Reputational Bargaining Under Knowledge of Rationality*

Alexander Wolitzky

M.I.T.

February 26, 2011

Abstract

Two players announce bargaining postures to which they may become committed and then bargain over the division of a surplus. The share of the surplus that a player can guarantee herself under first-order knowledge of rationality is determined (as a function of her probability of becoming committed), as is the bargaining posture that she must announce in order to guarantee herself this much. This "maxmin" share of the surplus is large relative to the probability of becoming committed (e.g., it equals 30% if the commitment probability is 1 in 10, and equals 7% if the commitment probability is 1 in 1 million), and the corresponding bargaining posture simply demands this share plus compensation for any delay in reaching agreement. The paper relates the outcome of the model to the outcomes of a broad class of discrete-time bargaining procedures with frequent offers.

Keywords: *bargaining, knowledge of rationality, posturing, reputation*

*I thank my advisors, Daron Acemoglu, Glenn Ellison, and Muhamet Yildiz, for extensive discussions and comments as well as continual advice and support. For additional helpful comments, I thank Sandeep Baliga, Abhijit Banerjee, Alessandro Bonatti, Jeff Ely, Drew Fudenberg, Bob Gibbons, Michael Grubb, Peter Klibanoff, Anton Kolotilin, Mihai Manea, Parag Pathak, Jeroen Swinkels, Juuso Toikka, Iván Werning, and (especially) Gabriel Carroll, as well seminar audiences at Berkeley, Booth, Chicago, Harvard, Kellogg, Microsoft Research, Minnesota, M.I.T., Penn, Princeton, Stanford, Stanford GSB, and Yale. I thank the NSF for financial support.

1 Introduction

Economists have long been interested in how individuals split gains from trade. The division of surplus often determines not only equity, but also efficiency, as it affects individuals' ex ante incentives to make investments; this effect of surplus division on efficiency is a major theme of, for example, property rights theories of the firm (Grossman and Hart, 1986), industrial organization models of cumulative innovation (Green and Scotchmer, 1995), and search-and-matching models of the labor market (Hosios, 1990). Recently, "reputational" models of bargaining have been developed that make sharp prediction about the division of the surplus independently of many details of the bargaining procedure (Myerson, 1991; Abreu and Gul, 2000; Kambe, 1999; Compte and Jehiel, 2002). In these models, players may be committed to a range of possible bargaining strategies, or "postures," before the start of bargaining, and bargaining consists of each player attempting to convince her opponent that she is committed to a strong posture. These models assume that the probabilities with which the players are committed to various bargaining postures (either ex ante or after a stage where players strategically announce bargaining postures) are common knowledge, and that play constitutes a (sequential) equilibrium. In this paper, I study reputational bargaining while assuming only that the players know that each other is rational, and show that each player can guarantee herself a relatively large share of the surplus—even if her probability of being committed is small—by announcing the posture that simply demands this share plus compensation for any delay in reaching agreement. Furthermore, announcing any other posture does not guarantee as much.

A key feature of my model is the existence of a positive number ε such that, if a player announces any bargaining posture (i.e., any infinite path of demands) at the beginning of the game, she then becomes committed to that posture with probability at least ε (or, equivalently, she convinces her opponent that she is committed to that posture with probability at least ε). I derive the highest payoff that a player can guarantee herself by announcing any posture, regardless of her opponent's beliefs about her bargaining strategy, *so long as her opponent is rational and believes that she is committed to her announced posture with probability at least ε*. More precisely, player 1's "highest guaranteed," or "maxmin," payoff is the highest payoff u_1 with the property that there exists a corresponding posture (the "maxmin posture") and bargaining strategy such that player 1 receives at least u_1 whenever she announces this posture and follows this strategy and player 2 plays *any* best-response to *any* belief about player 1's strategy that assigns probability at least ε to player 1 following her announced posture. In particular, player 2 need not play a best-response to player 1's actual strategy, or vice versa; thus, player 1's maxmin payoff is below her lowest Nash equilibrium payoff.

The main result of this paper characterizes the maxmin payoff and posture when only one player may become committed to her announced posture; as discussed below, a very similar characterization applies when both players may become committed. While the maxmin payoff may be very small when ε is small in general two-person games, it is relatively large in my model: in particular, it equals $1/(1-\log\varepsilon)$. This equals 1 when $\varepsilon = 1$ (i.e., when the player makes a take-it-or-leave-it offer) and, more interestingly, goes to 0 very slowly as ε goes to 0 (more precisely, it goes to 0 at a logarithmic rate, which is slower than any polynomial rate). For example, a bargainer can guarantee herself approximately 30% of the surplus if her commitment probability is 1 in ten; 13% if it is 1 in 1 thousand; and 7% if it is 1 in 1 million. The second part of the main result is that the unique bargaining posture that guarantees this share of the surplus simply demands this share in addition to compensation for any delay; that is, it demands a share of the surplus that increases at rate equal to the discount rate, r. This compensation amounts to the entire surplus after a long enough delay, so the unique maxmin posture demands

$$\min\left\{e^{rt}/(1-\log\varepsilon), 1\right\}$$

at every time t. This posture is depicted in Figure 1, for commitment probability $\varepsilon = 1/1000$ and discount rate $r = 1$.

The intuition for the result that the unique maxmin posture demands compensation for delay involves two key ideas. First, when player 2's beliefs are those that lead him to reject player 1's demand for as long as possible, player 1's demand is accepted sooner when it is lower. This is analogous to the argument in the existing reputational bargaining literature that player 1 builds reputation more quickly *in equilibrium* when her current demand is

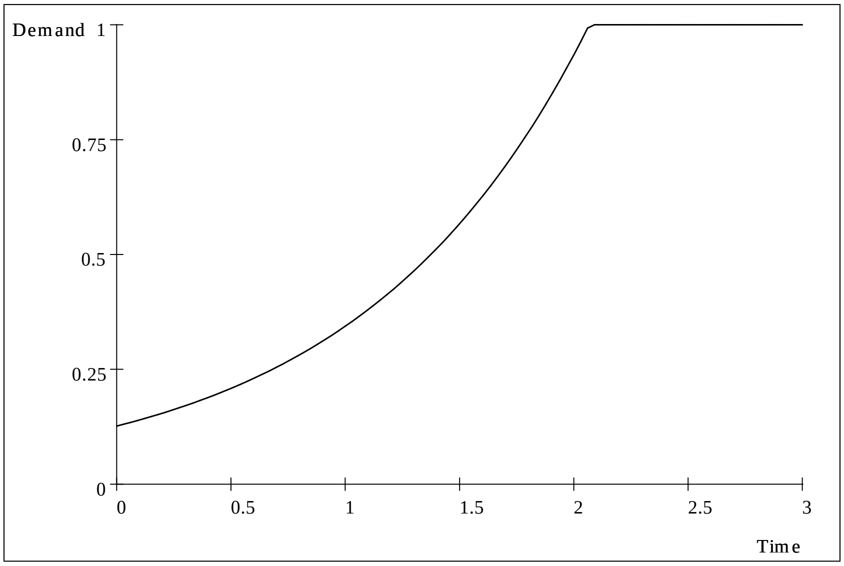

Figure 1: The Unique Maxmin Bargaining Posture for $\varepsilon = 1/1000$ and $r = 1$

lower, though my analysis is not based on equilibrium. Second, the maxmin posture can never make demands that would give player 1 less than her maxmin payoff if they were accepted, because player 2 could simply accept some such demand and give player 1 a payoff below her maxmin payoff, which was supposed to be guaranteed to player 1 (though it must be verified that such behavior by player 2 is rational). Combining these ideas implies that player 1 must always demand at least her maxmin level of *utility* (hence, compensation for delay), but no more.

Three distinctive features of my approach are the timing of commitment (players freely choose which bargaining postures to announce, but may become bound by their announcements), the range of bargaining postures players may announce (all possible paths of demands for the duration of bargaining), and the solution concept (first-order knowledge of rationality). The timing of commitment is appropriate if players are rational but may credibly announce bargaining postures. This assumption has many precedents in the literature, starting with Schelling (1956),[1] who discusses observable factors that make announced pos-

[1]For game-theoretic models, see Crawford (1982), Fershtman and Seidmann (1993), Muthoo (1996), Ellingsen and Miettinen (2008), and (especially) Kambe (1999). Staw (1981, 1997) discusses psychological and sociological factors that lock individuals and organizations into costly courses of actions.

tures more credible, corresponding to a higher value of ε in my model.[2] The assumption that a player may announce any path of demands and that all such announcements are equally credible seems unappealing a priori, because announcing a "simple" posture may be more credible than announcing a "complicated" posture. Fortunately, the unique maxmin posture is simply announcing, "I want a certain share of the surplus, and if you make me wait to get it then you must compensate me for the delay."[3] Thus, allowing players to credibly announce complicated postures ensures that my characterization of the maxmin posture does not depend on ad hoc restrictions on the range of credible postures, but my characterization would still apply if only "simple" postures were credible. In addition, the techniques I develop would allow one to characterize the maxmin payoff and posture in a more general model where credibility varies across announcements.

A player's maxmin payoff is her lowest payoff consistent with (first-order) knowledge of rationality (at the start of the game). In bargaining, a player cannot guarantee herself any positive payoff without knowledge of her opponent's rationality, as, for example, she receives payoff 0 if her opponent always rejects her offer and demands the entire surplus. Thus, knowledge of rationality is the weakest solution concept consistent with positive guaranteed payoffs. Furthermore, I show that any feasible payoff greater than the maxmin payoff is consistent with knowledge of rationality, which implies that the maxmin payoff and posture are the key objects of interest under knowledge of rationality.

Imposing only knowledge of rationality rather than a stronger solution concept, such as rationalizability or equilibrium, leads to more robust predictions. In particular, predictions under knowledge of rationality (such as the prediction that each player receives at least her

[2] For example, an announcement is more credible if the stakes in the current negotiation are small relative to the stakes in potential future negotiations; if the announcement is observable to a large number of third parties; if the bargainer can side-contract with third parties to bind herself to her announcements; if the bargainer may be acting as an agent for a third party and does not have independent authority to change her posture; or if the bargainer displays emotions that suggest an unwillingness to modify her posture.

[3] To my knowledge, this is the first bargaining model that predicts that such a posture will be adopted, though it seems like a reasonable bargaining position to stake out. For example, in most U.S. states defendants must pay "prejudgment interest" on damages in torts cases, which amounts to plaintiffs demanding the initial damages in addition to compensation for any delay (e.g., Knoll, 1996); similarly, unions sometimes include payment for strike days among their demands.

maxmin payoff) do not depend on each player's beliefs about her opponent's strategies, so long as each player believes that her opponent is playing a best-response to *some* belief about her own play; and also do not depend on unmodelled strategic considerations that do not affect a player's payoffs or her beliefs about her opponent's payoffs, but may affect higher-order beliefs about payoffs. Such considerations arise naturally in bargaining: for example, a local union that strategically announces that it will strike until offered a wage of at least \$25 an hour may be concerned that the firm's management may believe that it is actually required by the national union to strike until offered a wage of at least \$20 an hour. The maxmin payoff and posture may also be viewed as the predicted payoff and posture of a positive theory of bargaining in which each player is either maximally pessimistic (in a Bayesian sense) about her opponent's strategy or expects her opponent to play her "worst-case" strategy (in a maxmin sense), given her knowledge of her opponent's rationality. These two approaches are equivalent in my model, though they differ in general games.

I consider two main extensions of the model. First, I characterize the maxmin payoffs and postures when both players may become committed to their announced postures. I find that each player's maxmin posture is exactly the same as in the one-sided commitment model, and that each player's maxmin payoff is close to her maxmin payoff in the one-sided commitment model as long as her opponent's commitment probability is small. Thus, the one-sided commitment analysis applies to each player separately.

Second, I consider the role of the (continuous-time or discrete-time) bargaining procedure. Here, I provide a result showing that the maxmin payoff and posture are robust to details of the bargaining procedure such as the order and relative frequency of offers, so long as both players have the opportunity to make offers frequently.[4]

The paper proceeds as follows: Section 2 relates this paper to the literature. Section 3 presents the model and defines maxmin payoffs and postures. Section 4 analyzes the baseline case with one-sided commitment and presents the main characterization of maxmin payoffs and postures. Section 5 presents four brief extensions. Section 6 considers two-

[4] Abreu and Gul (2000) prove an analogous "independence of procedures" result in an equilibrium reputational bargaining model. No such result holds in complete-information bargaining models in the tradition of Rubinstein (1982).

sided commitment. Section 7 considers discrete-time bargaining games with frequent offers. Section 8 concludes. Omitted proofs for Sections 3 and 4 are in the appendix, and omitted proofs for Sections 5 and 7 are in the supplementary appendix.

2 Related Literature

The seminal paper on reputational bargaining is Abreu and Gul (2000), which generalizes Myerson (1991).[5] In their model, there is a vector of probabilities for each player corresponding to the probability that she is committed to each of a variety of behavioral types (which are analogous to bargaining postures in my model), and these vectors are common knowledge. In the (effectively unique) sequential equilibrium, players randomize over mimicking different behavioral types, with mixing probabilities determined by the prior, and play proceeds according to a war of attrition, where each player hopes that her opponent will concede. A player's equilibrium payoff is higher when she is more likely to be committed to strong behavioral types and when she is more patient. Thus, Abreu and Gul present a complete and elegant bargaining theory in which the bargaining procedure is unimportant and sharp predictions are driven by the vector of prior commitment probabilities.

The main difference between my analysis and Abreu and Gul's is that I characterize maxmin payoffs and postures rather than sequential equilibria. My approach entails weaker assumptions on knowledge of commitment probabilities (i.e., second-order knowledge that each player is committed to her announced posture with probability at least ε, rather than common knowledge of a vector of commitment probabilities) and on behavior (i.e., first-order knowledge of rationality, rather than sequential equilibrium), and does not yield unique predictions about the division of surplus or about the details of how bargaining will proceed. One motivation for this complementary approach is that behavioral types are sometimes viewed as "perturbations" reflecting the fact that a player (or an outside observer) cannot be sure that the model captures all of the other player's strategic considerations. Thus, it seems

[5] Other important antecedents of Abreu and Gul (2000) include Kreps and Wilson (1982) and Milgrom and Roberts (1982), who pioneered the incomplete information approach to reputation-formation, and Chatterjee and Samuelson (1987, 1988), who study somewhat simpler reputational bargaining models.

reasonable to assume that players realize that their opponents' type may be perturbed in some manner (e.g., that a rich set of types have positive prior weight), but assuming that the distribution over perturbations is common knowledge goes against the spirit of introducing perturbations.[6]

The paper most closely related to mine is Kambe (1999). In Kambe's model, each player first strategically announces a posture and then becomes committed to her announced posture with probability ε, as in my model. Thus, Kambe endogenizes the behavioral types of Abreu and Gul. The structure of equilibrium and the determinants of the division of the surplus are similar to those in Abreu and Gul's model. There are two differences between Kambe's model and mine. First, Kambe requires that players announce postures that demand a constant share of the surplus (as do Abreu and Gul), while I allow players to announce non-constant postures (and players do benefit from announcing non-constant postures in my model). Second, and more fundamentally, Kambe studies sequential equilibria, while I study maxmin payoffs and postures. These differences lead my analysis and results to be quite different from Kambe's, with the exception that Kambe's calculation of bounds on the set of sequential equilibrium payoffs resembles my calculation of the maxmin payoff in the special case where players can only announce constant postures (Section 5.1).

There are also a number of earlier bargaining models in which players try to commit themselves to advantageous postures. Crawford (1982) studies a two-stage model in which players first announce demands and then learn their private costs of changing these demands, and shows that such a model can lead to impasse. Fershtman and Seidmann (1993) show that agreement is delayed until an exogenous deadline if each player is unable to accept an offer that she has previously rejected. Muthoo (1996) studies a two-stage model related to Crawford (1982), with the feature that making a larger change to one's initial demand is more costly, and shows that a player's equilibrium payoff is increasing in her marginal cost

[6] In games with a long-run player facing a series of short-run players, Watson (1993) and Battigalli and Watson (1997) show that common knowledge of the mere fact that the long-run player is committed to a certain strategy with probability bounded away from 0 determines the division of the surplus. However, Wolitzky (2011) shows that common knowledge of the relative probabilities with which each player is committed to each strategy is needed for equilibrium selection in games with two long-run players, even when binding contracts are available (as is the case in bargaining).

of changing her demand. Ellingsen and Miettinen (2008) point out that if commitment is costly in Crawford's model then impasse not only can result, but must. These papers study equilibrium and do not involve reputation formation.

Finally, this paper is also related to the literature on bargaining with incomplete information either without common priors (Yildiz 2003, 2004; Feinberg and Skrzypacz, 2005) or with rationalizability rather than equilibrium (Cho, 1994; Watson, 1998), in that players may disagree about the distribution over outcomes of bargaining. I briefly discuss a connection with the literature on reputation in repeated games in the conclusion.

3 Model and Key Definitions

This section describes the model and defines maxmin payoffs and postures, which are the main objects of analysis.

3.1 Model

Two players ("she," "he") bargain over one unit of surplus in two phases: a "commitment phase" followed by a "bargaining phase." I describe the bargaining phase first. It is intended to capture a continuous bargaining process where players can change their demands and accept their opponents' demands at any time, but in order to avoid well-known technical issues that emerge when players can condition their play on "instantaneous" actions of their opponents (Simon and Stinchcombe, 1989; Bergin and MacLeod, 1993) I assume that players can revise their paths of demands only at integer times (while letting them accept their opponents' demands at any time).

Time runs continuously from $t = 0$ to ∞. At every integer time $t \in \mathbb{N}$ (where $\mathbb{N}$ is the natural numbers), each player $i \in \{1, 2\}$ chooses a path of demands for the next length-1 period of time, $u_i^t : [t, t+1) \to [0, 1]$, which is required to be the restriction to $[t, t+1)$ of a continuous function on $[t, t+1]$. Let $\mathcal{U}^t$ be the set of all such functions. The interpretation is that $u_i^t(\tau)$ is the demand that player i makes at time τ (this is simply denoted by $u_i(\tau)$ when t is understood; note that $u_i(\tau)$ can be discontinuous at integer times but is right-continuous everywhere[7]). Even though player i's path of demands for $[t, t+1)$ is decided at t, player j only observes demands as they are made. Intuitively, each player i may *accept* her opponent's demand $u_j(t)$ at any time t, which ends the game with payoffs $\left(e^{-rt}(1-u_j(t)), e^{-rt}u_j(t)\right)$, where $r \in \mathbb{R}_+$ is the common discount rate (throughout, $j = -i$). Formally, every instant of time t is divided into three *dates*, $(t, -1)$, $(t, 0)$, and $(t, 1)$, with the following timing:[8] First, at date $(t, -1)$, each player i announces *accept* or *reject*. If both players reject, the game continues; if only player i accepts, the game ends with payoffs $\left(e^{-rt}(1-\lim_{\tau\uparrow t} u_j(\tau)), e^{-rt}\lim_{\tau\uparrow t} u_j(\tau)\right)$; and if both players accept, the games ends with payoffs determined by the average of the two demands, $\lim_{\tau\uparrow t} u_1(\tau)$ and $\lim_{\tau\uparrow t} u_2(\tau)$. Next, at date $(t, 0)$, both players simultaneously announce their time-t demands $(u_1(t), u_2(t))$ (which were determined at the most recent integer time); if t is an integer, this is also the date where each player chooses a path of demands for the next length-1 period. Finally, at date $(t, 1)$, each player i again announces *accept* or *reject*. If both players reject, the game continues; if only player i accepts, the game ends with payoffs $\left(e^{-rt}(1-u_j(t)), e^{-rt}u_j(t)\right)$; and if both players accept, the game ends and the demands $u_1(t)$ and $u_2(t)$ are averaged. This timing ensures that there is a first and last date at which each player can accept each of her opponent's demands. In particular, at integer time t, player i may accept *either* her opponent's "left" demand, $\lim_{\tau\uparrow t} u_j(\tau)$, or her time-$t$ demand, $u_j(t)$.

