right]^{\lambda_2^A/\lambda^B}. \tag{6}$$

The WOA should also deliver $\mathcal{B}_2$ his expected value $v_2 e^{r_2^B t}$. An optimal strategy for $\mathcal{B}_2$ is to concede at $t+\tau_1$ if A has not conceded yet. Therefore, $\mathcal{B}_2$'s expected value in the WOA is

$$\begin{aligned} v_2 e^{r_2^B t} &= \mu_2(t) b + (1-\mu_2(t))(1-a), \\ &= \left[\mu_1(t) + (1-\mu_1(t)) \frac{\lambda_1^A}{\lambda_1^A + r_2^B}(1-E)\right] b + (1-\mu_1(t)) E(1-a) \end{aligned}$$

where

$$E = e^{-(\lambda_1^A + r_2^B)\tau_1} \qquad \text{so} \qquad \int_0^{\tau_1} e^{-r_1^B \tau} \lambda_1^A e^{-\lambda_1^A \tau} d\tau = \frac{\lambda_1^A}{\lambda_1^A + r_2^B}(1-E).$$

Subtracting b from both sides in the previous equation, we obtain

$$\bar{\mu}_2(t)(a+b-1) = \bar{\mu}_1(t) \left[\left[1 - \frac{\lambda_1^A}{\lambda_1^A + r_2^B}(1-E)\right] b - E(1-a)\right]. \tag{7}$$

Let $D = z^B \pi^B(b,t) + (1-z^B)[\beta_1^B \varphi_1(t) + \beta_2 \varphi_2^B(t)]$ and $U = z^B \pi^B(b,t) + (1-z^B)\beta_2 \varphi_2^B(t)$ so that $\hat{z}^B(t) + \hat{\beta}_2(t) = U/D$. Substituting the expressions for the corresponding posteriors in (6) we obtain

$$\bar{\mu}_1(t) = \frac{\hat{z}^A D^{\lambda_1^A/\lambda^B}}{U^{(\lambda_1^A - \lambda_2^A)/\lambda^B} [z^B \pi^B(b,t)]^{\lambda_2^A/\lambda^B}}. \tag{8}$$

Solving for E from (7), we conclude that $\mathcal{B}_2$'s expected value is attained when

$$\begin{aligned} E &= \frac{(\lambda_1^A + r_2^B)(a+b-1)\bar{\mu}_2(t) - r_2^B b \bar{\mu}_1(t)}{[\lambda_1^A b - (\lambda_1^A + r_2^B)(1-a)]\bar{\mu}_1(t)} \\ &= e^{-(\lambda_1^A + r_2^B)\tau_1} = [\hat{z}^B(t) + \hat{\beta}_2(t)]^{[\lambda_1^A + r_2^B]/\lambda^B} = \left[\frac{U}{D}\right]^{\frac{\lambda_1^A + r_2^B}{\lambda^B}} \end{aligned} \tag{9}$$

Let $c = (\lambda_1^A + r_2^B)(a+b-1)$ and $d = r_2^B b$. Note that $c > d$ since $\lambda_1^A > \lambda_2^A$. Also let $\gamma_k = \lambda_k^A/\lambda^B$, $k = 1,2$, and $\rho = r_2^B/\lambda^B$. Then (8) and (9) imply

$$\left[\frac{D^{\gamma_1}}{U^{\gamma_1-\gamma_2}}\right]^{\frac{\gamma_1+\rho}{\gamma_1-\gamma_2}} \left[\frac{U}{D}\right]^{\gamma_1+\rho} = D^{\frac{\gamma_2(\gamma_1+\rho)}{\gamma_1-\gamma_2}} = \left[\frac{\bar{\mu}_1[z^B\pi^B(b,t)]^{\gamma_2}}{\hat{z}^A}\right]^{\frac{\gamma_1+\rho}{\gamma_1-\gamma_2}} \left[\frac{c\bar{\mu}_2 - d\bar{\mu}_1}{(c-d)\bar{\mu}_1}\right]$$

$$\left[\frac{D^{\gamma_1}}{U^{\gamma_1-\gamma_2}}\right]^{\frac{\gamma_1+\rho}{\gamma_1}} \left[\frac{U}{D}\right]^{\gamma_1+\rho} = U^{(\gamma_1+\rho)(\gamma_1+\gamma_2)/\gamma_1} = \left[\frac{\bar{\mu}_1[z^B\pi^B(b,t)]^{\gamma_2}}{\hat{z}^A}\right]^{\frac{\gamma_1+\rho}{\gamma_1}} \left[\frac{c\bar{\mu}_2 - d\bar{\mu}_1}{(c-d)\bar{\mu}_1}\right]$$

which can be solved for D and U to get

$$D = \left[\frac{\bar{\mu}_1[z^B\pi^B(b,t)]^{\gamma_2}}{\hat{z}^A}\right]^{1/\gamma_2} \left[\frac{c\bar{\mu}_2 - d\bar{\mu}_1}{(c-d)\bar{\mu}_1}\right]^{\frac{\gamma_1-\gamma_2}{\gamma_2(\gamma_1+\rho)}}$$

$$U = \left[\frac{\bar{\mu}_1[z^B\pi^B(b,t)]^{\gamma_2}}{\hat{z}^A}\right]^{1/\gamma_2} \left[\frac{c\bar{\mu}_2 - d\bar{\mu}_1}{(c-d)\bar{\mu}_1}\right]^{\frac{\gamma_1}{\gamma_2(\gamma_1+\rho)}}$$

Finally $(1-z^B)\beta_1\varphi_1^B(b,t|a;v) = D-U$ and $(1-z^B)\beta_2\varphi_2^B(b,t|a;v) = U - z^B\pi^B(b,t)$. It follows that

$$\varphi_1^B(t) = \frac{Z(z^B,b,t)}{\beta_1}\,[g(t)-h(t)] \quad \text{and} \tag{10}$$

$$\varphi_2^B(t) = \frac{Z(z^B,b,t)}{\beta_2}\,[h(t)-1], \quad \text{where} \tag{11}$$

$$N = \frac{\lambda^B}{\lambda_2^A}, \quad m_1 = N\,\frac{\lambda_1^A-\lambda_2^A}{\lambda_1^A+r_2^B}, \quad m_2 = N\,\frac{\lambda_1^A}{\lambda_1^A+r_2^B},$$

$$g(t) = \left[\frac{\bar\mu_1(t)}{\hat z^A}\right]^N \left[\frac{c\bar\mu_2(t)-d\bar\mu_1(t)}{(c-d)\bar\mu_1(t)}\right]^{m_1} \quad \text{and} \tag{12}$$

$$h(t) = \left[\frac{\bar\mu_1(t)}{\hat z^A}\right]^N \left[\frac{c\bar\mu_2(t)-d\bar\mu_1(t)}{(c-d)\bar\mu_1(t)}\right]^{m_2} \tag{13}$$

It is clear that $\varphi_k^B(t)$, $k=1,2$, (formulas (4)–(5) and (10)–(11)) depend on the equilibrium values (v_1,v_2). They also depend on the demands (a,b) (assumed fixed early on) and the reputations $(\hat z^A, z^B)$. We may find it convenient to make this dependence explicit sometimes and write instead $\varphi_k(t|\hat z^A,z^B,a,b,v_1,v_2)$ (or $\varphi_k(t|v_1,v_2)$ if we only want to highlight the dependence on (v_1,v_2)). The function $\varphi_k(t|\hat z^A,z^B,a,b,v_1,v_2)$ is continuous in $(t,\hat z^A,z^B,a,b,v_1,v_2)$.

By definition, $\bar\mu_2(t_3) = \hat z^A$, so $\varphi_2^B(t_3) = 0$. Since $\bar\mu_2(t)$ is a decreasing function of t, (5) implies that $\varphi_2^B(t)/\pi^B(b,t)$ is decreasing in $t\in(t_2,t_3)$. Tedious but simple computations also show that $h'(t)>0$, so by (11), $\varphi_2^B(t)/\pi^B(b,t)$ is strictly increasing in $t\in(t_1,t_2)$. Hence, $\varphi_2^B(t)/\pi^B(b,t)$ is single-peaked at $t=t_2$. Also, $h(t_1)=1$, so $\varphi_2^B(t_1)=0$, as required.

For completeness, let us also include here a simple expression for the sneaking-in value function that we developed in Lemma 8 above:

$$\underline{v}_2(t) = (1-\lfloor b_2^*(a)\rfloor) - \frac{\bar\mu_1(t)}{\lambda_1^A+r_2^B}\left[d+(c-d)\left[\frac{\hat z^A}{\bar\mu_1(t)}\right]^{1+r_2^B/\lambda_1^A}\right]. \tag{14}$$

Linearization.

In this section we develop approximations for $\varphi_k^B(t)$, $k=1,2$, and their integrals. As in the previous subsection, we maintain fixed $a\in\mathbb{A}$ and $b\in\mathbb{B}$, where $a+b>1$ and $\varphi^A(a)>0$. As always, let $b_z = (1-\hat z^A)b + \hat z^A(1-a)$. The values of $\varphi_1^B(t)$ and $\varphi_2^B(t)$ depend on (v_1,v_2) nonlinearly, through $\bar\mu_k(t)$, $k=1,2$. Fix $v_1\in[1-a,b_z)$ and let t^* be such that $v_1e^{r_1^Bt^*} = b_z$. We now develop linear approximations for $\bar\mu_1(t)$ and $\bar\mu_2(t)$ near t^*. When the times t_j, $j=1,2,3$, are close to t^*, the linear approximations produce good approximations for $\varphi_1^B(t)$ and $\varphi_2^B(t)$.