The public history up to time t excluding the time-t demands is denoted by $h^{t-} = (u_1(\tau), u_2(\tau))_{\tau<t}$, and the public history up to time t including the time-t demands is denoted by $h^{t+} = (u_1(\tau), u_2(\tau))_{\tau\leq t}$ (with the convention that this corresponds to all offers having been rejected, as otherwise the game would have ended). A generic time-t history is denoted by h^t. Since $\lim_{\tau\uparrow t} u_j(\tau) = u_j(t)$ if t is not an integer, I generally distinguish between h^{t-} and h^{t+} only for integer t. Formally, *a bargaining phase (behavior) strategy* for player i is a pair $\sigma_i = (F_i, G_i)$ such that F_i maps histories into $[0, 1]$ with the property that $F_i(h^t) \leq F_i\left(h^{t'}\right)$ whenever $h^{t'}$ is a successor of h^t, and G_i maps histories of the form h^{t-} with $t \in \mathbb{N}$ into $\Delta(\mathcal{U}^t)$. Let Σ_i be the set of player i's bargaining phase strategies.

[7]A function $f : \mathbb{R} \to \mathbb{R}$ is right-continuous if, for every $x \in \mathbb{R}$ and every $\eta > 0$, there exists $\delta > 0$ such that $|f(x) - f(x')| < \eta$ for all $x' \in (x, x+\delta)$.

[8]This is similar to the notion of date introduced by Abreu and Pearce (2007).

The interpretation is that $F_i(h^{t-})$ is the probability that player i accepts player j's demand at or before date $(t,-1)$, $F_i(h^{t+})$ is the probability that player i accepts player j's demand at or before date $(t,1)$, and $G_i(h^{t-})$ is the probability distribution over paths of demands $u_i^t : [t, t+1) \to [0,1]$ chosen by player i at date $(t,0)$. This formalism implies that player i's *hazard rate of acceptance at history* h^t, $f_i(h^t) / (1 - F_i(h^t))$, is well-defined at any time t at which the realized distribution function F_i admits a density f_i (in which case $F_i(h^{t-}) = F_i(h^{t+})$); and in addition player i's *probability of acceptance at history* h^{t+} (resp., h^{t-}), $F_i(h^{t+}) - F_i(h^{t-})$ (resp., $F_i(h^{t-}) - \lim_{\tau \uparrow t} F_i(h^{\tau-})$), is well-defined for all times t. However, so long as one bears in mind these formal definitions, it suffices for the remainder of the paper to omit the notation (F_i, G_i) and instead simply view a (bargaining phase) strategy σ_i as a function that maps every history h^t to a hazard rate of acceptance, a discrete probability of acceptance, and (if $h^t = h^{t-}$ for some $t \in \mathbb{N}$) a probability distribution over paths of demands u_i^t. I say that *agreement is reached at time* t if the game ends at time t (i.e., at date $(t,-1)$ or $(t,1)$). Both players receive payoff 0 if agreement is never reached.

At the beginning of the bargaining phase, player i has an initial belief π_i about the *behavior* of her opponent. That is, $\pi_i \in \Delta(\Sigma_j)$, so π_i is a probability distribution over behavior strategies σ_j; note that π_i can alternatively be viewed as an element of Σ_j by reducing compound lotteries over pure strategies. Let $\mathrm{supp}(\pi_i) \subseteq \Sigma_j$ be the support of π_i, let $u_i(\sigma_i, \sigma_j)$ be player i's expected utility given strategy profile (σ_i, σ_j), let $u_i(\sigma_i, \pi_i)$ be player i's expected utility given strategy σ_i and belief π_i, and let $\Sigma_i^*(\pi_i) \equiv \mathrm{argmax}_{\sigma_i} u_i(\sigma_i, \pi_i)$ be the set of player i's best-responses to belief π_i.

At the beginning of the game (prior to time 0), player 1 (but not player 2) publicly announces a *bargaining posture* $\gamma : [0, \infty) \to [0,1]$, which must be continuous at non-integer times t, be right-continuous everywhere, and have well-defined left limits everywhere. Slightly abusing notation, a posture γ is identified with the strategy of player 1's that demands $\gamma(t)$ for all $t \in \mathbb{R}_+$ and always rejects player 2's demand; with this notation, $\gamma \in \Sigma_1$. In other words, a posture is a pure bargaining phase strategy that does not condition on player 2's play or accept player 2's demand. After announcing posture γ, player 1 becomes committed to γ with some probability $\varepsilon > 0$, meaning that she must follow strategy γ in the bargaining phase. With probability $1 - \varepsilon$, she is free to play any strategy in the bargaining phase. Whether or not player 1 becomes committed to γ is observed only by player 1.

3.2 Defining the Maxmin Payoff and Posture

This subsection defines player 1's maxmin payoff and posture. Intuitively, player 1's maxmin payoff is the highest payoff she can guarantee herself when all she knows about player 2 is that he is rational (i.e., maximizes his expected payoff given his belief about player 1's behavior, and updates his belief according to Bayes' rule when possible) and that he believes that player 1 follows her announced posture γ with probability at least ε.

Formally, that player 2 is rational and assigns probability at least ε to player 1 following her announced posture γ means that his strategy satisfies the following condition:

Definition 1 *A strategy σ_2 of player 2's is* rational given posture γ *if there exists a belief π_2 of player 2's such that $\pi_2(\gamma) \geq \varepsilon$ and $\sigma_2 \in \Sigma_2^*(\pi_2)$.*

A belief π_1 of player 1's is *consistent with knowledge of rationality given posture γ* if every strategy $\sigma_2 \in \operatorname{supp}(\pi_1)$ is rational given posture γ. In other words, the set of beliefs π_1 that are consistent with knowledge of rationality given posture γ is $\Pi_1^\gamma \equiv \Delta\{\sigma_2 : \sigma_2 \text{ is rational given posture } \gamma\}$.

Given that her belief is consistent with knowledge of rationality, the highest payoff that player 1 can guarantee herself after announcing posture γ is the following:

Definition 2 *Player 1's* maxmin payoff given posture γ *is*

$$u_1^*(\gamma) \equiv \sup_{\sigma_1} \inf_{\pi_1 \in \Pi_1^\gamma} u_1(\sigma_1, \pi_1).$$

A strategy $\sigma_1^*(\gamma)$ of player 1's is a maxmin strategy given posture γ *if*

$$\sigma_1^*(\gamma) \in \operatorname*{argmax}_{\sigma_1} \inf_{\pi_1 \in \Pi_1^\gamma} u_1(\sigma_1, \pi_1).$$

Equivalently, $u_1^*(\gamma)$ is the highest payoff player 1 can receive when she chooses a strategy σ_1 and then player 2 chooses a rational strategy σ_2 that minimizes $u_1(\sigma_1, \sigma_2)$; that is,

$$u_1^*(\gamma) = \sup_{\sigma_1} \inf_{\sigma_2 : \sigma_2 \text{ is rational given posture } \gamma} u_1(\sigma_1, \sigma_2).$$

In particular, to guarantee herself a high payoff, player 1 must play a strategy that does well against *any* rational strategy of player 2's.[9]

Finally, I define player 1's maxmin payoff, the highest payoff that player 1 can guarantee herself before announcing a posture, as well as the corresponding maxmin posture.

Definition 3 *Player* 1*'s* maxmin payoff *is*

$$u_1^* \equiv \sup_{\gamma} u_1^* (\gamma) .$$

A posture γ^* *is a* maxmin posture *if there exists a sequence of postures* $\{\gamma_n\}$ *such that* $\gamma_n (t) \to \gamma^* (t)$ *for all* $t \in \mathbb{R}_+$ *and* $u_1^* (\gamma_n) \to u_1^*$*.*

I sometimes emphasize the dependence of u_1^* and γ^* on ε by writing $u_1^* (\varepsilon)$ and γ_ε^*.[10] Both the set of maxmin strategies given any posture γ and the set of maxmin postures are non-empty, though at this point this is not obvious.

The reason why γ^* is defined as a limit of postures $\{\gamma_n\}$ such that $u_1^* (\gamma_n) \to u_1^*$, rather than as an element of $\text{argmax}_\gamma u_1^* (\gamma)$, is that the latter set may be empty because of an openness problem that is standard in bargaining models. To see the problem, consider the ultimatum bargaining game, where player 1 makes a take-it-or-leave-it demand in $[0, 1]$ to player 2. By knowledge of rationality, player 1 knows that any demand strictly below 1 will be accepted, but demanding 1 does not guarantee her a positive payoff because it is a best-response for player 2 to reject. Definition 3 is analogous to specifying that in this game player 1's maxmin payoff is 1 and her maxmin strategy is demanding 1.

Note that Definitions 2 and 3 are "non-Bayesian" in the sense that they characterize the largest payoff that player 1 can guarantee herself, rather than the maximum payoff that she can obtain given some belief. The following is the "Bayesian" version of these definitions:

[9] A potential criticism of the concept of the maxmin payoff given posture γ is that it appears to neglect the fact that, in the event that player 1 does become committed to posture γ, she is guaranteed only $\inf_{\pi_1 \in \Pi_1^\gamma} E^{\pi_1} [u_1 (\gamma, \sigma_2)]$ in the bargaining phase, rather than $\sup_{\sigma_1} \inf_{\pi_1 \in \Pi_1^\gamma} E^{\pi_1} [u_1 (\sigma_1, \sigma_2)]$. However, I show in Section 4.3 that these two numbers are actually identical.

[10] The notation $\gamma^* (\cdot)$ is already taken by the time-t demand of posture γ^*. I apologize for abusing notation in writing $u_1^* (\gamma)$ and $u_1^* (\varepsilon)$ for different objects and hope that this will not cause confusion.

Definition 4 *Player* 1*'s* pessimistic payoff given posture γ *is*

$$u_1^{pess}(\gamma) \equiv \inf_{\pi_1 \in \Pi_1^\gamma} \sup_{\sigma_1} u_1(\sigma_1, \pi_1).$$

Player 1*'s* pessimistic payoff *is* $u_1^{pess} \equiv \sup_\gamma u_1^{pess}(\gamma)$. *A posture* γ^{pess} *is a* pessimistic posture *if there exists a sequence of postures* $\{\gamma_n\}$ *such that* $\gamma_n(t) \to \gamma^{pess}(t)$ *for all* $t \in \mathbb{R}_+$ *and* $u_1^{pess}(\gamma_n) \to u_1^{pess}$.

Player 1's pessimistic payoff is the worst payoff she can receive by best-responding to a fixed rational strategy of player 2's. Player 1's maxmin payoff is weakly lower than her pessimistic payoff, because in the definition of the pessimistic payoff player 1 "knows" the distribution over player 2's strategies when she chooses her strategy, while in the definition of the maxmin payoff her strategy is evaluated with respect to the worst-case response of player 2's. However, I show in Section 4.3 that these payoffs are in fact identical in my model. For expositional consistency, I focus on maxmin payoffs and strategies.

Another reason for studying player 1's maxmin payoff is that it determines the entire range of payoffs that are consistent with knowledge of rationality, as shown by the following proposition.

Proposition 1 *For any posture* γ *and any payoff* $u_1 \in [u_1^*(\gamma), 1)$*, there exists a belief* $\pi_1 \in \Pi_1^\gamma$ *such that* $\max_{\sigma_1} u_1(\sigma_1, \pi_1) = u_1$.

4 Characterization of the Maxmin Payoff and Posture

This section states and proves Theorem 1, the main result of the paper, which solves for player 1's maxmin payoff and posture. Section 4.1 states and discusses Theorem 1, and Sections 4.2 through 4.4 provide the proof. The approach is as follows: In Section 4.2, I fix a posture γ and find a belief of player 2's, π_2^γ (satisfying $\pi_2^\gamma(\gamma) \geq \varepsilon$), and corresponding best-response, $\sigma_2^\gamma \in \Sigma_2^*(\pi_2^\gamma)$, that minimize $u_1(\gamma, \sigma_2)$, player 1's payoff from mimicking γ in the bargaining phase; π_2^γ and σ_2^γ are called the γ*-offsetting belief* and γ*-offsetting strategy*, respectively, and play a key role in the analysis. Section 4.3 shows that γ itself is a maxmin strategy given posture γ, for any γ, which implies that $u_1^*(\gamma) = u_1(\gamma, \sigma_2^\gamma)$ for any posture γ.

That is, player 1's maxmin payoff given posture γ is the payoff she receives from mimicking γ when her opponent follows his γ-offsetting strategy. Section 4.4 maximizes $u_1(\gamma, \sigma_2^\gamma)$ over γ to prove Theorem 1.

4.1 Main Result

The main result is the following:

Theorem 1 *Player* 1*'s maxmin payoff is*

$$u_1^*(\varepsilon) = 1/(1 - \log \varepsilon),$$

and the unique maxmin posture γ_ε^* *is given by*

$$\gamma_\varepsilon^*(t) = \min\left\{e^{rt}/(1 - \log \varepsilon), 1\right\} \text{ for all } t \in \mathbb{R}_+.$$

A priori, one might have expected player 1's maxmin payoff to be very small when ε is small (because player 2's beliefs and strategy may be chosen quite freely in the definition of the maxmin payoff), and might have expected player 1's maxmin posture to be complicated (as player 1 is not restricted to announcing monotone, continuous, or otherwise well-behaved postures). Theorem 1 shows that, on the contrary, player 1's maxmin payoff is relatively high for even very small commitment probabilities ε, as shown by Table 1, and that player 1's unique maxmin posture is simply demanding the maxmin payoff plus compensation for any delay in reaching agreement.[11]

[11] The importance of non-constant postures is a difference between this paper and existing reputational bargaining models, where it is usually assumed that players may only be committed to strategies that demand a constant share of the surplus (as in Abreu and Gul (2000), Compte and Jehiel (2002), and Kambe (1999)). A notable exception is Abreu and Pearce (2007), where players may be committed to non-constant postures that can also condition their play on their opponents' behavior. However, Abreu and Pearce's main result is that a particular posture that demands a constant share of the surplus is approximately optimal in their model, when commitment probabilities are small.

ε	$u_1^*(\varepsilon)$
.25	.42
.1	.30
10^{-3}	.13
10^{-6}	.07
10^{-9}	.05

Table 1: The Maxmin Payoff $u_1^*(\varepsilon)$ for Different Commitment Probabilities ε

The intuition for why the unique maxmin posture is given by $\gamma_\varepsilon^*(t) = \min\{e^{rt}u_1^*(\varepsilon), 1\}$ was outlined in the introduction. The most basic intuition for why $u_1^*(\varepsilon)$ is large relative to ε is that, when player 1 announces a posture that offers player 2 a large share of the surplus and then mimics this posture, player 2 must accept player 1's offer unless he believes that he will be rewarded with high probability for rejecting. In the latter case, if player 1 does not reward player 2 for rejecting, then player 2 quickly updates his belief toward player 1's being committed to her announced posture (i.e., player 1 builds reputation quickly), and player 2 accepts player 1's offer when he becomes convinced that she is committed. Thus, player 1 builds reputation quickly when her demand is small, so a small commitment probability need not lead to much delay before her demand is accepted.

4.2 Offsetting Beliefs and Strategies

In this subsection, I fix an arbitrary posture for player 1, γ, and find a rational strategy of player 2's, σ_2^γ (the "γ-offsetting strategy"), that minimizes player 1's payoff when she announces posture γ and then mimics γ in the bargaining phase. That is, I solve the problem

$$\inf_{(\pi_2,\sigma_2):\pi_2(\gamma)\geq\varepsilon,\sigma_2\in\Sigma_2^*(\pi_2)} u_1(\gamma,\sigma_2). \tag{1}$$

The resulting strategy σ_2^γ always demands the entire surplus and rejects player 1's demand until some time t^*, and accepts player 1's smaller time-t^* demand, $\min\{\lim_{\tau\uparrow t^*}\gamma(\tau), \gamma(t^*)\}$ (henceforth denoted by $\underline{\gamma}(t^*)$), if player 1 follows γ until time t^*. If player 1 ever deviates from γ, then σ_2^γ rejects player 1's demand forever. The corresponding belief π_2^γ (the "γ-offsetting belief") is that player 1 plays γ with probability ε, and with probability $1-\varepsilon$ plays a particular strategy $\tilde{\gamma}$ that mixes between mimicking γ and accepting player 2's demand up until time t^*, and subsequently mimics γ.

The key step in solving (1) is computing the smallest time T by which agreement must be reached under strategy profile $(\gamma, \sigma_2^\gamma)$. I then show that the value of (1) is simply $\min_{t\leq T} e^{-rt}\underline{\gamma}(t)$, and that the time t^* at which the strategy σ_2^γ accepts $\underline{\gamma}(t)$ is a time before T that minimizes $e^{-rt}\underline{\gamma}(t)$.

Toward computing T, let $v(t,-1)$ be the continuation value of player 2 from best-responding to γ starting from date $(t,-1)$, and let $v(t)$ be the corresponding continuation value starting from date $(t,1)$:

$$\begin{aligned} v(t,-1) &\equiv \max_{\tau\geq t} e^{-r(\tau-t)}\left(1-\underline{\gamma}(\tau)\right), \\ v(t) &\equiv \max\left\{1-\gamma(t), \sup_{\tau>t} e^{-r(\tau-t)}\left(1-\underline{\gamma}(\tau)\right)\right\}, \end{aligned} \tag{2}$$

where $\underline{\gamma}(\tau) \equiv \min\{\lim_{s\uparrow\tau}\gamma(s), \gamma(\tau)\}$. Thus, the difference between $v(t,-1)$ and $v(t)$ is that only $v(t,-1)$ gives player 2 the opportunity to accept the demand $1-\lim_{\tau\uparrow t}\gamma(\tau)$; in particular, $v(t,-1)=v(t)$ if $\gamma(t)$ (or $v(t)$) is continuous at t. Note that $\max_{\tau\geq t} e^{-r(\tau-t)}\left(1-\underline{\gamma}(\tau)\right)$ is well-defined because $\underline{\gamma}(\tau)$ is lower semi-continuous and $\lim_{\tau\to\infty} e^{-r(\tau-t)}\left(1-\underline{\gamma}(\tau)\right)=0$, and that $v(t)$ is continuous at all non-integer times t; let $\{s_1, s_2, \ldots\} \equiv S \subseteq \mathbb{N}$ be the set of discontinuity points of $v(t)$. Finally, note that $v(t)$ can increase at rate no faster than r. That is, $v(t) \geq e^{-r(t'-t)}v(t')$ for all $t' \geq t$, because if $v(t') = e^{-r(\tau-t')}\left(1-\underline{\gamma}(\tau)\right)$ for some $\tau \geq t'$, then $v(t) \geq e^{-r(\tau-t)}\left(1-\underline{\gamma}(\tau)\right) = e^{-r(t'-t)}v(t')$. This implies that $v(t)$ is continuous but for downward jumps,[12] and that $v(t)$ is differentiable almost everywhere.[13] These are but two of the useful properties of the function $v(t)$ (which are not shared by $\gamma(t)$) that reward working with $v(t)$ rather than $\gamma(t)$ in the subsequent analysis.

Next, I introduce two functions $\lambda(t)$ and $p(t)$ with the property that if player 1 mixes between mimicking γ and conceding the entire surplus to player 2, then $\lambda(t)$ (resp., $p(t)$)

[12] A function $f : \mathbb{R}\to\mathbb{R}$ is *continuous but for downward jumps* if $\liminf f_{x\uparrow x^*}(x) \geq f(x^*) \geq \limsup_{x\downarrow x^*} f(x)$ for all $x\in\mathbb{R}$.

[13] Proof: Let $f(t) = e^{-rt}v(t)$. Then $f(t)$ is non-increasing, which implies that $f(t)$ is differentiable almost everywhere (e.g., Royden, 1988, p. 100). Hence, $v(t)$ is differentiable almost everywhere.

is the smallest non-negative hazard rate (resp., discrete probability) at which player 1 must concede in order for player 2 to be willing to reject player 1's time-t demand, $\underline{\gamma}(t)$. Let

$$\lambda(t) = \frac{rv(t) - v'(t)}{1 - v(t)} \tag{3}$$

if $v(t)$ is differentiable at t and $v(t) < 1$, and let $\lambda(t) = 0$ otherwise; note that $\lambda(t) \geq 0$ for all t, because $v(t)$ cannot increase at rate faster than r. Also, let

$$p(t) = \frac{v(t,-1) - v(t)}{1 - v(t)} \tag{4}$$

if $v(t) < 1$, and let $p(t) = 0$ otherwise. To see intuitively why the aforementioned property holds, note that accepting player 1's time-t demand gives player 2 flow payoff $rv(t)$, while rejecting gives player 2 flow payoff $\lambda(t)(1 - v(t)) + v'(t)$, and equalizing these flow payoffs yields (3);[14,15] similarly, accepting player 1's demand at date $(t,-1)$ gives player 2 payoff $v(t,-1)$, while delaying acceptance until date $(t,1)$ gives player 2 payoff $p(t)(1) + (1 - p(t))v(t)$, and equalizing these payoffs yields (4).

When player 2 expects player 1 to accept his demand at rate (resp., probability) $\lambda(t)$ (resp., $p(t)$), he becomes convinced that player 1 is committed to posture γ at the time $\tilde{T}$ defined in the following lemma, which leads him to accept player 1's demand no later than the time T defined in the lemma. In the lemma, and throughout the paper, maximization or minimization over times t should be read as taking place over $t \in \mathbb{R}_+ \cup \{\infty\}$ (i.e., as allowing $t = \infty$, with the convention that $e^{-r\infty}\underline{\gamma}(\infty) \equiv 0$ for all postures γ).

Lemma 1 *Let*

$$\tilde{T} \equiv \sup \left\{ t : \exp\left(-\int_0^t \lambda(s)\, ds \right) \prod_{s \in S \cap [0,t)} (1 - p(s)) > \varepsilon \right\},$$

[14] This intuition is correct when $v(t) = 1 - \underline{\gamma}(t)$. When $v(t) > 1 - \underline{\gamma}(t)$, player 2 prefers to reject player 1's time-t demand even when player 1 concedes at rate 0. At these times, $rv(t) = v'(t)$, which implies that $\lambda(t) = 0$. Hence, $\lambda(t)$ is always the smallest non-negative hazard rate at which player 1 must concede in order for player 2 to be willing to reject $\underline{\gamma}(t)$.

[15] If $v'(t) = 0$, then $\lambda(t)$ becomes the concession rate that makes player 2 indifferent between accepting and rejecting the constant offer $v(t)$, which is familiar from the literatures on wars of attrition and reputational bargaining. However, in these literatures $\lambda(t)$ is the rate at which player 1 concedes in equilibrium, while here is the rate at which player 1 concedes according to player 2's offsetting beliefs, as will become clear.

and let

$$T \equiv \max \operatorname*{argmax}_{t \geq \tilde{T}} \begin{cases} e^{-rt}\left(1-\gamma\left(t\right)\right) & \text{if } t = \tilde{T} \\ e^{-rt}\left(1-\underline{\gamma}\left(t\right)\right) & \text{if } t > \tilde{T} \end{cases}.$$

Then, for any π_2 such that $\pi_2(\gamma) \geq \varepsilon$ and any $\sigma_2 \in \Sigma_2^*(\pi_2)$, agreement is reached no later than time T under strategy profile (γ, σ_2). In particular,

$$\inf_{(\pi_2,\sigma_2):\pi_2(\gamma)\geq\varepsilon,\sigma_2\in\Sigma_2^*(\pi_2)} u_1(\gamma,\sigma_2) \geq \min_{t\leq T} e^{-rt}\underline{\gamma}(t). \tag{5}$$

Lemma 1 amounts to the statement that $\tilde{T}$ is the latest time at which it is possible that agreement has not yet been reached *and* player 2 is not certain that player 1 is playing γ, when player 2's initial belief is some π_2 such that $\pi_2(\gamma) \geq \varepsilon$, under strategy profile (γ, σ_2) for some $\sigma_2 \in \Sigma_2^*(\pi_2)$. The first step of the proof (of Lemma 1) shows that, in computing this time, one can restrict attention to beliefs π_2 that assign probability 1 to player 1's accepting player 2's demand whenever player 1 deviates from strategy γ, and to strategies σ_2 that always demand the entire surplus. This is because giving more surplus to player 2 *in the event that player* 1 *deviates from* γ makes player 2 more willing to reject player 1's demand, without changing player 2's beliefs about the probability that player 1 is playing γ at any history. The second step shows that, with beliefs of this form, if $v(t)$ is always equal to player 2's continuation payoff from delaying acceptance until he becomes convinced that player 1 is playing γ, then player 1's concession rate and probability must be given by $\lambda(t)$ and $p(t)$, and player 2 becomes convinced that player 1 is playing γ at time $\tilde{T}$; this formalizes the motivation for $\lambda(t)$ and $p(t)$ given above. The proof is completed by showing that if $v(t)$ is ever strictly less than player 2's continuation payoff from delaying acceptance until he becomes convinced that player 1 is playing γ, then player 2 becomes convinced that player 1 is playing γ no later than $\tilde{T}$.

The remainder of this subsection is devoted to showing that (5) holds with equality, which proves that (1) equals $\min_{t\leq T} e^{-rt}\underline{\gamma}(t)$. The idea is that player 2 may hold a belief that induces him to demand the entire surplus until time $t^* \equiv \min \operatorname{argmin}_{t\leq T} e^{-rt}\underline{\gamma}(t)$ and then accept player 1's offer; this is the γ-offsetting belief.[16] I first define the γ-offsetting

[16]Note that $\min \operatorname{argmin}_{t\leq T} e^{-rt}\underline{\gamma}(t)$ is well-defined, because $\underline{\gamma}(t)$ is lower semi-continuous (though it may equal ∞, if $T=\infty$). This particular choice of t^* is for concreteness; any element of $\operatorname{argmin}_{t\leq T} e^{-rt}\underline{\gamma}(t)$ would suffice for the analysis.

belief, and then show that (5) holds with equality.

I begin by introducing a strategy, $\tilde{\gamma}$, which is used in defining the γ-offsetting belief.[17] Let

$$\chi(t) = \max\left\{ \frac{\exp\left(-\int_0^t \lambda(s)\,ds\right)\prod_{s\in S\cap[0,t)}(1-p(s)) - \varepsilon}{\exp\left(-\int_0^t \lambda(s)\,ds\right)\prod_{s\in S\cap[0,t)}(1-p(s))}, 0 \right\}; \tag{6}$$

let

$$\hat{\lambda}(t) = \frac{\lambda(t)}{\chi(t)} \tag{7}$$

if $\chi(t) > 0$, and let $\hat{\lambda}(t) = 0$ otherwise; and let

$$\hat{p}(t) = \min\left\{ \frac{p(t)}{\chi(t)}, 1 \right\} \tag{8}$$

if $\chi(t) > 0$, and let $\hat{p}(t) = 0$ otherwise. Intuitively, $\chi(t)$ is the posterior probability that player 2 assigns to player 1's playing a strategy other than γ at time t when player 1's *unconditional* concession rate and probability are $\lambda(t)$ and $p(t)$, and $\hat{\lambda}(t)$ and $\hat{p}(t)$ are the *conditional* (on not playing γ) concession rate and probability needed for the unconditional concession rate and probability to equal $\lambda(t)$ and $p(t)$.

Definition 5 *$\tilde{\gamma}$ is the strategy that demands $u_1(t) = \gamma(t)$ for all $t \in \mathbb{R}_+$, accepts with hazard rate $\hat{\lambda}(t)$ for all $t < t^*$, accepts with probability $\hat{p}(t)$ at date $(t, 1)$ for all $t \leq t^*$, and rejects for all $t > t^*$, for all histories h^t.*

I now define the γ-offsetting belief.

Definition 6 *The* γ-offsetting belief, *denoted π_2^γ, is given by $\pi_2^\gamma(\gamma) = \varepsilon$ and $\pi_2^\gamma(\tilde{\gamma}) = 1 - \varepsilon$. The* γ-offsetting strategy, *denoted σ_2^γ, is the strategy that demands $u_2(t) = 1$ for all t and accepts or rejects player 1's demand as follows:*

1. *If h^t is consistent with γ, then reject if $t < t^*$; accept at date $(t^*, -1)$ if and only if $\lim_{\tau \uparrow t^*} \gamma(\tau) \leq \gamma(t^*)$; accept at date $(t^*, 1)$ if and only if $\lim_{\tau \uparrow t^*} \gamma(\tau) > \gamma(t^*)$; and reject if $t > t^*$.*[18]

[17] This approach is related to a construction in Wolitzky (2011).

[18] History h^t is *consistent* with strategy σ_1 if there exists a strategy σ_2 such that h^t is reached under (σ_1, σ_2). In particular, history h^{t-} (resp., h^{t+}) is consistent with γ if and only if $u_1(\tau) = \gamma(\tau)$ for all $\tau < t$ (resp., $\tau \leq t$).

2. If h^t is not consistent with γ, then reject.

Finally, I show that (5) holds with equality, and also that the γ-offsetting *(belief, strategy)* pair $(\pi_2^\gamma, \sigma_2^\gamma)$ is a solution to (1). If $t^* = \infty$, then the following statement that agreement is reached at time t^* means that agreement is never reached.

Lemma 2 *Agreement is reached at time t^* under strategy profile $(\gamma, \sigma_2^\gamma)$, and $\sigma_2^\gamma \in \Sigma_2^* (\pi_2^\gamma)$. In particular, the pair $(\pi_2^\gamma, \sigma_2^\gamma)$ is a solution to (1), and $u_1 (\gamma, \sigma_2^\gamma) = \min_{t \leq T} e^{-rt} \underline{\gamma} (t)$.*

Proof. It is immediate from Definition 6 that agreement is reached at t^* under strategy profile $(\gamma, \sigma_2^\gamma)$, which implies that $u_1 (\gamma, \sigma_2^\gamma)$ equals $\min_{t \leq T} e^{-rt} \underline{\gamma} (t)$, the right-hand side of (5). Since $\pi_2^\gamma (\gamma) \geq \varepsilon$, it remains only to show that $\sigma_2^\gamma \in \Sigma_2^* (\pi_2^\gamma)$.

If $t < \min \left\{ \tilde{T}, t^* \right\}$ and h^t is consistent with γ, then, by construction of $\tilde{\gamma}$, player 1 accepts player 2's demand of 1 with unconditional hazard rate $\lambda (t)$ and unconditional discrete probability $p (t)$ under π_2^γ. The proof of Lemma 1 implies that it is optimal for player 2 to demand $u_2 (t) = 1$ and reject at any time $t < \min \left\{ \tilde{T}, t^* \right\}$ when player 1 accepts player 2's demand of 1 at rate $\lambda (t)$ and probability $p (t)$ until time $\tilde{T}$; and that in addition if $t^* < \tilde{T}$ then player 2 is indifferent between between accepting and rejecting at time t^* when player 1 accepts with this rate and probability until time $\tilde{T}$. Therefore, it is optimal for player 2 to demand $u_2 (t) = 1$ and reject at time t when player 1 accepts with this rate and probability only until time $\min \left\{ \tilde{T}, t^* \right\}$.

If $t \in \left[\tilde{T}, t^* \right)$ and h^t is consistent with γ, then under π_2^γ player 2 is certain that player 1 is playing γ at h^t. Since $t^* \leq T$, this implies that it is optimal for player 2 to reject. If a history h^t is not reached under strategy profile $(\pi_2^\gamma, \sigma_2^\gamma)$ (as is the case if $t > t^*$), then any continuation strategy of player 2's is optimal. Finally, to see that accepting $\underline{\gamma} (t^*)$ (i.e., accepting at the more favorable of dates $(t^*, -1)$ and $(t^*, 1)$) is optimal, note that the fact that $t^* \in \operatorname{argmin}_{t \leq T} e^{-rt} \underline{\gamma} (t)$ implies that $\underline{\gamma} (t) \geq \underline{\gamma} (t^*)$ for all $t \in [t^*, T]$. Hence, $t^* \in \operatorname{argmax}_{t \in [t^*, T]} e^{-rt} \left(1 - \underline{\gamma} (t)\right)$. Because $\tilde{\gamma}^{t^*}$ coincides with γ after time t^*, it follows that, conditional on having reached time t^*, player 2 receives at most $\sup_{t \in (t^*, T]} e^{-rt} \left(1 - \underline{\gamma} (t)\right)$ if he rejects, and receives $e^{-rt^*} \left(1 - \underline{\gamma} (t^*)\right)$ if he accepts, which is weakly more. Therefore, $\sigma_2^\gamma \in \Sigma_2^* (\pi_2^\gamma)$. ■

4.3 Maxmin Strategies

This subsection shows that γ itself is a maxmin strategy given posture γ, and that in particular $u_1^*(\gamma) = u_1(\gamma, \sigma_2^\gamma) = \min_{t \leq T(\gamma)} e^{-rt}\underline{\gamma}(t)$, where I have made the dependence of T on γ explicit. The intuition is that player 1 is not guaranteed a positive payoff in any continuation game following a deviation from her announced posture, because at such histories player 2's beliefs and strategy are unrestricted.

The key result of this subsection is the following:

Lemma 3 *For any posture γ, $u_1^*(\gamma) = \min_{t \leq T(\gamma)} e^{-rt}\underline{\gamma}(t)$.*

Proof. By Lemma 2, $(\pi_2^\gamma, \sigma_2^\gamma)$ is a solution to (1), so

$$\sigma_2^\gamma \in \underset{\pi_1 \in \Pi_1^\gamma}{\operatorname{argmin}}\, u_1(\gamma, \pi_1). \tag{9}$$

Under strategy σ_2^γ, player 2 always demands $u_2(t) = 1$ and only accepts player 1's demand if player 1 conforms to γ through time t^*. Hence, $\sup_{\sigma_1} u_1(\sigma_1, \sigma_2^\gamma) = e^{-rt^*}\underline{\gamma}(t^*) = u_1(\gamma, \sigma_2^\gamma)$, and therefore

$$\gamma \in \underset{\sigma_1}{\operatorname{argmax}}\, u_1(\sigma_1, \sigma_2^\gamma). \tag{10}$$

(9) and (10) imply the following chain of inequalities:

$$\begin{aligned}
\sup_{\sigma_1} \inf_{\pi_1 \in \Pi_1^\gamma} u_1(\sigma_1, \pi_1) &\geq \inf_{\pi_1 \in \Pi_1^\gamma} u_1(\gamma, \pi_1) \\
&= u_1(\gamma, \sigma_2^\gamma) \quad \text{(by (9))} \\
&= \max_{\sigma_1} u_1(\sigma_1, \sigma_2^\gamma) \quad \text{(by (10))} \\
&\geq \max_{\sigma_1} \min_{\pi_1 \in \Pi_1^\gamma} u_1(\sigma_1, \pi_1).
\end{aligned}$$

This is possible only if both inequalities hold with equality (and the supremum and infimum in the first line are attained at γ and σ_2^γ, respectively). Therefore, $u_1^*(\gamma) = u_1(\gamma, \sigma_2^\gamma) = \min_{t \leq T(\gamma)} e^{-rt}\underline{\gamma}(t)$. ■

As an aside, a similar argument establishes the equivalence between the maxmin approach of Definitions 2 and 3 and the Bayesian approach of Definition 4.

Corollary 1 $u_1^{pess}(\gamma) = u_1^*(\gamma)$ *for all γ; $u_1^{pess} = u_1^*$; and γ is a maxmin posture if and only if it is a pessimistic posture.*

Proof. (9) and (10) imply that $\inf_{\pi_1 \in \Pi_1^\gamma} \sup_{\sigma_1} u_1(\sigma_1, \pi_1) = u_1(\gamma, \sigma_2^\gamma)$, by the same chain of inequalities that proves that $\sup_{\sigma_1} \inf_{\pi_1 \in \Pi_1^\gamma} u_1(\sigma_1, \pi_1) = u_1(\gamma, \sigma_2^\gamma)$. Hence, $u_1^{pess}(\gamma) = u_1(\gamma, \sigma_2^\gamma) = u_1^*(\gamma)$ for all γ, and the remainder of the result follows from Definition 4. ■

4.4 Proof of Theorem 1

I now sketch the remainder of the proof of Theorem 1. The details of the proof are deferred to the appendix.

The first part of the proof is constructing a sequence of postures $\{\gamma_n\}$ such that $\lim_{n\to\infty} u_1^*(\gamma_n) = 1/(1-\log\varepsilon)$ and $\{\gamma_n(t)\}$ converges pointwise to $\gamma^*(t) \equiv \min\{e^{rt}/(1-\log\varepsilon), 1\}$. Define γ_n by

$$\gamma_n(t) = \min\left\{\left(\frac{n}{n+1}\right)\frac{e^{rt}}{1-\log\varepsilon}, 1\right\} \text{ for all } t \in \mathbb{R}_+.$$

Let T_n^1 be the time where $\gamma_n(t)$ reaches 1. It can be shown that $T_n^1 > \tilde{T}(\gamma_n)$ for all $n \in \mathbb{N}$, where $\tilde{T}$ is defined as in Lemma 1 and I have emphasized the dependence of $\tilde{T}$ on γ. This implies that $\gamma_n(t) = \left(\frac{n}{n+1}\right)\frac{e^{rt}}{1-\log\varepsilon}$ for all $t \le \tilde{T}(\gamma_n)$, and that $\gamma_n\left(\tilde{T}(\gamma_n)\right) < 1$. Since $\gamma_n(t)$ is non-decreasing and $\gamma_n\left(\tilde{T}(\gamma_n)\right) < 1$, it follows from the definition of $T(\gamma_n)$ that $T(\gamma_n) = \tilde{T}(\gamma_n)$. Thus, by Lemma 3,

$$\begin{aligned} u_1^*(\gamma_n) &= \min_{t \le T(\gamma_n)} e^{-rt}\underline{\gamma_n}(t) \\ &= \min_{t \le \tilde{T}(\gamma_n)} \left(\frac{n}{n+1}\right)\frac{1}{1-\log\varepsilon} \\ &= \left(\frac{n}{n+1}\right)\frac{1}{1-\log\varepsilon}. \end{aligned}$$

Therefore, $\lim_{n\to\infty} u_1^*(\gamma_n) = 1/(1-\log\varepsilon)$.