Here we restrict to the case where $v_2>v_1$ and $v_2e^{r_2^Bt^*} = b_z-\Delta$ for some $\Delta>0$. In this case, (t_1,t_2) and (t_2,t_3) are nonempty intervals. Note that adjusting v_2 is equivalent to adjusting Δ. Hereafter we assume that Δ is small,. It follows that $|t_1 - t^*|$ and $|t_3 - t^*|$ are order $O(\Delta)$. Therefore, by Taylor series expansion around t^*, $e^{r_k^B t} = e^{r_k^B t^*}[1 + r_k^B(t - t^*)] + O(|t-t^*|^2)$ for all $t \in (t_1, t_3)$. Let $s = t - t^*$. Hence

$$v_1 e^{r_1^B t} = b_z(1 + r_1^B s) + O(s^2) \quad \Longrightarrow \quad \bar{\mu}_1(t) = -\frac{r_1^B b_z s}{a+b-1} + O(s^2)$$
$$v_2 e^{r_2^B t} = (b_z - \Delta)(1 + r_2^B s) + O(s^2) \quad \Longrightarrow \quad \bar{\mu}_2(t) = \frac{\Delta - r_2^B b_z s}{a+b-1} + O(s^2).$$

Let $s_i = t_i - t^*, i = 1, 2, 3$. The equation $\bar{\mu}_1(t_2) = \bar{\mu}_2(t_2)$ implies that $-r_1^B b_z s_2 = \Delta - r_2^B b_z s_2 + O(s_2^2)$, or

$$s_2 = \tilde{s}_2 + O(\Delta^2) \quad \text{where} \quad \tilde{s}_2 = \frac{-\Delta}{b_z(r_1^B - r_2^B)}. \tag{15}$$

Similarly, $v_2 e^{r_2^B t_3} = b_z$ leads to $(b - \Delta)(1 + r_2^B s_3) + O(s_3^2) = b_z$, or

$$s_3 = \tilde{s}_3 + O(\Delta^2) \quad \text{where} \quad \tilde{s}_3 = \frac{\Delta}{r_2^B b_z}. \tag{16}$$

If $v_2 \geq \underline{v}_2(0)$ the equation $\underline{v}_2(t_1) = v_2 e^{r_2^B t_1}$ leads to the approximation

$$\frac{r_1^B b_z s_1}{(\lambda_1^A + r_2^B)(a+b-1)}\left[d + (c-d)W\right] = -\Delta + b_z r_2^B s_1 + O(s_1^2)$$
$$\text{where} \qquad W = \left[\frac{\hat{z}^A(a+b-1)}{-r_1^B b_z s_1}\right]^{1 + r_2^B/\lambda_1^A}.$$

This is a nonlinear equation in s_1. If we view W as an exogenous parameter, the solution s_1 of this equation increases with W. When we make $W = 0$, we obtain the approximation

$$\tilde{s}_1 = \frac{(\lambda_1^A + r_2^B)}{r_2^B}\tilde{s}_2, \tag{17}$$

and hence $\tilde{s}_1 + O(\Delta^2) \leq s_1 \leq s_2$. Though this does not establish a tight estimate for s_1, it does establish that $|s_1| = |t_1 - t^*| = O(\Delta)$, as claimed earlier. If $v_2 > \underline{v}_2(0)$, then $t_1 = 0$ (by definition) and consequently we define $\tilde{s}_1 = \max\{(\lambda_1^A + r_2^B)\tilde{s}_2/r_2^B, -t^*\}$ in general. Obviously, when $-t^* > (\lambda_1^A + r_2^B)\tilde{s}_2/r_2^B$, we still have that $\tilde{s}_1 = O(\Delta)$. We discuss $\tilde{s}_1$ later in more detail, after we obtain an approximation for $\varphi_2^B(t)$.

Let

$$X = \frac{\lambda_1^A + r_2^B}{(r_1^B - r_2^B)\lambda_1^A b_z}, \qquad Y = \frac{r_2^B}{\lambda_1^A}, \quad \text{and} \quad n_j = N - m_j,\ j = 1, 2.$$

If $\xi(x,y) = x^n y^m$, then by Taylor approximation,

$$\xi(x+\epsilon_1, y+\epsilon_2) = \xi(x,y)\left[1 + \frac{n\epsilon_1}{x} + \frac{m\epsilon_2}{y}\right].$$

We use this approximation when $x = -s$, $y = X\Delta + Ys$, $\epsilon_1 = O(s^2) = \epsilon_2$, and $s \in [s_1, s_2]$. Since $s_1 < s_2 < 0$ and $s_2 = O(\Delta)$, $\xi(-s+O(s^2), X\Delta+Ys+O(s^2)) = \xi(-s, X\Delta+Ys)(1+O(\Delta)$. Therefore, for each function $\gamma \in \{g, h, \varphi_1^B, \varphi_2^B\}$

$$\gamma(t^* + s) = \tilde{\gamma}(s)(1 + O(\Delta)) \quad \text{for all } s \in [\tilde{s}_1, \tilde{s}_3],$$

where the corresponding functions $\tilde{\gamma}$ are defined as follows

$$\tilde{g}(s) = \left[\frac{r_1^B b_z}{\hat{z}^A(a+b-1)}\right]^N [-s]^{n_1}[X\Delta + Ys]^{m_1} \tag{18}$$

$$\tilde{h}(s) = \left[\frac{r_1^B b_z}{\hat{z}^A(a+b-1)}\right]^N [-s]^{n_2}[X\Delta + Ys]^{m_2} \tag{19}$$

$$\tilde{\varphi}_1^B(s) = \frac{Z(z^B, b, t^*+s)}{\beta_1}\left[\tilde{g}(s) - \tilde{h}(s)\right] \qquad s \in [\tilde{s}_1, \tilde{s}_2] \tag{20}$$

$$\tilde{\varphi}_2^B(s) = \begin{cases} \dfrac{Z(z^B, b, t^*+s)}{\beta_2}\tilde{h}(s) & s \in [\tilde{s}_1, \tilde{s}_2] \\ \dfrac{Z(z^B, b, t^*+s)}{\beta_2}\left[\dfrac{\Delta - r_2^B b_z s}{\hat{z}^A(a+b-1)}\right]^N & s \in [\tilde{s}_2, \tilde{s}_3], \end{cases} \tag{21}$$

and $\tilde{\varphi}_1^B(s) = 0$ for $s \in [s_1, s_3]\backslash[\tilde{s}_1, \tilde{s}_2]$ and $\tilde{\varphi}_2^B(s) = 0$ for $s \in [s_1, s_3]\backslash[\tilde{s}_1, \tilde{s}_3]$.

We now discuss the precision of the approximation $\tilde{s}_1$ when $v_2 \geq \underline{v}_2(0)$. Then, the point t_1 is also the left limit of the pooling region. Hence, $\varphi_2^B(t_1) = 0$. Since $\tilde{\varphi}_2^B(s)$ is a good approximation for $\varphi_2^B(t^*+s)$, an approximation for s_1 is obtained when we solve the equation $\tilde{\varphi}_2^B(s) = 0$. The approximation $\tilde{s}_1$ we obtained earlier is precisely the solution of this last equation. When $\hat{z}^A/\Delta = O(\Delta^2)$, $[\hat{z}^A]^{1+r_2^B/\lambda_1^A}/|s|^{r_2^B/\lambda_1^A} = o(\Delta^2)$ for all $s < s_2$ (note that $s_1 < s_2 < 0$). In this case $W = o(\Delta^2)$ and $s_1 = \tilde{s}_1 + O(\Delta^2)$. The case in which $b = \lfloor b_2^*(a)\rfloor = \lfloor b_2^*(a)\rfloor$ is particularly important for our analysis. Here, a minimum requirement for equilibrium is that $\Phi_2^B(\hat{z}^A, z^B, a, b, v_1, v_2) \leq 1$. If $\varphi^A(a) > 0$ and $(z^A, z^B) \in K(R, \bar{z})$ for $\bar{z}$ small, this inequality is satisfied only if $\Delta^{N+1}/[\hat{z}^A]^{N-1} = O(1)$ (see equation (23) below). Since $N > 1$ when $b = \lfloor b_2^*(a)\rfloor$, Δ is small when $\hat{z}^A$ is small. Moreover, if the grid $\mathbb{B}$ is sufficiently fine so that $N = r^A(1 - \lfloor b_2^*(a)\rfloor)/[r_2^B(1-a)] \leq 2$, this implies that $\hat{z}^A/\Delta$ is of order less than $O(\Delta^2)$.

We would also like to obtain good approximations for the integrals of φ_k^B, $k = 1, 2$. We do so by integrating $\tilde{\varphi}_k^B$, $k = 1, 2$. For the rest of this section, we assume that $v_2 \geq \underline{v}_2(0)$ and hence assume that $\tilde{s}_1 = (\lambda_1^A + r_2^B)\tilde{s}_2/r_2^B$. We consider the case $v_2 < \underline{v}_2(0)$ in Lemma 12, and there we only obtain a bound for the relevant integrals.