The second part is showing that no posture γ guarantees more than $1/(1-\log\varepsilon)$. Here, the crucial observation is that any posture γ such that $\gamma(t) \ge e^{rt}/(1-\log\varepsilon)$ for all $t \le T(\gamma)$ satisfies $\tilde{T}(\gamma) \ge T^1$, where T^1 is the time at which $\gamma^*(t)$ reaches 1. Since any posture that guarantees at least $1/(1-\log\varepsilon)$ must satisfy $\gamma(t) \ge e^{rt}/(1-\log\varepsilon)$ for all $t \le T(\gamma)$ (by Lemma 3), and $T(\gamma) \ge \tilde{T}(\gamma)$ for any posture γ, this implies that $u_1^*(\gamma) = \min_{t \le T(\gamma)} e^{-rt}\underline{\gamma}(t) \le e^{-rT^1} = 1/(1-\log\varepsilon)$ for any posture γ. The appendix shows that in addition no sequence of postures $\{\gamma_n'\}$ converging pointwise to any posture other than γ^* can correspond to a sequence of maxmin payoffs $\{u_1^*(\gamma_n')\}$ converging to $1/(1-\log\varepsilon)$.

5 Extensions

This section presents four extensions of Theorem 1. Section 5.1 characterizes the maxmin payoff when player 1 can only announce constant postures; Section 5.2 extends Theorem 1 to general convex bargaining sets; Section 5.3 considers heterogeneous discounting; and Section 5.4 extends Theorem 1 to higher-order knowledge of rationality.

5.1 Constant Postures

Theorem 1 shows that the unique maxmin posture is non-constant. In this subsection, I determine how much lower a player's maxmin payoff is when she is required to announce a constant posture. The purpose of this study is, first, to establish that announcing non-constant postures allows a player to guarantee herself a significantly higher payoff; second, to determine the share of the surplus that a player can guarantee herself in settings where announcing a non-constant posture might not be credible; and, third, to facilitate comparison with the existing reputational bargaining literature, in which typically players can only announce constant postures.

A posture γ is *constant* if $\gamma(t) = \gamma(0)$ for all t. If γ is constant, I slightly abuse notation by writing γ for the constant demand $\gamma(t)$ in addition to the posture itself. The constant posture γ that maximizes $u_1^*(\gamma)$ is the *maxmin constant posture*, denoted $\bar{\gamma}^*$,[19] and the corresponding payoff is the *maxmin constant payoff*, denoted $\bar{u}_1^*$. These can be derived using Lemmas 1 through 3, leading to the following:

Proposition 2 *For all $\varepsilon < 1$, the unique maxmin constant posture is $\bar{\gamma}_\varepsilon^* = \frac{2-\log\varepsilon-\sqrt{(\log\varepsilon)^2-4\log\varepsilon}}{2}$, and the maxmin constant payoff is $\bar{u}_1^*(\varepsilon) = \exp(-(1-\bar{\gamma}_\varepsilon^*))\,\bar{\gamma}_\varepsilon^*$.*

Proposition 2 solves for $\bar{\gamma}_\varepsilon^*$ and $\bar{u}_1^*(\varepsilon)$, but it does not yield a clear relationship between the maxmin constant payoff, $\bar{u}_1^*(\varepsilon)$, and the (overall) maxmin payoff, $u_1^*(\varepsilon)$. Therefore, I graph the ratio of $u_1^*(\varepsilon)$ to $\bar{u}_1^*(\varepsilon)$ in Figure 2. In addition, the following analytical result regarding the ratio of $u_1^*(\varepsilon)$ to $\bar{u}_1^*(\varepsilon)$ is straightforward:

[19]Such a posture exists for all $\varepsilon < 1$, so there is no need for a limit definition like Definition 3. When $\varepsilon = 1$, such a definition would imply that the maxmin constant posture equals 1.

Corollary 2 *$u_1^*(\varepsilon)/\bar{u}_1^*(\varepsilon)$ is decreasing in ε, $\lim_{\varepsilon\to 1} u_1^*(\varepsilon)/\bar{u}_1^*(\varepsilon) = 1$, and $\lim_{\varepsilon\to 0} u_1^*(\varepsilon)/\bar{u}_1^*(\varepsilon) = e$.*

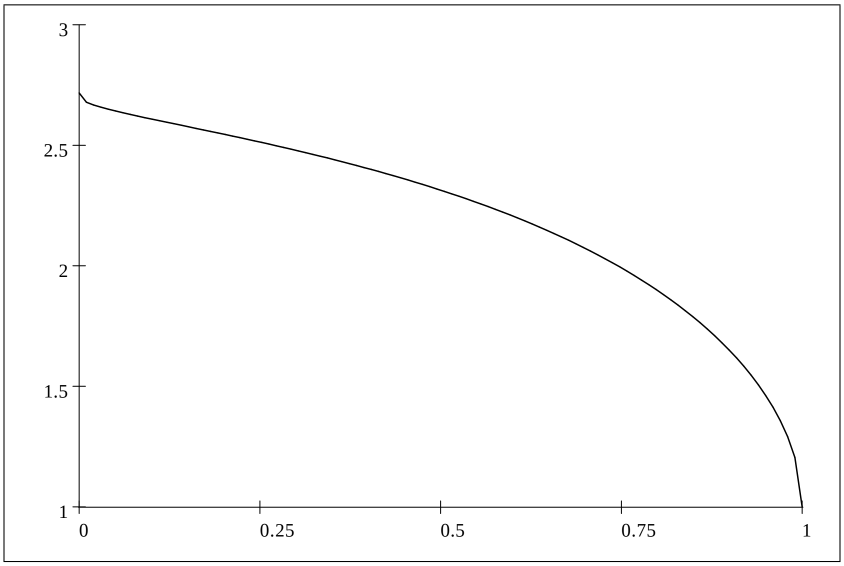

Figure 2: The Ratio $u_1^*(\varepsilon)/\bar{u}_1^*(\varepsilon)$ for $\varepsilon \in [0,1]$.

The most interesting part of Corollary 2 is that a player's maxmin payoff is approximately e times greater when she can announce non-constant postures than when she can only announce constant postures, when her commitment probability is small. Thus, there is a large advantage to announcing non-constant postures. However, a player can still guarantee herself a substantial share of the surplus when she can only announce constant postures, and her maxmin payoff goes to 0 with ε at the same rate in either case.

5.2 General Convex Bargaining Sets

This subsection shows that the maxmin payoff derived in Theorem 1 is a lower bound on the maxmin payoff with general convex bargaining sets, normalized so that each player's lowest and highest feasible payoffs are 0 and 1. More generally, taking a concave transformation of the Pareto frontier of the bargaining set weakly increases the maxmin payoff.

Formally, a decreasing function $\phi : [0,1] \to [0,1]$ is the *Pareto frontier (of the bargaining set)* if the game ends with payoffs $(e^{-rt}u_1(t), e^{-rt}\phi(u_1(t)))$ when player 2 accepts player 1's demand $u_1(t)$, and ends with payoffs $\left(e^{-rt}\phi^{-1}(u_2(t)), e^{-rt}u_2(t)\right)$ when player 1 accepts player 2's demand $u_2(t)$. Note that the definition of player 1's maxmin payoff is valid for any bargaining set. The result is the following:

Proposition 3 *Suppose that $\phi : [0,1] \to [0,1]$ is a decreasing and concave function satisfying $\phi(0) = 1$ and $\phi(1) = 0$, and that $\psi : [0,1] \to [0,1]$ is an increasing and concave function satisfying $\psi(0) = 0$ and $\psi(1) = 1$. Let u_1^{ϕ} be player 1's maxmin payoff when the Pareto frontier is ϕ, and let $u_1^{\psi \circ \phi}$ be player 1's maxmin payoff when the Pareto frontier is $\psi \circ \phi$. Then $u_1^{\psi \circ \phi} \geq u_1^{\phi}$.*

Proposition 3 shows that taking any concave transformation ψ of a Pareto frontier ϕ weakly increases player 1's maxmin payoff. The intuition is that any fixed demand of player 1's leaves more for player 2 when the Pareto frontier is $\psi \circ \phi$ than when it is ϕ, which implies that player 2 must believe that player 1 is conceding more rapidly in order for him to reject player 1's demand when the Pareto frontier is $\psi \circ \phi$. This in turn lets player 1 build reputation more quickly and thus guarantee herself a higher payoff.

5.3 Heterogeneous Discounting

I have assumed that the players have the same discount rate. This simplified notation and led to simple formulas for $u_1^*(\varepsilon)$ and γ_ε^* in Theorem 1. However, it is straightforward to generalize the model to the case where player i has discount rate r_i and $r_i \neq r_j$; one must only keep track of whose discount rate "r" stands for in the above analysis. Introducing heterogeneous discounting yields interesting comparative statics with respect to the players' relative patience, r_1/r_2 (as will become clear, u_1^* depends on r_1 and r_2 only through r_1/r_2). First, the standard result in the reputational bargaining literature (Abreu and Gul, 2000; Compte and Jehiel, 2002; Kambe, 1999) that player 1's sequential equilibrium payoff converges to 1 as r_1/r_2 converges to 0, and converges to 0 as r_1/r_2 converges to ∞, also applies to player 1's maxmin payoff. Thus, this important comparative static result continues to hold under knowledge of rationality, and in particular does not rely on equilibrium. This is analogous to the results on reputation in repeated games under knowledge of rationality of Watson (1993) and Battigalli and Watson (1997). However, I also derive player 1's maxmin payoff for fixed r_1/r_2 (rather than only in the limit). This leads to a second comparative static result, which indicates that a change in relative patience has a larger effect on the maxmin payoff than a much larger change in commitment probability. An analogous result holds in equilibrium in existing reputational bargaining models.

I first present the analog of Theorem 1 for heterogeneous discount rates, and then state the two comparative statics results as corollaries.

Proposition 4 *If player i's discount rate is r_i, then player 1's maxmin payoff, $u_1^*(\varepsilon)$, is the unique number u_1^* that solves*

$$u_1^* = \frac{1}{1 - \frac{r_1}{r_2}\log\varepsilon - \left(\frac{r_1}{r_2} - 1\right)\log u_1^*}. \tag{11}$$

Corollary 3 shows that the standard limit comparative statics on r_1/r_2 in reputational bargaining models require only first-order knowledge of rationality.

Corollary 3 $\lim_{r_1/r_2 \to 0} u_1^*(\varepsilon) = 1$. *If $\varepsilon < 1$, then in addition* $\lim_{r_1/r_2 \to \infty} u_1^*(\varepsilon) = 0$.

Corollary 4 shows that the commitment probability ε must decrease exponentially to (approximately) offset a geometric increase in relative patience $(r_1/r_2)^{-1}$. The result is stated for the case $r_1/r_2 \le 1$, where even an exponential decrease in ε does not fully offset a geometric increase in $(r_1/r_2)^{-1}$. If $r_1/r_2 > 1$, then an exponential decrease in ε more than offsets a geometric increase in $(r_1/r_2)^{-1}$.

Corollary 4 *Suppose that $r_1/r_2 \le 1$ and that r_1/r_2 and ε both decrease while $(r_1/r_2)\log\varepsilon$ remains constant. Then $u_1^*(\varepsilon)$ increases.*

5.4 Rationalizability

Theorem 1 derives the highest payoff that player 1 can guarantee herself under first-order knowledge of rationality, the weakest epistemic assumption consistent with the possibility of reputation-building. I now show that player 1 cannot guarantee herself more than this under the much stronger assumption of rationalizability (or under any finite-order knowledge of rationality), which reenforces Theorem 1 substantially. The intuition is that the γ-offsetting belief—and thus the γ-offsetting strategy—is not only rational but also rationalizable, and player 1 receives payoff u_1^* when she best-responds to the γ-offsetting strategy.

I consider the following definition of rationalizability:[20]

Definition 7 *A set of bargaining phase strategy profiles* $\Omega = \Omega_1 \times \Omega_2 \subseteq \Sigma_1 \times \Sigma_2$ *has the* best-response property given posture γ *if for all* $\sigma_1 \in \Omega_1$ *there exists some belief* $\pi_1 \in \Delta(\Omega_2)$ *such that* $\sigma_1 \in \Sigma_1^*(\pi_1)$*; and for all* $\sigma_2 \in \Omega_2$ *there exists some belief* $\pi_2 \in \Delta(\Omega_1 \cup \{\gamma\})$ *such that* $\pi_2(\gamma) \geq \varepsilon$*, with strict inequality only if* $\gamma \in \Omega_1$*, and* $\sigma_2 \in \Sigma_2^*(\pi_2)$*. The* set of rationalizable strategies given posture γ *is*

$$\Omega^{RAT}(\gamma) \equiv \bigcup \{\Omega : \Omega \textit{ has the best-response property given posture } \gamma\}.$$

Player 1*'s* rationalizable maxmin payoff given posture γ *is*

$$u_1^{RAT}(\gamma) \equiv \sup_{\sigma_1} \inf_{\sigma_2 \in \Omega_2^{RAT}(\gamma)} u_1(\sigma_1, \sigma_2).$$

Player 1*'s* rationalizable maxmin payoff *is*

$$u_1^{RAT} \equiv \sup_{\gamma} u_1^{RAT}(\gamma).$$

A posture γ^{RAT} *is a* rationalizable maxmin posture *if there exists a sequence of postures* $\{\gamma_n\}$ *such that* $\gamma_n(t) \to \gamma^{RAT}(t)$ *for all* $t \in \mathbb{R}_+$ *and* $u_1^{RAT}(\gamma_n) \to u_1^{RAT}$*.*

The result is the following:

Proposition 5 *Player* 1*'s rationalizable maxmin payoff equals her maxmin payoff, and the unique rationalizable maxmin posture is the unique maxmin posture. That is,* $u_1^{RAT} = u_1^*$*, and the unique rationalizable maxmin posture is* $\gamma^{RAT} = \gamma^*$*.*

Any rationalizable strategy given posture γ is also rational given posture γ. Therefore, Lemma 1 applies under rationalizability. The only additional fact used in the proof of Theorem 1 is that $u_1^*(\gamma) = \min_{t \leq T(\gamma)} e^{-rt}\underline{\gamma}(t)$ for any posture γ (Lemma 3). Supposing that the analogous equation holds under rationalizability (i.e., that $u_1^{RAT}(\gamma) = \min_{t \leq T(\gamma)} e^{-rt}\underline{\gamma}(t)$),

[20] While consistent with this paper's focus on normal-form rationality, this normal-form definition of rationalizability is weak in that it does not eliminate strategies that are dominated "off-path." However, I conjecture that Proposition 5 also holds under extensive-form rationalizability (Pearce, 1984; Battigalli and Siniscalchi, 2003).

the proof of Theorem 1 goes through as written. Hence, to prove Proposition 5 it suffices to prove the following lemma, the proof of which shows that the γ-offsetting belief and strategy are rationalizable:

Lemma 4 *For any posture γ, $u_1^{RAT}(\gamma) = \min_{t \leq T(\gamma)} e^{-rt} \underline{\gamma}(t)$.*

6 Two-Sided Commitment

This section introduces the possibility that both players may announce—and become committed to—postures prior to the start of bargaining. I show that each player i's maxmin payoff is close to that derived in Section 4 when her opponent's commitment probability, ε_j, is small in absolute terms (even if ε_j is large relative to ε_i). In addition, each player's maxmin posture is exactly as in Section 4. This shows that the analysis of Section 4 provides a two-sided theory of reputational bargaining. The results of this section contrast with the existing reputational bargaining literature, which emphasizes that *relative* commitment probabilities are crucial for determining *equilibrium* behavior and payoffs.

Formally, modify the model of Section 3 by assuming that in the announcement stage players simultaneously announce postures (γ_1, γ_2), to which they become committed with probabilities ε_1 and ε_2, respectively.[21] The bargaining phase is unaltered. Thus, at the beginning of the bargaining phase, player i believes that player j is committed to posture γ_j with probability ε_j and is rational with probability $1 - \varepsilon_j$ (though this fact is not common knowledge). The following definitions are analogs of Definitions 1 through 3 that allow for the fact that both players may become committed to the postures they announce:

Definition 8 *A belief π_i of player i's is* consistent with knowledge of rationality given postures (γ_i, γ_j) *if $\pi_i(\gamma_j) \geq \varepsilon_j$; $\pi_i(\gamma_j) > \varepsilon_j$ only if there exists π_j such that $\pi_j(\gamma_i) \geq \varepsilon_i$ and $\gamma_j \in \Sigma_j^*(\pi_j)$; and, for all $\sigma_j \neq \gamma_j$, $\sigma_j \in \operatorname{supp}(\pi_i)$ only if there exists π_j such that $\pi_j(\gamma_i) \geq \varepsilon_i$ and $\sigma_j \in \Sigma_j^*(\pi_j)$. Let $\Pi_i^{\gamma_i, \gamma_j}$ be the set of player i's beliefs that are consistent with knowledge of rationality given postures (γ_i, γ_j). Player i's* maxmin payoff given postures (γ_i, γ_j) *is*

$$u_i^*(\gamma_i, \gamma_j) \equiv \sup_{\sigma_i} \inf_{\pi_i \in \Pi_i^{\gamma_i, \gamma_j}} u_i(\sigma_i, \pi_i).$$

[21] The events that player 1 and player 2 become committed need not be independent.

Player i's maxmin payoff *is*

$$u_i^* \equiv \sup_{\gamma_i} \inf_{\gamma_j} u_i^* \left(\gamma_i, \gamma_j\right).$$

A posture γ_i^* is a maxmin posture (of player i's) *if there exists a sequence of postures $\{\gamma_n\}$ such that $\gamma_n (t) \to \gamma_i^* (t)$ for all $t \in \mathbb{R}_+$ and* $\inf_{\gamma_j} u_i^* \left(\gamma_n, \gamma_j\right) \to u_i^*$.

Note that if $\varepsilon_j = 0$ then all of these definitions (for player i) reduce to the corresponding definitions in the one-sided commitment model. Thus, writing $u_i^* (\varepsilon_i, \varepsilon_j)$ for player i's maxmin payoff in the two-sided commitment model when the commitment probabilities are ε_i and ε_j, it follows that $u_i^* (\varepsilon_i, 0) = u_i^* (\varepsilon_i)$, player i's maxmin payoff in the one-sided commitment model.

I now show that $u_i^* (\varepsilon_i, \varepsilon_j)$ is approximately equal to $u_i^* (\varepsilon_i)$ whenever ε_j is small, and that the maxmin posture is exactly as in the one-sided commitment model. This is simply because player i cannot guarantee herself anything in the event that player j is committed (e.g., if player j's announced posture always demands the entire surplus), which implies that player i guarantees herself as much as possible by conditioning on the event that player j is not committed. In this event, which occurs with probability $1 - \varepsilon_j$, player i can guarantee herself $u_i^* (\varepsilon_i)$, and the only way she can guarantee herself this much is by announcing $\gamma_{\varepsilon_i}^*$.

Theorem 2 *Player i's maxmin payoff is $u_i^* (\varepsilon_i, \varepsilon_j) = (1 - \varepsilon_j) u_i^* (\varepsilon_i)$, and player i's unique maxmin posture is $\gamma_{i,(\varepsilon_i, \varepsilon_j)}^* = \gamma_{\varepsilon_i}^*$.*

Proof. Let γ_j^0 be the posture of player j's given by $\gamma_j^0 (t) = 1$ for all t. Note that $u_i \left(\sigma_i, \gamma_j^0\right) = 0$ for all σ_i. Therefore, $\inf_{\gamma_j} u_i \left(\sigma_i, \gamma_j\right) = 0$ for all σ_i.