Both $\tilde{g}$ and $\tilde{h}$ are functions of the form $f(s) = (-s)^n(X\Delta + Ys)^m$ for some constants $n > 0$ and $m > 0$. We have that

$$\int_{\tilde{s}_1}^{\tilde{s}_2} f(s)ds = [X\Delta]^m \int_{\tilde{s}_1}^{\tilde{s}_2} (-s)^n \left[1 + \frac{Y}{X\Delta}s\right]^m ds = \frac{[X\Delta]^{n+m+1}}{Y^{n+1}} \int_{1+\frac{Y}{X\Delta}\tilde{s}_1}^{1+\frac{Y}{X\Delta}\tilde{s}_2} (1-t)^n\, t^m dt,$$

with the change of variables $t = 1 + Ys/[X\Delta]$. Now

$$\int (1-t)^n\, t^m dt = \frac{t^{m+1}}{m+1} H(t,m,n), \quad 1 + \frac{Y}{X\Delta}\tilde{s}_1 = 0 \quad \text{and} \quad 1 + \frac{Y}{X\Delta}\tilde{s}_2 = \frac{\lambda_1^A}{\lambda_1^A + r_2^B},$$

where $H(t,m,n)$ is the hypergeometric function (usually denoted by ${}_2F_1(m+1, -n, m+2, t)$) defined by the series expansion:

$$H(t,m,n) = 1 + \sum_{k=1}^{\infty} h_k t^k \quad \text{where} \quad h_k = \frac{m+1}{m+k+1} \frac{(-n)(1-n)\cdots(k-1-n)}{k!}.$$

Since $H(0) = 1$,

$$\int_{\tilde{s}_1}^{\tilde{s}_2} f(s)ds = \frac{[X\Delta]^{n+m+1}}{(m+1)Y^{n+1}} \left[t^{m+1} H(t,m,n)\right]_{t=\lambda_1^A/(\lambda_1^A + r_2^B)}.$$

Since $\Delta - r_2^B b_z \tilde{s}_3 = 0$, it follows that

$$\begin{aligned} \int_{\tilde{s}_2}^{\tilde{s}_3} [\Delta - r_2^B b_z s]^N ds &= \frac{[\Delta - r_2^B b_z \tilde{s}_2]^{N+1}}{[N+1] r_2^B b_z} \\ &= \frac{1}{r_2^B b_z} \frac{\lambda_2^A}{\lambda_2^A + \lambda^B} \left[\frac{r_1^B \Delta}{r_1^B - r_2^B}\right]^{N+1}. \end{aligned}$$

We assumed that $\pi^B(b,t) \geq \underline{\pi}$ is continuous in $t \in [0,T]$ for each b. Thus, $Z(z^B, b, t)$ is absolutely continuous in t. Note that $\tilde{s}_1 = -\omega_1\Delta$ and $\tilde{s}_3 = \omega_2\Delta$ for some positive constants ω_1 and ω_2. Thus, for each $\delta > 0$ there exists $\Delta \in (0,\delta)$ such that $Z(z^B, b, t+s) = Z(z^B, b, t)(1 + O(\delta))$ for all $s \in [\tilde{s}_1, \tilde{s}_3]$ and all $t \in [-\tilde{s}_1, T - \tilde{s}_3]$. Therefore

$$\begin{aligned} \int_{s_1}^{s_2} \varphi_1^B(t^* + s)ds &= \int_{\tilde{s}_1}^{\tilde{s}_2} \frac{Z(z^B, b, t^* + s)}{\beta_1} [\tilde{g}(s) - \tilde{h}(s)](1 + O(\Delta))ds \\ &= \frac{Z(z^B, b, t^*)}{\beta_1}(1 + O(\epsilon)) \int_{\tilde{s}_1}^{\tilde{s}_2} [\tilde{g}(s) - \tilde{h}(s)]ds. \end{aligned}$$

Thus, for any $\delta > 0$, there exists $\Delta \in (0, \delta)$ such that

$$\int_{s_1}^{s_2} \varphi_1^B(t^* + s)ds = \frac{Z(z^B, b, t^*)}{\beta_1[\hat{z}^A]^N} \Delta^{N+1}\theta_1[1 + O(\delta)] \tag{22}$$

$$\int_{s_1}^{s_3} \varphi_2^B(t^* + s)ds = \frac{Z(z^B, b, t^*)}{\beta_2[\hat{z}^A]^N} \Delta^{N+1}\theta_2[1 + O(\delta)] \quad \text{where} \tag{23}$$

$$x = \frac{\lambda_1^A}{\lambda_1^A + r_2^B}, \quad \rho(m, n) = \frac{x^{m+1}H(x, m, n)}{(m+1)Y^{n+1}}, \quad J = \frac{\lambda_1^A}{(r_1^B - r_2^B)(1-a)}$$

$$\theta_1 = X\left[\frac{J}{x}\right]^N [\rho(m_1, n_1) - \rho(m_2, n_2)]$$

$$\theta_2 = X\left[\frac{J}{x}\right]^N [\rho(m_2, n_2) + \frac{\lambda_1^A}{\lambda_2^A + \lambda^B}\, x^{N+1}]$$

Define

$$\theta(a, b) = \frac{\theta_1}{\theta_2} \qquad (a, b) \in \mathbb{A} \times \mathbb{B} \quad \textit{with} \quad a + b > 1.$$

The functions θ_1 and θ_2 depend on a and b, and are independent of (v_1, v_2). Furthermore, they also depend on $\hat{z}^A$ but only through b_z (note that X depends on b_z). Since b_z cancels out when we take the ratio, $\theta(a, b)$ is indeed a function of (a, b) alone.

We now continue with the proofs of Lemmas 12–14.

Proof of Lemma 12

Let $b = \lfloor b_2^*(a) \rfloor$ and $\Delta = b_z - v_2 e^{r_2^B t^*}$. By assumption, $0 < \Delta < \delta$. In the previous linearization section we defined $s_i = t_i - t^*$ and show that $s_i = O(\Delta)$ for $i = 1, 2, 3$ (see equations (15)–(17)). Therefore, we can choose $\delta > 0$ sufficiently small so that $|s_i| < \epsilon$ for $i = 1, 2, 3$. Now, if $v_2 \le \underline{v}_2(0)$, then $t_1 = 0$ (by definition) and $t^* = t^* - t_1$. Hence, $t^* = O(\Delta)$ and $v_1 = e^{-r_1^B t^*} b_z \ge b_z(1 - O(\Delta))$. Again, we can choose $\delta > 0$ sufficiently small so that $v_1 \ge b_z - \epsilon$. □

Proof of Lemma 13

Note that $\varphi^A(a|z) \le 1$ implies that

$$\hat{z}^A \ge \frac{z^A \pi^A(a)}{z^A \pi^A(a) + (1 - z^A)} \ge z^A \pi^A(a) \ge z^A \underline{\pi}.$$

Since $\lambda_1^A > \lambda^B$ and $z^A \ge z^B/R$, (4) implies that there exists $\bar{z}$ such that for all $z \in K(R, \bar{z})$ and each $t \in (0, t_1)$,

$$\varphi_1^B(t) \le \frac{R\,\pi^B(\lfloor b_2^*(a) \rfloor, t)}{\beta_1} \left[\frac{\bar{\mu}_1(t)}{\underline{\pi}}\right]^{\lambda^B/\lambda_1^A} \bar{z}^{1-\lambda^B/\lambda_1^A}.$$

Therefore, for some constant $M > 0$,

$$\int_0^{t_1^\ell} \varphi_1^B(t)dt \leq M\,\bar{z}^{1-\lambda^B/\lambda_1^A},$$

and clearly we can choose $\bar{z}$ small enough so that $M\,\bar{z}^{1-\lambda^B/\lambda_1^A} \leq \delta$.

By Lemma 11, $\varphi_1^B(t) = 0$ for $t \in (t_2, t_3)$. Hence, by the proof of Lemma 9 (which uses the assumption that $\varphi^A(a|z) \geq \epsilon$), the condition $\int_{t_2}^{t_3} \varphi_2^B(t) \leq \Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)$ $= 1$ implies $v_2 e^{r_2^B t^*} \geq b_z - \delta$ when z is small enough. □

Proof of Lemma 14

Choose $\delta > 0$ as required by Lemma 12 so that $|t_i - t^*| < \epsilon$, $i = 1, 2, 3$. Since $b_2 = \lfloor b_2^*(a) \rfloor = \lfloor b_2^*(a) \rfloor$, $\theta(a) \equiv \theta(a, \lfloor b_2^*(a) \rfloor)$ is only a function of a. By assumption, $v_2 \geq \underline{v}_2(0)$, so $\tilde{s}_1 < -t^*$ and the equations (22)–(23) are valid. Thus,

$$\frac{\int_{t_1}^{t_2} \varphi_1^B(t|\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)dt}{\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v_1, v_2)} = \frac{\beta_2}{\beta_1}\,\theta(a) + O(\delta),$$

and we can choose $\delta > 0$ small enough so that $|O(\delta)| < \epsilon$. □

Proof of Lemma 15

Assume that $v_2 < \underline{v}_2(0)$, so that $v_2 e^{r_2^B t}$ and $\underline{v}_2(t)$ do not intersect in $[0, t^*)$. In this case $t_1 = 0$ and $s_1 = -t^*$ is not the left limit of integration assumed in equations (22)–(23), and $(0, t_1) = \emptyset$. When $s_1 = -t^*$, the pooling region is "truncated" in the left at $t = 0$. For $s > 0$, let $(\tilde{v}_1, \tilde{v}_2)$ be such that $\tilde{v}_k e^{r_k^B s} = v_k$ for $k = 1, 2$. Then $\tilde{v}_k e^{r_k^B (s+t)} = v_k e^{r_k^B t}$ for $t \geq 0$, $k = 1, 2$. Thus, adjusting $(\tilde{v}_1, \tilde{v}_2)$ this way corresponds to "moving the vertical axes" and extending the original diagram to the left. Corresponding to $(\tilde{v}_1, \tilde{v}_2)$ there is a new sneaking-in function that also satisfies $\underline{v}_2(s + t|\tilde{v}) = \underline{v}_2(t|v)$ for all $t \geq 0$. Since the indifference curve $\tilde{v}_2 e^{r_2^B t}$ is steeper than the sneaking-in function $\underline{v}_2(t|\tilde{v})$ at each t, there exists $s^* > 0$ large enough so that when $s = s^*$, $\tilde{v}_2 = \underline{v}_2(0|\tilde{v})$. Let v_k^* be such that $v_k^* e^{r_k^B s^*} = v_k$, $k = 1, 2$. Then $\varphi_k^B(s^* + t|v^*) = \varphi_k^B(t|v)$ for all $t \geq 0$, $k = 1, 2$. By Lemma 14, for any $\epsilon > 0$ small we can choose $\delta > 0$ such that