Next, let $\Pi_i^{\gamma_i, \gamma_j} (\varepsilon_i, \varepsilon_j)$ be the set of beliefs π_i that are consistent with knowledge of rationality for commitment probabilities $(\varepsilon_i, \varepsilon_j)$, and let $\Pi_i^{\gamma_i} (\varepsilon_i)$ be the analogous set in the one-sided commitment model. I claim that if $\pi_i \in \Pi_i^{\gamma_i, \gamma_j} (\varepsilon_i, \varepsilon_j)$, then there exists $\pi_i' \in \Pi_i^{\gamma_i} (\varepsilon_i)$ such that $\pi_i = (1 - \varepsilon_j) \pi_i' \oplus \varepsilon_j \gamma_j$, where $(1 - \alpha) x \oplus \alpha y$ is the compound lottery that puts weight $1 - \alpha$ on x and α on y. To see this, note that $\pi_i \left(\gamma_j\right) \geq \varepsilon_j$, so there exists a probability distribution π_i' such that $\pi_i = (1 - \varepsilon_j) \pi_i' \oplus \varepsilon_j \gamma_j$. Furthermore, by definition of $\Pi_i^{\gamma_i, \gamma_j} (\varepsilon_i, \varepsilon_j)$, $\sigma_j \in \text{supp} (\pi_i')$ only if there exists π_j such that $\pi_j (\gamma_i) \geq \varepsilon_i$ and $\sigma_j \in \Sigma_j^* (\pi_j)$

(whether or not σ_j equals γ_j).[22] By definition of $\Pi_i^{\gamma_i}(\varepsilon_i)$, this implies that $\pi_i' \in \Pi_i^{\gamma_i}(\varepsilon_i)$.

Combining the above observations,

$$\begin{aligned}
\inf_{\gamma_j} u_i^*\left(\gamma_i,\gamma_j\right) &= \inf_{\gamma_j}\sup_{\sigma_i} \inf_{\pi_i \in \Pi_i^{\gamma_i,\gamma_j}(\varepsilon_i,\varepsilon_j)} u_i\left(\sigma_i,\pi_i\right) \\
&= \inf_{\gamma_j}\sup_{\sigma_i} \inf_{\pi_i' \in \Pi_i^{\gamma_i}(\varepsilon_i)} \left(1-\varepsilon_j\right) u_i\left(\sigma_i,\pi_i'\right) + \varepsilon_j u_i\left(\sigma_i,\gamma_j\right). \\
&= \sup_{\sigma_i} \inf_{\pi_i' \in \Pi_i^{\gamma_i}(\varepsilon_i)} \left(1-\varepsilon_j\right) u_i\left(\sigma_i,\pi_i'\right) + \varepsilon_j (0) \\
&= \left(1-\varepsilon_j\right) u_i^*\left(\gamma_i\right).
\end{aligned}$$

Therefore, the definitions of $u_i^*(\varepsilon_i,\varepsilon_j)$ and $u_i^*(\varepsilon_i)$ imply that $u_i^*(\varepsilon_i,\varepsilon_j) = \sup_{\gamma_i}(1-\varepsilon_j)\,u_i^*(\gamma_i) = (1-\varepsilon_j)\,u_i^*(\varepsilon_i)$. Similarly, the definition of a maxmin posture in the one-sided commitment model implies that $\gamma^*_{i,(\varepsilon_i,\varepsilon_j)}$ is a maxmin posture in the two-sided commitment model if and only if it is a maxmin posture in the one-sided commitment model with $\varepsilon = \varepsilon_i$. ■

Theorem 2 implies that the qualitative insights of Theorem 1 also apply with two-sided commitment. For example, fixing any ε_2 bounded away from 0, $u_1^*(\varepsilon_1,\varepsilon_2)$ goes to 0 at a logarithmic rate in ε_1. Thus, Theorem 2 says much more than that $u_1^*(\varepsilon_1,\varepsilon_2)$ is continuous in ε_2 at $\varepsilon_2 = 0$. Table 2 displays the maxmin payoff for both the one-sided commitment model and the two-sided commitment model in the case where $\varepsilon_1 = \varepsilon_2 = \varepsilon$:

ε	$u_i^*(\varepsilon)$	$(1-\varepsilon)\,u_i^*(\varepsilon)$
.25	.42	.31
.1	.30	.27
10^{-3}	.13	.13
10^{-6}	.07	.07
10^{-9}	.05	.05

Table 2: The Maxmin Payoff for Different Commitment

Probabilities ε with One- and Two-Sided Commitment

Finally, Definition 8 specifies that player i's belief is consistent with knowledge of rationality only if it assigns probability exactly ε_j to the event that player j is committed to ε_j.

[22] Here, the weaker statement that $\pi_i'\left(\gamma_j\right) > 0$ only if there exists π_j such that $\pi_j(\gamma_i) \geq \varepsilon_i$ and $\sigma_j \in \Sigma_j^*(\pi_i)$ is immediate, and this can be strengthened to the statement that $\gamma_j \in \operatorname{supp}(\pi_i')$ only if there exists such a π_j because the best-response correspondence is upper hemi-continuous in beliefs.

If this were relaxed by specifying that a belief is consistent with knowledge of rationality if it assigns any probability $\varepsilon'_j \leq \varepsilon_j$ to the event that player j is committed to ε_j (and assigns probability $1 - \varepsilon'_j$ to player j's being rational), Theorem 2 and its proof would go through with trivial modifications. Thus, Theorem 2 requires only that player i believes that player j's commitment probability is not more than ε_j, not that player i believes that player j's commitment probability is exactly ε_j.

7 Discrete-Time Bargaining with Frequent Offers

This section considers discrete-time bargaining procedures in which both players can make offers frequently. This includes procedures with any order and relative frequency of offers. I show that, with one-sided commitment, for any sequence of discrete-time bargaining games that converges to continuous time (in the sense that each player may make an offer close to any given time), the corresponding sequence of maxmin payoffs and postures converges to the continuous-time maxmin payoff and posture given by Theorem 1 (the analogous result with two-sided commitment is immediate and is omitted to simplify the exposition). Abreu and Gul (2000) provide a similar independence-of-procedures result for sequential equilibrium outcomes of reputational bargaining. Because my result concerns maxmin payoffs and postures rather than equilibria, my proof is very different from Abreu and Gul's.

Formally, replace the (continuous time) bargaining phase of Section 3 with the following procedure: There is a (commonly known) function $g : \mathbb{R}_+ \to \{0, 1, 2\}$ that specifies who makes an offer at each time. If $g(t) = 0$, no player takes an action at time t. If $g(t) = i \in \{1, 2\}$, then player i makes a demand $u_i(t) \in [0, 1]$ at time t, and player j immediately accepts or rejects. If player j accepts, the game ends with payoffs $(e^{-rt}u_i(t), e^{-rt}(1 - u_i(t)))$; if player j rejects, the game continues. Let $I_i^g = \{t : g(t) = i\}$, and assume that $I_i^g \cap [0, t]$ is finite for all t and that I_i^g is infinite. The announcement phase is correspondingly modified so that player 1 announces a posture $\gamma : I_i^g \to [0, 1]$, and if player 1 becomes committed to posture γ (which continues to occur with probability ε), she demands $\gamma(t)$ at time t and rejects all of player 2's demands. I refer to the function g as a *discrete-time bargaining game*.

I now define convergence to continuous time. This definition is very similar to that of Abreu and Gul (2000), as is the above model of discrete-time bargaining and the corresponding notation.

Definition 9 *A sequence of discrete-time bargaining games* $\{g_n\}$ converges to continuous time *if for all* $\Delta > 0$*, there exists* N *such that for all* $n \geq N$*,* $t \in \mathbb{R}_+$*, and* $i \in \{1, 2\}$*,* $I_i^{g_n} \cap [t, t + \Delta] \neq \emptyset$*.*

The maxmin payoff and posture in a discrete-time bargaining game are defined exactly as in Section 3. Let $u_1^{*,g}$ be player 1's maxmin payoff in discrete-time bargaining game g, and let $u_1^{*,g}(\gamma)$ be player 1's maxmin payoff given posture γ in g. The independence-of-procedures results states that, for any sequence of discrete-time bargaining games converging to continuous time, the corresponding sequence of maxmin payoffs $\{u_1^{*,g_n}\}$ converges to u_1^*, and any corresponding sequence of postures $\{\gamma^{g_n}\}$ such that $u_1^{*,g_n}(\gamma^{g_n}) \to u_1^*$ "converges" to γ^*, where u_1^* and γ^* are the maxmin payoff and posture identified in Theorem 1. The nature of the convergence of the sequence $\{\gamma^{g_n}\}$ to γ^* is slightly delicate. For example, there may be (infinitely many) times $t \in \mathbb{R}_+$ such that $\lim_{n\to\infty} \gamma^{g_n}(t)$ exists and is greater than $\gamma^*(t)$, because these demands may be "non-serious" (in that they are followed immediately by lower demands).[23] Thus, rather than stating the convergence in terms of $\{\gamma^{g_n}\}$ and γ^*, I state it in terms of the corresponding continuation values of player 2, which are the economically more important variables. Formally, given a posture γ^{g_n} in discrete-time bargaining game g^n, let

$$v^{g_n}(t) \equiv \max_{\tau \geq t : \tau \in I_1^{g_n}} e^{-r(\tau - t)} \left(1 - \gamma^{g_n}(\tau)\right).$$

Let $v^*(t) \equiv \max\{1 - e^{rt}/(1 - \log \varepsilon), 0\}$, the continuation value corresponding to γ^* in the continuous-time model of Section 3. The independence-of-procedures result is as follows:

Theorem 3 *Let* $\{g_n\}$ *be a sequence of discrete-time bargaining games converging to continuous time. Then* $u_1^{*,g_n} \to u_1^*$*, and if* $\{\gamma^{g_n}\}$ *is a sequence of postures with* γ^{g_n} *a posture in* g_n *and* $u_1^{*,g_n}(\gamma^{g_n}) \to u_1^*$*, then* $v^{g_n}(t) \to v^*(t)$ *for all* $t \in \mathbb{R}_+$*.*

[23] The reason this complication does not arise in Theorem 1 is that the assumption that $\gamma(t)$ is continuous at non-integer times rules out "non-serious" demands.

The key fact behind the proof of Theorem 3 is that for any sequence of discrete-time postures $\{\gamma^{g_n}\}$ converging to some continuous-time posture γ, $\lim_{n\to\infty} u_1^{*,g_n}(\gamma^{g_n}) = \lim_{n\to\infty} u_1^{*}(\gamma^{g_n})$ (where $u_1^{*}(\gamma^{g_n})$ is the maxmin payoff given a natural embedding of γ^{g_n} in continuous time, defined formally in the supplementary appendix). This fact is proved by constructing a belief that is similar to the γ^{g_n}-offsetting belief in each discrete-time game g_n and then showing that these beliefs converge to the γ-offsetting belief in the limiting continuous-time game.

8 Conclusion

This paper analyzes a model of reputational bargaining in which players initially announce postures to which they may become committed and then bargain over a unit of surplus. It characterizes the highest payoff that a player can guarantee herself under first-order knowledge of rationality, along with the bargaining posture that she must announce in order to guarantee herself this much. A key step in the characterization is showing that this maxmin payoff is the payoff a player receives when her opponent holds the "offsetting belief" that she mixes between following her announced posture and accepting her opponent's demand at a specific rate. Technically, this intermediate result lets one evaluate a posture in terms of its performance against an opponent who holds the corresponding offsetting belief, rather than having to check it's performance against every rational opposing strategy. Conceptually, it shows that the maxmin payoff is also the lowest payoff that can be obtained by a player who knows her opponent's strategy, establishing an equivalence between "maxmin" and "Bayesian" definitions of the highest guaranteed payoff.

I find that a player can guarantee herself a relatively high share of the surplus even if her probability of becoming committed is very small, and that the unique bargaining posture that guarantees this much is simply demanding this share of the surplus in addition to compensation for any delay in reaching agreement. These insights apply for one- or two-sided commitment, for any bargaining procedure with frequent offers, for general bargaining sets, for heterogeneous discount factors, and for any level of knowledge of rationality. In addition, if a player could only announce postures that always demand the same share of the surplus (as in most of the existing literature), her maxmin payoff would be approximately e

times lower.

These results are intended to complement the existing equilibrium analysis of reputational bargaining models. Consider the fundamental question, "What posture should a bargainer stake out?" In equilibrium analysis, the answer to this question depends on her opponent's beliefs about her continuation play following every possible announcement. Yet it may be impossible for either the bargainer or an outside observer to learn these beliefs, especially when bargaining is one-shot. Hence, an appealing alternative approach is to look for a posture that guarantees a high payoff against *any* belief of one's opponent, and for the highest payoff that each player can guarantee herself. This paper shows that this approach yields sharp and economically plausible results, while addressing important concerns about robustness.

The results of this paper are particularly applicable in models where the division of the surplus is of primary importance (rather than the details of how bargaining proceeds, which depend on additional behavioral assumptions). A leading example is the class of models where two parties make costly ex ante investments and then bargain over the resulting surplus. It is a direct consequence of Theorem 2 that, if the players' commitment probabilities are small, both players benefit from comparable increases in their commitment probabilities.[24] Note that this is distinct from the idea that both players benefit from reducing delay; rather, both players benefit because higher commitment probabilities reduce the scope for pessimism about how bargaining will proceed (which would not be possible in an equilibrium analysis). This implies that, for example, comparably increasing both players' commitment powers increases investments whenever investments are complementary. It seems likely that additional insights could be derived by further studying non-equilibrium models of bargaining both in this class of models and in other applied theory models involving bargaining.

Finally, I discuss two additional interesting issues for future research. First, an earlier version of this paper extends the model to multilateral bargaining, where $n \geq 3$ players

[24] To see this somewhat more formally, recall that Theorem 2 states that player i's maxmin payoff is $u_i^*(\varepsilon_i, \varepsilon_j) = \frac{1-\varepsilon_j}{1-\log \varepsilon_i}$. Therefore, $\frac{\partial u_i^*(\varepsilon_i,\varepsilon_j)}{\partial \varepsilon_i} = \frac{1-\varepsilon_j}{\varepsilon_i(1-\log \varepsilon_i)^2}$, while $\frac{\partial u_i^*(\varepsilon_i,\varepsilon_j)}{\partial \varepsilon_j} = -\frac{1}{1-\log \varepsilon_i}$. Since $\lim_{\varepsilon_i \to 0} \varepsilon_i (1 - \log \varepsilon_i) = 0$, it follows that $\lim_{\varepsilon_i \to 0} \left| \frac{\partial u_i^*(\varepsilon_i,\varepsilon_j)/\partial \varepsilon_i}{\partial u_i^*(\varepsilon_i,\varepsilon_j)/\partial \varepsilon_j} \right| = \infty$. In this sense, an increase in ε_i increases player i's maxmin payoff by much more than an increase in ε_j decreases it.

must unanimously agree on the division of the surplus. In such a model, player j may reject any proposal if he expects player k to do so as well, and vice versa. Hence, a player with commitment power cannot guarantee herself a positive payoff under knowledge of rationality. It therefore remains to be seen whether reputational models can make sharp and robust predictions in multilateral bargaining.

Second, it would be interesting to analyze commitment and reputation-building under knowledge of rationality in dynamic games other than bargaining, noting that the definition of a player's maxmin payoff and posture extends to general games. One intriguing observation is that the rate at which the reputation-builder's maxmin payoff converges to her lowest feasible payoff as her commitment probability converges to 0 is slower in my model than in existing repeated game models. In particular, the reputation-builder's (player 1's) maxmin payoff converges to her minimum payoff of 0 at a logarithmic rate in ε in my model, while in repeated game models this convergence is at a polynomial rate in ε.[25] To understand this difference, recall from the proof of Proposition 2 that if player 1 announces a constant posture γ, then her maxmin payoff given posture γ equals $\varepsilon^{\gamma/(1-\gamma)}\gamma$, which is polynomial in ε. However, the maxmin constant posture is increasing in ε (and goes to 0 as $\varepsilon \to 0$), and Corollary 2 shows that player 1 can guarantee herself a payoff that goes to 0 at a logarithmic rate in ε by appropriately recalibrating her announced posture as $\varepsilon \to 0$. Thus, roughly speaking, the reason why reputation bounds in the repeated games literature converge to 0 more quickly than in my model is that there is generally no way to continuously moderate one's posture as one's commitment probability decreases in repeated games.

Appendix: Omitted Proofs for Sections 3 and 4

Proof of Proposition 1. The proof uses results from Section 4, and therefore should not be read before reading Section 4.

[25] Fudenberg and Levine (1989) show that if player 1 is committed to her Stackelberg action with probability ε and player 2 is myopic, then player 1's payoff in any Nash equilibrium is at least $\varepsilon^{r_1\alpha}\tilde{u}_1 + (1-\varepsilon^{r_1\alpha})\,\underline{u}_1$, for some constant $\alpha > 0$, where $\tilde{u}_1$ is her Stackelberg payoff and $\underline{u}_1$ is her lowest feasible payoff. This bound is the basis for most of the subsequent literature; for example, convergence to $\underline{u}_1$ is also polynomial in ε in Schmidt (1993), Cripps, Schmidt, and Thomas (1996), and Evans and Thomas (1997).

Fix a posture γ and payoff $u_1 \in [u_1^*(\gamma), 1)$. If $u_1 \neq \gamma(0)$, then let $\hat{\sigma}_2^\gamma$ be identical to the γ-offsetting strategy defined in Definition 6, with the modification that player 1's demand is accepted at any history h^t at which player 1 has demanded u_1 at all previous dates. If $u_1 = \gamma(0)$, then let $\hat{\sigma}_2^\gamma$ be identical to the γ-offsetting strategy defined in Definition 6, with the modification that player 1's demand is accepted at date $(-\log(u_1)/r, -1)$ if player 1 has demanded 1 at all previous dates. In either case, let π_2^γ be as in Definition 6, and note that $\pi_2^\gamma(\gamma) \geq \varepsilon$. If $u_1 \neq \gamma(0)$, no strategy under which $u_1(0) = u_1$ is in the support of π_2^γ; similarly, if $u_1 = \gamma(0)$, no strategy under which $u_1(0) = 1$ is in the support of π_2^γ (since $u_1 < 1$). Therefore, the same argument as in the proof of Lemma 2 shows that $\hat{\sigma}_2^\gamma \in \Sigma_2^*(\pi_2^\gamma)$. Hence, the belief $\hat{\pi}_1$ given by $\hat{\pi}_1(\hat{\sigma}_2^\gamma) = 1$ is an element of Π_1^γ. Furthermore, under strategy $\hat{\sigma}_2^\gamma$, player 2 always demands 1 and only accepts player 1's demand if player 1 has either conformed to γ through time t^* (defined in Section 4.2) or has always demanded u_1 (in the $u_1 \neq \gamma(0)$ case) or 1 (in the $u_1 = \gamma(0)$ case). Note that $\exp(-r(-\log(u_1)/r)) = u_1$. Hence, in either case, $u_1(\sigma_1, \hat{\pi}_1) \in \{0, u_1^*(\gamma), u_1\}$ for every strategy σ_1. Let $\hat{\sigma}_1$ be the strategy of player 1's that always demands u_1 (if $u_1 \neq \gamma(0)$) or 1 (if $u_1 = \gamma(0)$) and never accepts player 2's demand. Then $u_1(\hat{\sigma}_1, \hat{\pi}_1) = u_1 = \max_{\sigma_1} u_1(\sigma_1, \hat{\pi}_1)$, completing the proof.
■

Proof of Lemma 1. I prove the result for pure strategies σ_2, which immediately implies the result for mixed strategies.

Fix π_2 such that $\pi_2(\gamma) \geq \varepsilon$ and pure strategy $\sigma_2 \in \Sigma_2^*(\pi_2)$. The plan of the proof is to show that if $\pi_2(\gamma) \geq \varepsilon$ and agreement is not reached by $\tilde{T}$ under strategy profile (γ, σ_2), then player 2 must be certain that player 1 is playing γ at any time $t > \tilde{T}$. This suffices to prove the lemma, because $\sigma_2 \in \Sigma_2^*(\pi_2)$ implies that player 2 accepts $\underline{\gamma}(t)$ no later than time $t = T$ if at any time $t > \tilde{T}$ agreement has not been reached and he is certain that player 1 is playing γ.

Let $\chi^{(\pi_2,\sigma_2)}(t)$ be the probability that player 2 assigns to player 1 *not* playing γ at date $(t, -1)$ when his initial belief is π_2 and play up until date $(t, -1)$ is given by player 1's following strategy γ and player 2's following (pure) strategy σ_2; this is determined by Bayes' rule, because $\pi_2(\gamma) \geq \varepsilon > 0$. By convention, if agreement is reached at time τ, let $\chi^{(\pi_2,\sigma_2)}(t) = \chi^{(\pi_2,\sigma_2)}(\tau)$ for all $t > \tau$. Let $t(\gamma, \sigma_2)$ be the time at which agreement is reached under strategy profile (γ, σ_2) (with the convention that $t(\gamma, \sigma_2) \equiv \infty$ if agreement is never reached under (γ, σ_2)); and let

$$\hat{t}(\gamma, \sigma_2) \equiv \sup \left\{ t : \chi^{(\pi_2, \sigma_2)}(t) > 0 \right\},$$

the latest time at which player 2 is not certain that player 1 is playing γ under strategy profile (γ, σ_2) with belief π_2. Let

$$\hat{T} \equiv \sup_{(\pi_2, \sigma_2): \pi_2(\gamma) \geq \varepsilon, \sigma_2 \in \Sigma_2^*(\pi_2), t(\gamma, \sigma_2) \geq \hat{t}(\gamma, \sigma_2)} \hat{t}(\gamma, \sigma_2). \tag{12}$$

That is, $\hat{T}$ is the latest possible time t at which player 2 is not certain that player 1 is following γ and agreement is not reached by t. I will show that $\hat{T} = \tilde{T}$, which completes the proof.

I first claim that in the definition of $\hat{T}$ it is without loss of generality to restrict attention to (π_2, σ_2) such that σ_2 always demands $u_2(t) = 1$, π_2 puts probability 1 on player 1 conceding at any history h^{t+} at which $u_1(t) \neq \gamma(t)$, and π_2 puts probability 0 on player 1 conceding at any history h^{t-}; that is, that the right-hand side of (12) continues to equal $\hat{T}$ when this additional constraint is imposed. To see this, suppose that (π_2', σ_2') satisfies $\pi_2'(\gamma) \geq \varepsilon$, $\sigma_2' \in \Sigma_2^*(\pi_2')$, and $t(\gamma, \sigma_2') \geq \hat{t}(\gamma, \sigma_2')$ (the constraints of (12)). Let π_2 be the belief under which player 1 demands $u_1(t) = \gamma(t)$ for all $t \in \mathbb{R}_+$; accepts player 2's demand at every history of the form $(\gamma(\tau), 1)_{\tau \leq t}$ at the same rate and probability at which player 1 deviates from γ at time t (i.e., at date $(t, -1)$, $(t, 0)$, or $(t, 1)$) under strategy profile (π_2', σ_2') (viewing π_2' as a mixed strategy of player 1's); and rejects player 2's demand at every other history. Clearly, there exists a strategy $\sigma_2 \in \Sigma_2^*(\pi_2)$ that always demands $u_2(t) = 1$. Note that player 1's rate and probability of deviating from γ at history $(\gamma(\tau), 1)_{\tau \leq t}$ under belief π_2 is the same as at time t under strategy profile (π_2', σ_2'), and that player 2's continuation payoff after such a deviation is weakly higher in the former case. Recall that strategy γ never accepts player 2's demand, so agreement is reached only if player 2 accepts player 1's demand or if player 1 has deviated from γ. Therefore, since rejecting player 1's demand $\gamma(t)$ under strategy profile (π_2', σ_2') is optimal for all $t < t(\gamma, \sigma_2')$, it follows that rejecting player 1's demand $\gamma(t)$ at history $(\gamma(\tau), 1)_{\tau \leq t}$ is optimal under belief π_2, for all $t < t(\gamma, \sigma_2')$. This implies that $t(\gamma, \sigma_2) \geq t(\gamma, \sigma_2')$. Furthermore, $\chi^{(\pi_2, \sigma_2)}(t) = \chi^{(\pi_2', \sigma_2')}(t)$ for all $t \in \mathbb{R}_+$, so $\hat{t}(\gamma, \sigma_2) = \hat{t}(\gamma, \sigma_2')$. Hence, $t(\gamma, \sigma_2) \geq \hat{t}(\gamma, \sigma_2)$. Finally, $\pi_2(\gamma) \geq \varepsilon$. Therefore, (π_2, σ_2) satisfies the constraints of (12); σ_2 always demands $u_2(t) = 1$; π_2 puts probability 1 on player 1 conceding at any history h^{t+} at which $u_1(t) \neq \gamma(t)$; π_2 puts probability 0 on player 1 conceding at any history h^{t-}; and $\hat{t}(\gamma, \sigma_2) \geq \hat{t}(\gamma, \sigma_2')$; so the right-hand side of (12) continues to equal $\hat{T}$ when the additional constraint is imposed.

Thus, fix a belief π_2 that puts probability 1 on player 1 conceding at any history h^{t+} at which $u_1(t) \neq \gamma(t)$, and puts probability 0 on player 1 conceding at any history h^{t-}. Let $\lambda^{\pi_2}(t)$ and $p^{\pi_2}(t)$ be the concession rate and probability of player 1 at history $(\gamma(\tau), 1)_{\tau \leq t}$ when her strategy is given by π_2; let S^{π_2} be the (countable) set of times s such that $p^{\pi_2}(s) > 0$; and let $\hat{t}(\pi_2) \equiv \hat{t}(\gamma, \sigma_2^0)$, where σ_2^0 is the strategy that always demands $u_2(t) = 1$ and always rejects player 1's demand. Fixing a strategy $\sigma_2 \in \Sigma_2^*(\pi_2)$ that always demands $u_2(t) = 1$, note that (γ, σ_2) and (γ, σ_2^0) induce the same path of play until time $t(\gamma, \sigma_2)$, and therefore $t(\gamma, \sigma_2) \geq \hat{t}(\gamma, \sigma_2)$ if and only if $t(\gamma, \sigma_2) \geq \hat{t}(\pi_2)$. Hence, $t(\gamma, \sigma_2) \geq \hat{t}(\gamma, \sigma_2)$ if and only if it is optimal for player 2 to reject player 1's offer until time $\hat{t}(\pi_2)$, when his initial belief is π_2 and player 1 plays γ. I claim that this holds if and only if

$$
\begin{aligned}
& 1 - \gamma(t) \\
\leq \; & \int_t^{\hat{t}(\pi_2)} \exp\left(-r(\tau - t) - \int_t^{\tau} \lambda^{\pi_2}(s)\, ds\right) \left(\prod_{s \in S^{\pi_2} \cap (t,\tau)} (1 - p^{\pi_2}(s))\right) \lambda^{\pi_2}(\tau)\, d\tau \\
& + \sum_{s \in S^{\pi_2} \cap \left(t, \hat{t}(\pi_2)\right)} \exp\left(-r(s - t) - \int_t^{s} \lambda^{\pi_2}(q)\, dq\right) \left(\prod_{q \in S^{\pi_2} \cap (t,s)} (1 - p^{\pi_2}(q))\right) p^{\pi_2}(s) \\
& + \exp\left(-r\left(\hat{t}(\pi_2) - t\right) - \int_t^{\hat{t}(\pi_2)} \lambda^{\pi_2}(s)\, ds\right) \left(\prod_{s \in S^{\pi_2} \cap \left(t, \hat{t}(\pi_2)\right)} (1 - p^{\pi_2}(s))\right) v\left(\hat{t}(\pi_2)\right)
\end{aligned}
$$

$$
\text{for all } t < \hat{t}(\pi_2). \tag{13}
$$

The left-hand side of (13) is player 2's payoff from accepting player 1's demand at date $(t, 1)$ when $p^{\pi_2}(t) = 0$. The right-hand side of (13) is player 2's continuation payoff from rejecting player 1's demand until time $\hat{t}(\pi_2)$ when $p^{\pi_2}(t) = 0$. Thus, (13) must hold if $t(\gamma, \sigma_2) \geq \hat{t}(\pi_2)$. It remains to show that (13) implies that it is optimal for player 2 to reject at times where $p^{\pi_2}(t) > 0$. Suppose that $p^{\pi_2}(t) > 0$. At date $(t, -1)$, the fact that S^{π_2} is countable and (13) holds at all times before t that are not in S^{π_2} implies that

$\lim_{\tau\uparrow t}(1-\gamma(\tau))$ is weakly less than player 2's continuation payoff from rejecting playing 1's demand until time $\hat{t}(\pi_2)$. Furthermore, the fact that player 1 concedes with probability 0 at date $(t,-1)$ implies that $\lim_{\tau\uparrow t}(1-\gamma(\tau))$ is indeed player 2's payoff from accepting at date $(t,-1)$. Thus, rejecting is optimal at date $(t,-1)$. At date $(t,1)$, player 2's payoff from accepting is $(1-p^{\pi_2}(t)/2)(1-\gamma(t))+(p^{\pi_2}(t)/2)(1)$, while his continuation payoff from rejecting until time $\hat{t}(\pi_2)$ is $1-p^{\pi_2}(t)$ times the right-hand side of (13) plus $p^{\pi_2}(t)(1)$. Hence, (13) implies that rejecting is optimal at date $(t,1)$ as well.

In addition, (13) holds if and only if

$$\begin{aligned}
& v(t) \\
\leq\ & \int_t^{\hat{t}(\pi_2)} \exp\left(-r(\tau-t)-\int_t^{\tau}\lambda^{\pi_2}(s)\,ds\right)\left(\prod_{s\in S^{\pi_2}\cap(t,\tau)}\left(1-p^{\pi_2}(s)\right)\right)\lambda^{\pi_2}(\tau)\,d\tau \\
& +\sum_{s\in S^{\pi_2}\cap\left(t,\hat{t}(\pi_2)\right)}\exp\left(-r(s-t)-\int_t^{s}\lambda^{\pi_2}(q)\,dq\right)\left(\prod_{q\in S^{\pi_2}\cap(t,s)}\left(1-p^{\pi_2}(q)\right)\right)p^{\pi_2}(s) \\
& +\exp\left(-r\left(\hat{t}(\pi_2)-t\right)-\int_t^{\hat{t}(\pi_2)}\lambda^{\pi_2}(s)\,ds\right)\left(\prod_{s\in S^{\pi_2}\cap\left(t,\hat{t}(\pi_2)\right)}\left(1-p^{\pi_2}(s)\right)\right)v\left(\hat{t}(\pi_2)\right) \\
& \text{for all } t < \hat{t}(\pi_2). \qquad (14)
\end{aligned}$$

To see this, note that (14) immediately implies (13) because $v(t)\geq 1-\gamma(t)$ for all t. For the converse, suppose that (13) holds. If $v(t)>1-\gamma(t)$ then $v(t)=e^{-r(\tau-t)}\left(1-\underline{\gamma}(\tau)\right)$ for some $\tau>t$ such that $v(\tau,-1)=1-\underline{\gamma}(\tau)$, which implies that $v(\tau,-1)$ is weakly less than the limit as $s\uparrow\tau$ of the right-hand side of (14) evaluated at time s (with the convention that the right-hand side of (14) equals $v(s)$ if $s\geq\hat{t}(\pi_2)$). Now the right-hand side of (14) at time t is at least $e^{-r(\tau-t)}$ times as large as is this limit, which implies that the right-hand side of (14) at time t is at least $e^{-r(\tau-t)}v(\tau,-1)=v(t)$. Hence, (14) holds.

By the previous two paragraphs, (12) may be rewritten as

$$\hat{T}=\sup_{\substack{\pi_2:\pi_2(\gamma)\geq\varepsilon,\\ (14)\ holds}}\ \sup\left\{t:\chi^{\pi_2}(t)\equiv\frac{\exp\left(-\int_0^t\lambda^{\pi_2}(s)\,ds\right)\prod_{s\in S^{\pi_2}\cap[0,t)}\left(1-p^{\pi_2}(s)\right)-\varepsilon}{\exp\left(-\int_0^t\lambda^{\pi_2}(s)\,ds\right)\prod_{s\in S^{\pi_2}\cap[0,t)}\left(1-p^{\pi_2}(s)\right)}>0\right\}. \qquad (15)$$

I first show that there exists some belief π_2 that both attains the (outer) supremum in (15) (with the convention that the supremum is attained at π_2 if $\hat{t}(\pi_2) = \hat{T} = \infty$) and also maximizes $\lim_{t \uparrow \hat{T}} \chi^{\pi_2}(t)$ over all beliefs π_2 that attain the supremum (note that this limit exists for all π_2, because $\chi^{\pi_2}(t)$ is non-increasing). I also show that (14) must hold with equality (at all $t < \hat{T}$) under any such belief π_2, which implies that (15) may be solved under the additional constraint that (14) holds with equality.

First, fix a sequence $\{\chi^{\pi_2^n}\}$ such that $\hat{t}(\pi_2^n) \uparrow \hat{T}$, $\pi_2^n(\gamma) \geq \varepsilon$ for all n, and (14) holds for all n. Note that $\chi^{\pi_2^n}(t)$ is non-increasing in t, for all n. Since the space of monotone functions from $\mathbb{R}_+$ to $[0,1]$ is sequentially compact (by Helly's selection theorem (see, e.g., Billingsley (1995) Theorem 25.9)), there exists a subsequence $\{\chi^{\pi_2^m}\}$ that converges pointwise to some (non-increasing) function χ^{π_2}.[26] Furthermore, $\chi^{\pi_2}(0) = 1 - \varepsilon$, because $\chi^{\pi_2^m}(0) = 1 - \varepsilon$ for all m. Combined with the fact that χ^{π_2} is non-increasing, this implies that there exists a pair of functions $(\lambda^{\pi_2} : \mathbb{R}_+ \to \mathbb{R}_+, p^{\pi_2} : \mathbb{R}_+ \to [0,1])$ such that $p^{\pi_2}(t) = 0$ for all t outside of a countable set S^{π_2} and

$$\chi^{\pi_2}(t) = \frac{\exp\left(-\int_0^t \lambda^{\pi_2}(s)\,ds\right)\prod_{s \in S^{\pi_2} \cap [0,t)}(1 - p^{\pi_2}(s)) - \varepsilon}{\exp\left(-\int_0^t \lambda^{\pi_2}(s)\,ds\right)\prod_{s \in S^{\pi_2} \cap [0,t)}(1 - p^{\pi_2}(s))}$$

for all t. Therefore, there exists a belief $\pi_2 \in \Delta(\Sigma_1)$ corresponding to concession rate (resp., probability) $\lambda^{\pi_2}(t)$ (resp., $p^{\pi_2}(t)$) such that $\pi_2(\gamma) \geq \varepsilon$. Finally, the fact that $\chi^{\pi_2^m}(t) \to \chi^{\pi_2}(t)$ for all t implies that

$$\exp\left(-\int_0^t \lambda^{\pi_2^m}(s)\,ds\right)\prod_{s \in S^{\pi_2} \cap [0,t)}\left(1 - p^{\pi_2^m}(s)\right) \to \exp\left(-\int_0^t \lambda^{\pi_2}(s)\,ds\right)\prod_{s \in S^{\pi_2} \cap [0,t)}\left(1 - p^{\pi_2}(s)\right)$$

for all t. Since for all $t < \hat{T}$, there exists $M > 0$ such that (14) holds at time t under π_2^m for all $m > M$, this implies that (14) holds at all times $t < \hat{T}$ under π_2.

[26] Showing that the space of monotone functions from $\mathbb{R}_+ \to [0,1]$ is sequentially compact requires a slightly different version of Helly's selection theorem than that in Billingsley (1995), so here is a direct proof: If $\{f_n\}$ is a sequence of monotone functions $\mathbb{R}_+ \to [0,1]$, then there exists a subsequence $\{f_m\} \subseteq \{f_n\}$ that converges on $\mathbb{Q}_+$ to a monotone function $f : \mathbb{Q}_+ \to [0,1]$. Let $\tilde{f} : \mathbb{R}_+ \to [0,1]$ be given by $\tilde{f}(x) = \lim_{l \to \infty} f(x_l)$, where $\{x_l\}_{l=1}^{\infty} \uparrow x$ and $x_l \in \mathbb{Q}_+$ for all l. Then $\tilde{f}$ is monotone, which implies that there is a countable set $\mathcal{S}$ such that $\tilde{f}$ is continuous on $\mathbb{R}_+ \backslash \mathcal{S}$. Since $\mathcal{S}$ is countable, there exists a sub-subsequence $\{f_k\} \subseteq \{f_m\}$ such that $\{f_k\}$ converges on $\mathcal{S}$. Finally, let $\hat{f}(x) = \tilde{f}(x)$ if $x \in \mathbb{R}_+ \backslash \mathcal{S}$ and $\hat{f}(x) = \lim_{k \to \infty} f_k(x)$ if $x \in \mathcal{S}$. Then $\{f_k\} \to \hat{f}$.

I now show that if there exists a time $t < \hat{T}$ at which (14) holds with *strict* inequality under belief π_2, then there exists an alternative belief π_2' that attains the supremum in (15). Suppose such a time t exists. I claim that it follows that there exists a time $t_1 < \hat{t}(\pi_2)$ at which (14) holds with strict inequality and in addition either $\int_{t_1}^{t_1+\Delta} \lambda^{\pi_2}(s)\,ds > 0$ for all $\Delta > 0$ or $\sum_{s \in S^{\pi_2} \cap [t_1, t_1+\Delta)} p^{\pi_2}(s) > 0$ for all $\Delta > 0$. To see this, note that there must exist a time $t' \in \left(t, \hat{t}(\pi_2)\right)$ such that either $\int_{t'}^{t'+\Delta} \lambda^{\pi_2}(s)\,ds > 0$ for all $\Delta > 0$ or $p^{\pi_2}(t') > 0$ (because otherwise (14) could not hold with strict inequality at t). Let t_1 be the infimum of such times t', and note that either $\int_{t_1}^{t_1+\Delta} \lambda^{\pi_2}(s)\,ds > 0$ for all $\Delta > 0$ or $\sum_{s \in S^{\pi_2} \cap [t_1, t_1+\Delta)} p^{\pi_2}(s) > 0$ for all $\Delta > 0$. Then the fact that (14) holds with strict inequality at time t implies that (14) holds with strict inequality at time t_1, because otherwise the fact that $\int_t^{t_1} \lambda^{\pi_2}(s)\,ds = 0$ and $p^{\pi_2}(t'') = 0$ for all $t'' \in [t, t_1)$ would imply that (14) could not hold with strict inequality at time t.