$$\frac{\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v^*)}{\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v^*)} < \frac{\beta_2}{\beta_1}\,\theta(a) + \epsilon < 1.$$

Assume by contradiction that $\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v)/\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v) \geq 1$. The function $\varphi_1^B(\cdot|v^*) : [0, s^* + t_2]$ is quasiconcave and attains a maximum at some point in $(0, s^* + t_2)$. The function $\varphi_2^B(\cdot|v^*) : [0, s^* + t_3]$ is quasiconcave and attains its maximum at $s^* + t_2$. Moreover, $\varphi_1^B(0|v^*) > 0 = \varphi_2^B(0|v^*)$ and $\varphi_2^B(s^* + t_2|v^*) > 0 = \varphi_1^B(s^* + t_2|v^*)$. Hence, there exists a unique $t_0 \in (0, s^* + t_2)$ such that $\varphi_1^B(t_0|v^*) = \varphi_2^B(t_0|v^*)$, and

$\varphi_1^B(t|v^*) > \varphi_2^B(t|v^*)$ for $t \in (0, t_0)$ and $\varphi_1^B(t|v^*) < \varphi_2^B(t|v^*)$ for $t \in (t_0, s^* + t_3)$. Therefore, $\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v)/\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, v) \geq 1$ implies that $t_0 - s^* > 0$.

As we move the vertical axes to the left, s increases from 0 to s^* and $(\tilde{v}_1, \tilde{v}_2)$ decreases from (v_1, v_2) to (v_1^*, v_2^*). We now study how the ratio

$$\gamma(s) = \frac{\Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, \tilde{v})}{\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, \tilde{v})}$$

varies with s. By assumption, when $s = 0$, $\tilde{v} = v$ and $\gamma(0) \geq 1$. We now argue that $\gamma(s) > 1$ for all $s \in (0, s^*]$. Suppose that at some $s \in (0, s^*]$, $\gamma(s) = 1$. Let $\underline{s}$ be the smallest such point and $\underline{v}_k$ be such that $\underline{v}_k e^{r_k^B \underline{s}} = v_k$, $k = 1, 2$. Then, $\gamma(s) > 1$ for $s \in (0, \underline{s})$, and

$$\begin{aligned}
\gamma'(\underline{s}) &= \frac{\varphi_1^B(s^* - \underline{s}|v^*)\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, \underline{v}) - \Phi_1^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, \underline{v})\varphi_2^B(s^* - \underline{s}|v^*)}{\Phi_2^B(\hat{z}^A, z^B, \underline{v})^2} \\
&= \frac{1}{\Phi_2^B(\hat{z}^A, z^B, a, \lfloor b_2^*(a) \rfloor, \underline{v})}[\varphi_1^B(s^* - \underline{s}|v^*) - \gamma(\underline{s})\varphi_2^B(s^* - \underline{s}|v^*)] > 0.
\end{aligned}$$

Thus, $\gamma(s) < 1$ for s just smaller than $\underline{s}$, which is a contradiction. Therefore, $\gamma(s) > 1$ for all $s \in (0, s^*]$. But this implies that $\gamma(s^*) > 1$ which is also a contradiction because the assumption $\beta_2\theta(a)/\beta_1 < 1$ implies that $\gamma(s^*) < 1$, as we argued above. $\square$