Thus, let $t_0 < \hat{T}$ be such that (14) holds with strict inequality at time t_0 and in addition $\int_{t_0}^{t_0+\Delta} \lambda^{\pi_2}(s)\,ds > 0$ for all $\Delta > 0$ (the case where $\sum_{s \in S^{\pi_2} \cap [t_0, t_0+\Delta)} p^{\pi_2}(s) > 0$ is similar, and thus omitted). Since $v(t)$ is continuous but for downward jumps, there exist $\eta > 0$ and $\Delta > 0$ such that (14) holds with strict inequality at t for all $t \in [t_0, t_0 + \Delta)$ when $\lambda^{\pi_2}(t)$ is replaced by $(1-\eta)\lambda^{\pi_2}(t)$ for all $t \in [t_0, t_0+\Delta)$. Define $\lambda^{\pi_2\prime}(t)$ by $\lambda^{\pi_2\prime}(t) \equiv \lambda^{\pi_2}(t)$ for all $t \notin [t_0, t_0+\Delta)$ and $\lambda^{\pi_2\prime}(t) \equiv (1-\eta)\lambda^{\pi_2}(t)$ for all $t \in [t_0, t_0+\Delta)$. Next, I claim that at time t_0 player 2's continuation payoff from rejecting γ until $\hat{t}(\pi_2)$ is strictly lower when player 1's concessions are given by $\left(\lambda^{\pi_2}(t), p^{\pi_2}(t)\right)$ than when they are given by $\left(\lambda^{\pi_2\prime}(t), p^{\pi_2\prime\prime}(t)\right)$, where $p^{\pi_2\prime\prime}(t)$ is defined by $p^{\pi_2\prime\prime}(t) \equiv p^{\pi_2}(t)$ for all $t \neq t_0$, and $p^{\pi_2\prime\prime}(t_0) \equiv 1 - \exp\left(-\eta \int_t^{t+\Delta} \lambda^{\pi_2}(s)\,ds\right)\left(1 - p^{\pi_2}(t_0)\right) > 0$. This follows because the total probability with which player 1 concedes in the interval $[t_0, t_0+\Delta)$ is the same under $\left(\lambda^{\pi_2}(t), p^{\pi_2}(t)\right)$ and under $\left(\lambda^{\pi_2\prime}(t), p^{\pi_2\prime\prime}(t)\right)$, and some probability mass of concession is moved earlier to t_0 under $\left(\lambda^{\pi_2\prime}(t), p^{\pi_2\prime\prime}(t)\right)$. Therefore, there exists $\zeta > 0$ such that at time t_0 player 2's continuation payoff from rejecting γ until $\hat{t}(\pi_2)$ is the same when player 1's concessions are given by $\left(\lambda^{\pi_2}(t), p^{\pi_2}(t)\right)$ and when they are given by $\left(\lambda^{\pi_2\prime}(t), p^{\pi_2\prime}(t)\right)$, where $p^{\pi_2\prime}(t)$ is defined by $p^{\pi_2\prime}(t) \equiv p^{\pi_2}(t)$ for all $t \neq t_0$, and $p^{\pi_2\prime}(t_0) \equiv (1-\zeta)\,p^{\pi_2\prime\prime}(t_0) < p^{\pi_2\prime\prime}(t_0)$. The fact that (14) holds at all $t < \hat{T}$ when player 1's concessions are given by $\left(\lambda^{\pi_2}(t), p^{\pi_2}(t)\right)$ now implies that (14) holds at all $t < \hat{T}$ when player 1's concessions are given by $\left(\lambda^{\pi_2\prime}(t), p^{\pi_2\prime}(t)\right)$. Furthermore, $\exp\left(-\int_0^{\hat{T}}\lambda^{\pi_2\prime}(t)\,dt\right)\prod_{s\in S^{\pi_2}\cap[0,\hat{T})}\left(1-p^{\pi_2\prime}(s)\right) > \exp\left(-\int_0^{\hat{T}}\lambda^{\pi_2}(t)\,dt\right)\prod_{s\in S^{\pi_2}\cap[0,\hat{T})}\left(1-p^{\pi_2}(s)\right)\geq\varepsilon$. Therefore, $\sup\left\{t:\chi^{\pi_2'}(t)>0\right\}\geq\hat{T}$, so by the definition of $\hat{T}$ it must be that $\sup\left\{t:\chi^{\pi_2'}(t)>0\right\}=\hat{T}$.

Next, suppose that (14) holds with equality under belief π_2 (defined above), and that in addition $v(t)$ is differentiable at some time $t<\hat{t}(\pi_2)$. Then the derivative of the right-hand side of (14) at t must exist and equal $v'(t)$. This implies that $p^{\pi_2}(t)=0$, and, by Leibniz's rule, the derivative of the right-hand side of (14) equals $-\lambda^{\pi_2}(t)+\left(r+\lambda^{\pi_2}(t)\right)v(t)$. Hence,

$$\lambda^{\pi_2}(t)=\frac{rv(t)-v'(t)}{1-v(t)}.$$

Since $v(t)$ is differentiable almost everywhere, this implies that

$$\int_0^{\tau}\lambda^{\pi_2}(s)\,ds=\int_0^{\tau}\lambda(s)\,ds \tag{16}$$

for all $\tau<\hat{t}(\pi_2)$, where $\lambda(s)$ is defined by (3). Similarly, if (14) holds with equality then the difference between the limit as $s\uparrow t$ of the right-hand side of (14) evaluated at s and the limit as $s\downarrow t$ of the right-hand side of (14) evaluated at s must equal $v(t,-1)-v(t)$, for all $t<\hat{t}(\pi_2)$. By inspection, this difference equals $p^{\pi_2}(t)-p^{\pi_2}(t)\,v(t)$. Hence,

$$p^{\pi_2}(t)=\frac{v(t,-1)-v(t)}{1-v(t)}$$

for all $t<\hat{t}(\pi_2)$. Therefore,

$$\prod_{s\in S^{\pi_2}\cap[0,\tau)}\left(1-p^{\pi_2}(s)\right)=\prod_{s\in S\cap[0,\tau)}\left(1-p(s)\right) \tag{17}$$

for all $\tau<\hat{t}(\pi_2)$, where S is the set of discontinuity points of $v(t)$, and $p(s)$ is defined by (4). Combining (16) and (17), I conclude that if (14) holds with equality under belief π_2, then

$$\hat{t}(\pi_2)=\sup\left\{t:\exp\left(-\int_0^t\lambda(s)\,ds\right)\prod_{s\in S\cap[0,t)}\left(1-p(s)\right)>\varepsilon\right\},$$

which equals $\tilde{T}$. In addition, $\chi^{\pi_2}(t)<0$ for all $t\in\left(\tilde{T},\hat{T}\right)$, so $\hat{T}=\tilde{T}$ and the supremum in (15) is attained at π_2.

Combining the previous three paragraphs, it follows that the supremum in (15) is always attained at some belief π_2. I now show that there exists a belief that both attains the supremum in (15) and maximizes $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)$ over all beliefs π_2 that attain the supremum in (15). Let $\chi\in[0,1]$ be the supremum of $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)$ over all beliefs π_2 that attain the supremum in (15). If $\chi=0$, then any belief π_2 that attains the supremum in (15) also satisfies $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)=\chi$. Thus, suppose that $\chi>0$. Let $\{\pi_2^n\}$ be a sequence of beliefs that all attain the supremum in (15) such that $\lim_{t\uparrow\hat{T}}\chi^{\pi_2^n}(t)\uparrow\chi$. The above sequential compactness argument implies that there exists a subsequence $\{\pi_2^m\}\subseteq\{\pi_2^n\}$ and a belief π_2 satisfying the constraints of (15) such that $\chi^{\pi_2^m}(t)\to\chi^{\pi_2}(t)$ for all t. Furthermore, $\chi^{\pi_2}(t)$ is non-increasing, so $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)$ exists. Because π_2 satisfies the constraints of (15), $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)\leq\chi$. Now suppose, toward a contradiction, that $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)<\chi$. Then there exists $\eta>0$ and $t'\leq\hat{T}$ such that $\chi^{\pi_2}(t')<\chi-\eta$. Since $\lim_{m\to\infty}\lim_{t\uparrow\hat{T}}\chi^{\pi_2^m}(t)=\chi$, there exists $M>0$ such that, for all $m>M$, $\lim_{t\uparrow\hat{T}}\chi^{\pi^m}(t)>\chi-\eta$. And $\chi^{\pi_2^m}(t)$ is non-increasing for all m, so this implies that $\chi^{\pi_2^m}(t')>\chi-\eta$ for all $m>M$. Now $\chi^{\pi_2^m}(t')\to\chi^{\pi_2}(t')$ implies that $\chi^{\pi_2}(t')\geq\chi-\eta$, a contradiction. Therefore, $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)=\chi$. Furthermore, $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)>0$ implies that π_2 attains the supremum in (15).

Finally, if (14) holds with strict inequality at some time $t<\hat{T}$ under a belief π_2 such that $\hat{t}(\pi_2)=\hat{T}$, the same procedure for modifying π_2 described above yields a belief π_2' such that $\hat{t}(\pi_2')=\hat{T}$ and $\lim_{t\uparrow\hat{T}}\chi^{\pi_2'}(t)>\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)$. This implies that the only beliefs π_2 that both attain the supremum in (15) and maximize $\lim_{t\uparrow\hat{T}}\chi^{\pi_2}(t)$ (over all beliefs that attain the supremum in (15)) satisfy the additional constraint that (14) holds with equality. Since I have proved that such a belief exists, the value of (15) equals the value of (15) under this additional constraint, which I have shown to equal $\tilde{T}$. ■

Proof of Theorem 1. Let γ_n and γ^* be defined as in Section 4.4. Note that $\{\gamma_n\}$ converges pointwise to γ^*. To show that $\lim_{n\to\infty}u_1^*(\gamma_n)=1/(1-\log\varepsilon)$, it remains only to show that $T_n^1>\tilde{T}(\gamma_n)$ for all $n\in\mathbb{N}$. To see this, note that $T_n^1=\frac{1}{r}\log\left(\frac{n+1}{n}(1-\log\varepsilon)\right)$. Since $\gamma_n(t)=\left(\frac{n}{n+1}\right)\frac{e^{rt}}{1-\log\varepsilon}$ for all $t\leq T_n^1$ and $\gamma_n(t)$ is non-decreasing, it follows that $v(t)=$

$1-\left(\frac{n}{n+1}\right)\frac{e^{rt}}{1-\log\varepsilon}$ for all $t\leq T_n^1$. Therefore,

$$\begin{aligned}
&\exp\left(-\int_0^{T_n^1}\frac{rv(t)-v'(t)}{1-v(t)}dt\right)\prod_{s\in S\cap[0,T_n^1]}\left(\frac{1-v(s,-1)}{1-v(s)}\right)\\
=&\exp\left(-\int_0^{T_n^1}r\left(\frac{n+1}{n}\right)(1-\log\varepsilon)e^{-rt}dt\right)\\
=&\exp\left(-\left(\frac{n+1}{n}\right)(1-\log\varepsilon)\left(1-e^{-rT_n^1}\right)\right)\\
=&\exp\left(-\left(\frac{n+1}{n}\right)(1-\log\varepsilon)\left(1-\left(\frac{n}{n+1}\right)\frac{1}{1-\log\varepsilon}\right)\right)\\
=&\exp\left(-\frac{1}{n}(1-\log\varepsilon)\right)\varepsilon\\
<&\ \varepsilon.
\end{aligned}$$

Hence, by the definition of $\tilde{T}(\gamma_n)$, $T_n^1\geq\tilde{T}(\gamma_n)$. Furthermore, the fact that $\exp\left(-\int_0^\tau\frac{rv(t)-v'(t)}{1-v(t)}dt\right)$ is strictly decreasing in τ for all $\tau\in[0,T_n^1]$ implies that $T_n^1>\tilde{T}(\gamma_n)$.

To complete the proof of Theorem 1, I must show that if $\{\gamma_n\}$ is any sequence of postures converging pointwise to some posture γ satisfying $u_1^*(\gamma_n)\to u_1\geq 1/(1-\log\varepsilon)$, then $\gamma=\gamma^*$.[27] There are two steps. First, letting $\{v_n\}$ be the continuation value functions corresponding to the $\{\gamma_n\}$, and letting v^* be the continuation value function corresponding to γ^*, I show that $\sup_{t\in\mathbb{R}_+}e^{-rt}|v^*(t)-v_n(t)|\to 0$. Second, I show that this implies that $\gamma'=\gamma^*$.

Step 1:

Suppose that $u_1^*(\gamma)\geq 1/(1-\log\varepsilon)-\zeta$ for some posture γ and some $\zeta\in(0,1/(1-\log\varepsilon))$. Let $T^1\equiv(1/r)\log(1-\log\varepsilon)$ (which equals $\lim_{n\to\infty}T_n^1$). Then it must be that $\tilde{T}(\gamma)\leq T^1-(1/r)\log(1-\zeta(1-\log\varepsilon))$, for otherwise it would follow from $T(\gamma)\geq\tilde{T}(\gamma)$ that

$$\begin{aligned}
u_1^*(\gamma) &= \min_{t\leq T(\gamma)}e^{-rt}\underline{\gamma}(t)\leq e^{-r\tilde{T}(\gamma)}\underline{\gamma}\left(\tilde{T}(\gamma)\right)\\
&< \exp\left(-rT^1+\log(1-\zeta(1-\log\varepsilon))\right)(1)=\frac{1}{1-\log\varepsilon}-\zeta.
\end{aligned}$$

Furthermore, if $u_1^*(\gamma)\geq 1/(1-\log\varepsilon)-\zeta$, it must also be that $\gamma(t)\geq e^{rt}(1/(1-\log\varepsilon)-\zeta)$ for all $t\leq T(\gamma)$, for otherwise $\min_{t\leq T(\gamma)}e^{-rt}\underline{\gamma}(t)$ would be strictly less than $1/(1-\log\varepsilon)-\zeta$.

[27] Technically, I must also show that $u_1^*(\gamma^*)\leq 1/(1-\log\varepsilon)$. In fact, $u_1^*(\gamma^*)=0$, by Lemma 3 and the observation that $T(\gamma^*)=\infty$ (which follows because $\gamma^*(t)=1$ for all $t\geq\tilde{T}(\gamma^*)$).

I will show that, for all $\delta > 0$, there exists $\zeta > 0$ such that, if both $\gamma(t) \geq e^{rt}(1/(1-\log\varepsilon) - \zeta)$ for all $t \leq T(\gamma)$ and $\tilde{T}(\gamma) \leq T^1 - (1/r)\log(1-\zeta(1-\log\varepsilon))$, then $\sup_{t\leq\tilde{T}(\gamma)} |v^*(t) - v(t)| \leq \delta$.

If $\tilde{T}(\gamma) \leq T^1 - (1/r)\log(1-\zeta(1-\log\varepsilon))$ then $\tilde{T}(\gamma)$ is finite, and therefore

$$\exp\left(-\int_0^{\tilde{T}(\gamma)} \frac{rv(t) - v'(t)}{1-v(t)}dt\right) \prod_{s\in S\cap\left[0,\tilde{T}(\gamma)\right]} \left(\frac{1-v(s,-1)}{1-v(s)}\right) \leq \varepsilon.$$

It is straightforward to check that $\gamma(t) \geq e^{rt}(1/(1-\log\varepsilon) - \zeta)$ for all $t \leq T(\gamma)$ only if $v(t) \leq 1 - e^{rt}(1/(1-\log\varepsilon) - \zeta)$ for all $t \leq T(\gamma)$. Recall that $e^{-rt}v(t)$ is non-increasing. Thus, if $\gamma(t) \geq e^{rt}(1/(1-\log\varepsilon) - \zeta)$ for all $t \leq T(\gamma)$ and $\tilde{T}(\gamma)$ is finite, then

$$\inf_{\substack{v(t):\\ e^{-rt}v(t)\text{ non-increasing,}\\ v(t)\leq 1-e^{rt}\left(\frac{1}{1-\log\varepsilon}-\zeta\right)}} \exp\left(-\int_0^{\tilde{T}(\gamma)} \frac{rv(t) - v'(t)}{1-v(t)}dt\right) \prod_{s\in S\cap\left[0,\tilde{T}(\gamma)\right]} \left(\frac{1-v(s,-1)}{1-v(s)}\right) \leq \varepsilon. \quad (18)$$

I first show that any attainable value of the program on the left-hand side of (18) can be arbitrarily closely approximated by the value attained by a *continuous* function $v(t)$ satisfying the constraints of (18); hence, in calculating the infimum over such values, attention may be restricted to continuous functions. To see this, fix $\eta \in (0,1)$ and let

$$S^\eta \equiv \bigcup_{s\in S\cap\left[0,\tilde{T}(\gamma)\right]} [s-\eta, s].$$

Define the function $v^\eta(t)$ by $v^\eta(t) \equiv v(t)$ for all $t \notin S^\eta$, and

$$v^\eta(t) \equiv v(s-\eta) - \frac{t-(s-\eta)}{\eta}(v(s-\eta) - v(s)) \text{ for all } t \in S^\eta.$$

Observe that $v^\eta(t)$ is continuous. Furthermore, for all $s \in S$,

$$\exp\left(\int_{s-\eta}^{s} \frac{v^{\eta\prime}(t)}{1-v^\eta(t)}dt\right) = \frac{1-v^\eta(s-\eta)}{1-v^\eta(s)} = \frac{1-v(s-\eta)}{1-v(s)}.$$

Also, since $v^\eta(t) \leq 1 - (1/(1-\log\varepsilon) - \zeta) < 1$ for all $t \in \left[0, \tilde{T}(\gamma)\right]$, and the measure of S^η goes to 0 as $\eta \to 0$,

$$\lim_{\eta\to 0} \exp\left(-\int_{S^\eta} \frac{rv^\eta(t)}{1-v^\eta(t)}dt\right) = 1.$$

Therefore,

$$
\begin{aligned}
& \lim_{\eta\to 0} \exp\left(-\int_0^{\tilde{T}(\gamma)} \frac{rv^\eta(t) - v^{\eta\prime}(t)}{1-v^\eta(t)}dt\right) \\
= \; & \lim_{\eta\to 0} \exp\left(-\int_{\left[0,\tilde{T}(\gamma)\right]\setminus S^\eta} \frac{rv^\eta(t) - v^{\eta\prime}(t)}{1-v^\eta(t)}dt\right) \exp\left(-\int_{S^\eta} \frac{rv^\eta(t)}{1-v^\eta(t)}dt\right) \prod_{s\in S\cap\left[0,\tilde{T}(\gamma)\right]} \left(\frac{1-v(s-\eta)}{1-v(s)}\right) \\
= \; & \exp\left(-\int_0^{\tilde{T}(\gamma)} \frac{rv(t) - v'(t)}{1-v(t)}dt\right) \prod_{s\in S\cap\left[0,\tilde{T}(\gamma)\right]} \left(\frac{1-v(s,-1)}{1-v(s)}\right).
\end{aligned}
$$