References

1. D. Abreu and F. Gul, *Bargaining and Reputation*, Econometrica **68** (2000), 85–117.
2. D. Abreu and D. Pearce, *Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts*, Econometrica **75** (2007), 653–710.
3. L. Ausubel, P. Cramton and R. Deneckere, *Bargaining with Incomplete Information*, Handbook of Game Theory **3** (2002), Elsevier Science B.V., Amsterdam, 1897–1945.
4. L. Ausubel and R. Deneckere, *Reputation in Bargaining and Durable Goods Monopoly*, Econometrica **57** (1989), 511–531.
5. S. Bikhchandani, *A Bargaining Model with Incomplete Information*, Review of Economic Studies **59** (1992), 187–203.
6. J. Bulow, *Durable Goods Monopolists*, Journal of Political Economy **90** (1982), 314–322.
7. K. Chatterjee and L. Samuelson, *Bargaining with Two-Sided Incomplete Information: An Infinite Horizon Model with Alternating Offers*, Review of Economic Studies **54** (1987), 175–192.
8. K. Chatterjee and L. Samuelson, *Bargaining under Two-Sided Incomplete Information: The Unrestricted Offers Case*, Operations Research **36** (1988), 605–618.
9. I.-K. Cho, *Uncertainty and Delay in Bargaining*, Review of Economic Studies **57** (1990), 575–596.
10. R. Coase, *Durability and Monopoly*, Journal of Law and Economics **15** (1971), 143–149.
11. O. Compte and P. Jehiel, *On the Role of Outside Options in Bargaining with Obstinate Parties*, Econometrica **70** (2002), 1477–1517.
12. P. Cramton, *Bargaining with Incomplete Information: An Infinite-Horizon Model with Continuous Uncertainty*, Review of Economic Studies **51** (1984), 579–593.
13. R. Deneckere and Liang, *Bargaining with Interdependent Values*, Econometrica **74** (2006), 1309–1364.
14. R. Evans, *Sequential Bargaining with Correlated Values*, Review of Economic Studies **56** (1989), 499–510.
15. J. Fanning, *Reputational Bargaining and Deadlines* (2013), working paper, New York University.
16. W. Fuchs and A. Skrzypacz, *Bargaining with Arrival of New Traders*, American Economic Review **100** (2010), 802–836.
17. W. Fuchs and A. Skrzypacz, *Bridging the Gap: Bargaining with Interdependent Values* (2012), forthcoming in *Journal of Economic Theory.*
18. D. Fudenberg, D. Levine and J. Tirole, *Infinite-Horizon Models of Bargaining with One-Sided Incomplete Information*, Game-Theoretic Models of Bargaining (A. Roth, eds.), Cambridge University Press, 1985, pp. 73–98.
19. D. Fudenberg and J. Tirole, *Sequential Bargaining with Incomplete Information*, Review of Economic Studies **50** (1983), 221-247.
20. F. Gul and H. Sonnenschein, *On Delay in Bargaining with One-Sided Uncertainty*, Econometrica **56** (1988), 601–611.
21. F. Gul, H. Sonnenschein and R. Wilson, *Foundations of Dynamic Monopoly and the Coase Conjecture*, Journal of Economic Theory **39** (1986), 155–190.
22. R. Inderst, *Bargaining with a Possibly Committed Seller*, Review of Economic Dynamics **8** (2005), 927–944.
23. C. M. Kahn, *The Durable Goods Monopolist and Consistency with Increasing Costs*, Econometrica **54** (1986), 275–294.
24. S. Kambe, *Bargaining with Imperfect Commitment*, Games and Economic Behavior **28** (1999), 217–237.
25. K. Kim, *The Coase Conjecture with Incomplete Information on the Monopolist's commitment*, Theoretical Economics **4** (2009), 17–44.
26. D. Kreps and R. Wilson, *Reputation and Imperfect Information*, Journal of Economic Theory **27** (1982), 253–279.

27. D. Kreps, P. Milgrom, J. Roberts and R. Wilson, *Rational Cooperation in the Finitely Repeated Prisoner's Dilemma*, Journal of Economic Theory **27** (1982), 245–252.
28. G. Mailath and L. Samuelson, *Repeated Games and Reputations: Long-Run Relationships*, Oxford University Press, 2006.
29. P. McAfee and T. Wiseman, *Capacity Choice Counters the Coase Conjecture*, Review of Economic Studies **75** (2008), 317–332.
30. P. Milgrom and J. Roberts, *Predation, Reputation and Entry Deterrence*, Journal of Economic Theory **27** (1982), 280–312.
31. R. Myerson, *Game Theory: Analysis of Conflict*, Harvard University Press, 1991.
32. J. Ortner, *Durable Goods Monopoly with Stochastic Costs* (2012), working paper.
33. A. Rubinstein, *Perfect Equilibrium in a Bargaining Model*, Econometrica **50** (1982), 97–109.
34. A. Rubinstein, *A Bargaining Model with Incomplete Information about Time Preferences*, Econometrica **53** (1985), 1151–1172.
35. J. Sobel, *Durable Goods Monopoly with Entry of New Consumers*, Econometrica **59** (1991), 1455–1485.
36. J. Sobel and I. Takahashi, *A Multi-Stage Model of Bargaining*, Review of Economic Studies **50** (1983), 411–426.
37. N. Stokey, *Rational Expectations and Durable Goods Pricing*, Bell Journal of Economics **12** (1981), 112–128.
38. J. Watson, *Alternating-Offer Bargaining with Two-Sided Incomplete Information*, Review of Economic Studies **65** (1998), 573–594.
39. A. Wolitzky, *Indeterminacy of Reputation Effects in Repeated Games with Contracts*, Games and Economic Behavior **73** (2011), 595–607.