I now derive a lower bound on the left-hand side (18) under the additional constraint that $v(t)$ is continuous. Using the fact that $v(s,-1) = v(s)$ for all s when v is continuous and integrating the $v'(t) / (1-v(t))$ term, this constrained program may be rewritten as

$$
\inf_{\substack{v(t) \text{ continuous:} \\ e^{-rt}v(t) \text{ non-increasing,} \\ v(t)\leq 1-e^{rt}\left(\frac{1}{1-\log\varepsilon}-\zeta\right)}} \exp\left(-\int_0^{\tilde{T}(\gamma)} \frac{rv(t)}{1-v(t)}dt\right)\left(\frac{1-v(0)}{1-v\left(\tilde{T}(\gamma)\right)}\right).
$$

Since $v(t) \geq 0$ for all t, the value of this program is bounded from below by the value of the program:

$$
\inf_{\substack{v(t) \text{ continuous:} \\ e^{-rt}v(t) \text{ non-increasing,} \\ v(t)\leq 1-e^{rt}\left(\frac{1}{1-\log\varepsilon}-\zeta\right)}} \exp\left(-\int_0^{\tilde{T}(\gamma)} \frac{rv(t)}{1-v(t)}dt\right)(1-v(0)). \tag{19}
$$

Note that (19) decreases whenever the value of $v(t)$ is increased on a subset of $\left[0,\tilde{T}(\gamma)\right]$ of positive measure, so the unique solution to (19) is $v(t) = 1 - e^{rt}(1/(1-\log\varepsilon) - \zeta)$ for all $t \leq \tilde{T}(\gamma)$. With this function $v(t)$, it can be checked that the value of $\tilde{T}(\gamma)$ such that (19) equals ε is

$$
T^1 - \frac{1}{r}\log(1-\zeta(1-\log\varepsilon)\log\varepsilon). \tag{20}
$$

This value is a lower bound on $\tilde{T}(\gamma)$ for any posture γ such that $\gamma(t) \geq e^{rt}(1/(1-\log\varepsilon)-\zeta)$ for all $t \leq T(\gamma)$. Thus, as $\zeta \to 0$, the unique solution to (19) converges to $v^*(t) = 1 - e^{rt}/(1-\log\varepsilon)$ for all $t \leq \tilde{T}(\gamma)$, and the corresponding lower bound on $\tilde{T}(\gamma)$ (i.e., (20)) converges to T^1. Furthermore, by the condition that $e^{-rt}v(t)$ is non-increasing, any function $v(t)$ satisfying the constraints of (19) yields a lower bound on $\tilde{T}(\gamma)$ that is greater than (20) by at least an amount proportional to $\sup_{t\leq\tilde{T}(\gamma)} |v^*(t) - v(t)|$. Therefore, for any fixed $\delta > 0$, there exists $\zeta > 0$ such that if both $\gamma(t) \geq e^{rt}(1/(1-\log\varepsilon) - \zeta)$ for all $t \leq T(\gamma)$ and $\tilde{T}(\gamma) \leq T^1 - (1/r)\log(1-\zeta(1-\log\varepsilon))$ (which converges to T^1 as $\zeta \to 0$), then $\sup_{t \leq \tilde{T}(\gamma)} |v^*(t) - v(t)| \leq \delta$.

Thus, I have shown that, for any $\delta > 0$ and $K > 1$, there exists $\zeta(K) > 0$ such that if $u_1^*(\gamma) \geq 1/(1-\log\varepsilon) - \zeta(K)$, then $\sup_{t \leq \tilde{T}(\gamma)} e^{-rt}|v(t) - v^*(t)| \leq \delta/K$. I now argue that, for K sufficiently large, there exists $\zeta' \in (0, \zeta(K))$ such that if $u_1^*(\gamma) \geq 1/(1-\log\varepsilon) - \zeta'$, then in addition $\sup_{t > \tilde{T}(\gamma)} e^{-rt}|v(t) - v^*(t)| \leq \delta$. To see this, note that as $K \to \infty$, $\tilde{T}(\gamma) \to T^1$ uniformly over all postures γ such that $\sup_{t \leq \tilde{T}(\gamma)} e^{-rt}|v(t) - v^*(t)| \leq \delta/K$. Choose $K^* > 1$ such that $\left|e^{-r\tilde{T}(\gamma)} - e^{-rT^1}\right| < \delta/2$ and $v^*\left(\tilde{T}(\gamma)\right) \leq e^{r\tilde{T}(\gamma)}\delta$ for any such posture γ, and suppose that a posture γ is such that $\sup_{t \leq \tilde{T}(\gamma)} e^{-rt}|v(t) - v^*(t)| \leq \delta/K$ but $e^{-rt_0}|v(t_0) - v^*(t_0)| > \delta$ for some $t_0 > \tilde{T}(\gamma)$. Then $v^*(t_0) \leq e^{rt_0}\delta$, so it follows that $e^{-rt_0}v^*(t_0) + \delta < e^{-rt_0}v(t_0)$. Therefore,

$$\max_{t \geq \tilde{T}(\gamma)} e^{-rt}\left(1 - \underline{\gamma}(t)\right) \geq e^{-rt_0}v(t_0) \geq \delta.$$

By the definition of $T(\gamma)$, this implies that there exists $t_1 \in \left[\tilde{T}(\gamma), T(\gamma)\right]$ such that $e^{-rt_1}\left(1 - \underline{\gamma}(t_1)\right) \geq \delta$, or equivalently $\underline{\gamma}(t_1) \leq 1 - e^{rt_1}\delta$. Hence,

$$\begin{aligned} u_1^*(\gamma) &= \min_{t \leq T(\gamma)} e^{-rt}\underline{\gamma}(t) \leq e^{-rt_1}\left(1 - e^{rt_1}\delta\right) \\ &\leq e^{-r\tilde{T}(\gamma)}\left(1 - e^{r\tilde{T}(\gamma)}\delta\right) = e^{-r\tilde{T}(\gamma)} - \delta \\ &< e^{-rT^1} - \delta/2 = 1/(1-\log\varepsilon) - \delta/2. \end{aligned}$$

Therefore, taking $\zeta' \equiv \min\{\zeta(K^*), \delta/2\}$ completes the first step of the proof.

Step 2:

I show that if $\gamma_n(t) \to \gamma(t)$ for all $t \in \mathbb{R}_+$ for some posture γ, and $\sup_{t \in \mathbb{R}_+} e^{-rt}|v^*(t) - v_n(t)| \to 0$, then $\gamma = \gamma^*$. First, note that if $\gamma(t) < \gamma^*(t)$ for some $t \in \mathbb{R}_+$, then there exist $N > 0$ and $\eta > 0$ such that $\gamma_n(t) < \gamma^*(t) - \eta$ for all $n > N$. Since $v_n(t) \geq 1 - \gamma_n(t)$, this implies that $v_n(t) \geq 1 - \gamma^*(t) + \eta = v^*(t) + \eta$ for all $n > N$, a contradiction.

It is more difficult to rule out the possibility that $\gamma(t) > \gamma^*(t)$ for some $t \in \mathbb{R}_+$. Suppose that this is so. Since γ and γ^* are right-continuous, there exist $\eta > 0$ and an open interval $I_0 \subseteq \mathbb{R}_+$ such that $\gamma(t) > \gamma^*(t) + \eta$ for all $t \in I_0$. If it were the case that $\gamma_n(t) \geq \gamma^*(t) + \eta/2$ for all $t \in I_0$ and n sufficiently large, then the condition $\sup_{t \in \mathbb{R}_+} e^{-rt}|v^*(t) - v_n(t)| \to 0$

would fail, so this is not possible.[28] Hence, there exists $t_1 \in I_0$ and $n_1 \geq 0$ such that $\gamma_{n_1}(t_1) < \gamma^*(t_1) + \eta/2$. Since γ_{n_1} and γ^* are right-continuous, there exists an open interval $I_1 \subseteq I_0$ such that $\gamma_{n_1}(t) < \gamma^*(t) + \eta/2$ for all $t \in I_1$. Next, it cannot be the case that $\gamma_n(t) \geq \gamma^*(t) + \eta/2$ for all $t \in I_1$ and $n > n_1$ (by the same argument as above), so there exists $t_2 \in I_1$ and $n_2 > n_1$ such that $\gamma_{n_2}(t_2) < \gamma^*(t_2) + \eta/2$. As above, this implies that there exists an open interval $I_2 \subseteq I_1$ such that $\gamma_{n_2}(t) < \gamma^*(t) + \eta/2$ for all $t \in I_2$. Proceeding in this manner yields a sequence of open intervals $\{I_m\}$ and integers $\{n_m\}$ such that $I_{m+1} \subseteq I_m$, $n_{m+1} > n_m$, and $\gamma_{n_m}(t) < \gamma^*(t) + \eta/2$ for all $t \in I_m$ and $m \in \mathbb{N}$. Let $I \equiv \bigcap_{m \in \mathbb{N}} I_m$, a non-empty set (possibly a single point), and fix $t \in I$. Then $\gamma_{n_m}(t) < \gamma^*(t) + \eta/2$ for all $m \in \mathbb{N}$, and since $n_{m+1} > n_m$ for all $m \in \mathbb{N}$ this contradicts the assumption that $\gamma_n(t) \to \gamma(t)$. ■

References

[1] Abreu, D., and Gul, F. (2000), "Bargaining and Reputation," *Econometrica*, 2000, 85-117.

[2] Abreu, D., and Pearce, D. (2007), "Bargaining, Reputation, and Equilibrium Selection in Repeated Games with Contracts," *Econometrica*, 75, 653-710.

[3] Battigalli, P., and Watson, J. (1997), "On 'Reputation' Refinements with Heterogeneous Beliefs," *Econometrica*, 65, 369-374.

[4] Battigalli, P., and Siniscalchi, M. (2003), "Rationalization and Incomplete Information," *Advances in Theoretical Economics*, 3, Iss. 1, Article 3.

[5] Bergin, J., and MacLeod, W.B. (1993), "Continuous Time Repeated Games," *International Economic Review*, 34, 21-37.

[28] To see this, fix $N > 0$, suppose that $\gamma_n(t) \geq \gamma^*(t) + \eta/2$ for all $t \in I_0$ and $n > N$, and denote the length of I_0 by 2Δ and the midpoint of I_0 by t_0. Noting that $\gamma^*(t_0) < 1 - \eta/2$, there exists $N' > 0$ such that $\gamma_n(t) \geq 1 - v_n(t) > 1 - v^*(t) - e^{rt}\left(1 - e^{-r\Delta}\right)(1 - \gamma^*(t_0) - \eta/2) = \gamma^*(t) - e^{rt}\left(1 - e^{-r\Delta}\right)(1 - \gamma^*(t_0) - \eta/2)$ for all $t \in \mathbb{R}_+$ and $n > N'$. Therefore, $v_n(t_0) \leq \max\left\{1 - \gamma^*(t_0) - \eta/2, e^{-r\Delta}(1 - \gamma^*(t_0 + \Delta)) + \left(1 - e^{-r\Delta}\right)(1 - \gamma^*(t_0) - \eta/2)\right\} \leq \max\left\{1 - \gamma^*(t_0) - \eta/2, 1 - \gamma^*(t_0) - \left(1 - e^{-r\Delta}\right)\eta/2\right\} = v^*(t_0) - \left(1 - e^{-r\Delta}\right)\eta/2$ for all $n > \max\{N, N'\}$, a contradiction.

[6] Billingsley, P. (1995), *Probability and Measure*. New York, NY: John Wiley & Sons, Inc..

[7] Chatterjee, K., and Samuelson, L. (1987), "Bargaining with Two-Sided Incomplete Information: An Infinite-Horizon Model with Alternating Offers," *Review of Economic Studies*, 59, 175-192.

[8] Chatterjee, K., and Samuelson, L. (1988), "Bargaining Under Two-Sided Incomplete Information: the Unrestricted Offers Case," *Operations Research*, 36, 605-618.

[9] Cho, I.-K. (1994), "Stationarity, Rationalizability, and Bargaining," *Review of Economic Studies*, 61, 357-374.

[10] Compte, O., and Jehiel, P. (2002), "On the Role of Outside Options in Bargaining with Obstinate Parties," *Econometrica*, 70, 1477-1517.

[11] Crawford, V. (1982), "A Theory of Disagreement in Bargaining," *Econometrica*, 50, 607-637.

[12] Cripps, M.W., Schmidt, K., and Thomas, J.P. (1996), "Reputation in Perturbed Repeated Games," *Journal of Economic Theory*, 69, 387-410.

[13] Ellingsen, T., and Miettinen, T. (2008), "Commitment and Conflict in Bilateral Bargaining," *American Economic Review*, 98, 1629–1635.

[14] Evans, R., and Thomas, J.P. (1997), "Reputation and Experimentation in Repeated Games with Two Long-Run Players," *Econometrica*, 65, 1153-1173.

[15] Feinberg, Y., and Skrzypacz, A. (2005), "Uncertainty about Uncertainty and Delay in Bargaining," *Econometrica*, 73, 69-91.

[16] Fershtman, C., and Seidmann, D.J. (1993), "Deadline Effects and Inefficient Delay in Bargaining with Endogenous Commitment," *Journal of Economic Theory*, 60, 306-321.

[17] Fudenberg, D., and Levine, D.K. (1989), "Reputation and Equilibrium Selection in Games with a Patient Player," *Econometrica*, 57, 759-778.

[18] Green, J.R., and Scotchmer, S. (1995), "On the Division of Profit in Sequential Innovation," *RAND Journal of Economics*, 26, 20-33.

[19] Grossman, S.J. and Hart, O.D. (1986), "The Costs and Benefits of Ownership: A Theory of Vertical and Lateral Integration," *Journal of Political Economy*, 94, 691-719.

[20] Hosios, A. (1990), "On the Efficiency of Matching and Related Models of Search and Unemployment," *Review of Economic Studies*, 57, 279-298.

[21] Kambe, S. (1999), "Bargaining with Imperfect Commitment," *Games and Economic Behavior*, 28, 217-237.

[22] Kreps, D.M. and Wilson, R. (1982), "Reputation and Imperfect Information," *Journal of Economic Theory*, 27, 253-279.

[23] Knoll, M.S. (1996), "A Primer on Prejudgment Interest," *Texas Law Review*, 75, 293-374.

[24] Milgrom, P., and Roberts, J. (1982), "Predation, Reputation, and Entry Deterrence," *Journal of Economic Theory*, 27, 280-312.

[25] Muthoo, A. (1996), "A Bargaining Model Based on the Commitment Tactic," *Journal of Economic Theory*, 69, 134-152.

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

[27] Pearce, D. (1984), "Rationalizable Strategic Behavior and the Problem of Perfection," *Econometrica*, 52, 1029-1050.

[28] Royden, H.L. (1988), *Real Analysis*, Third Edition. Upper Saddle River, NJ: Prentice Hall.

[29] Rubinstein, A. (1982), "Perfect Equilibrium in a Bargaining Model," *Econometrica*, 50, 97-109.

[30] Schelling, T.C. (1956), "An Essay on Bargaining," *American Economic Review*, 46, 281-306.

[31] Schmidt, K. (1993), "Reputation and Equilibrium Characterization in Repeated Games of Conflicting Interests," *Econometrica*, 61, 325-351.

[32] Simon, L.K., and Stinchcombe, M.B. (1989), "Extensive Form Games in Continuous time: Pure Strategies," *Econometrica*, 57, 1171-1214.

[33] Staw, B.M. (1981), "The Escalation of Commitment to a Course of Action," *Academy of Management Review*, 6, 577-587.

[34] Staw, B.M. (1997), "The Escalation of Commitment: An Update and Appraisal," in Shapira, Z., ed., *Organizational Decision Making*, 191-215. Cambridge: Cambridge University Press.

[35] Watson, J. (1993), "A 'Reputation' Refinement without Equilibrium," *Econometrica*, 61, 199-205.

[36] Watson, J. (1998), "Alternating-Offer Bargaining with Two-Sided Incomplete Information," *Review of Economic Studies*, 65, 573-594.

[37] Wolitzky, A. (2011), "Indeterminacy of Reputation Effects in Repeated Games with Contracts," *Forthcoming, Games and Economic Behavior.*

[38] Yildiz, M. (2003), "Bargaining without a Common Prior: An Immediate Agreement Theorem," *Econometrica*, 71, 793-811.

[39] Yildiz, M. (2004), "Waiting to Persuade," *Quarterly Journal of Economics*, 119, 223-248.

Supplementary Appendix: Omitted Proofs for Sections 5 and 7 (not for Publication)

Proof of Proposition 2. Lemmas 1 through 3 apply to any posture, whether or not it is constant. In addition, if γ is constant then $T(\gamma) = \tilde{T}(\gamma)$. Thus, Lemma 3 implies that $u_1^*(\gamma) = \min_{t \le T(\gamma)} e^{-rt}\gamma = e^{-r\tilde{T}(\gamma)}\gamma$. Furthermore, $\lambda(t) = r(1-\gamma)/\gamma$ and $p(t) = 0$ for all t, so, by the definition of $\tilde{T}(\gamma)$,

$$\exp\left(-r\left(\frac{1-\gamma}{\gamma}\right)\tilde{T}(\gamma)\right) = \varepsilon,$$

or

$$\tilde{T}(\gamma) = -\frac{1}{r}\left(\frac{\gamma}{1-\gamma}\right)\log\varepsilon$$

if $\gamma < 1$, and $\tilde{T}(\gamma) = \infty$ if $\gamma = 1$. Therefore,

$$\begin{aligned} \bar{u}_1^* &= \max_{\gamma\in[0,1]} e^{-r\tilde{T}(\gamma)}\gamma \\ &= \max_{\gamma\in[0,1)} \exp\left(\frac{\gamma}{1-\gamma}\log\varepsilon\right)\gamma. \end{aligned} \tag{21}$$

Note that (21) is concave in γ. Hence, the first-order condition

$$1 = -\frac{\bar{\gamma}_\varepsilon^*}{(1-\bar{\gamma}_\varepsilon^*)^2}\log\varepsilon, \tag{22}$$

which has a solution if $\varepsilon < 1$, is both necessary and sufficient. Solving this quadratic equation yields

$$\bar{\gamma}_\varepsilon^* = \frac{2 - \log\varepsilon - \sqrt{(\log\varepsilon)^2 - 4\log\varepsilon}}{2}.$$

Finally, substituting (22) into (21) yields $\bar{u}_1^* = \exp(-(1-\bar{\gamma}_\varepsilon^*))\bar{\gamma}_\varepsilon^*$. ■

Proof of Corollary 2. By (22), $-\bar{\gamma}_\varepsilon^*\log\varepsilon = (1-\bar{\gamma}_\varepsilon^*)^2$ for all ε. Therefore,

$$\begin{aligned} \frac{u_1^*(\varepsilon)}{\bar{u}_1^*(\varepsilon)} &= \frac{1}{(1-\log\varepsilon)\exp(-(1-\bar{\gamma}_\varepsilon^*))\bar{\gamma}_\varepsilon^*} \\ &= \frac{\exp(1-\bar{\gamma}_\varepsilon^*)}{\bar{\gamma}_\varepsilon^* + (1-\bar{\gamma}_\varepsilon^*)^2}. \end{aligned} \tag{23}$$

The derivative of (23) with respect to $\bar{\gamma}_\varepsilon^*$ is negative for all $\bar{\gamma}_\varepsilon^* \in [0,1]$. Since $\bar{\gamma}_\varepsilon^*$ is an increasing function of ε, this implies that $u_1^*(\varepsilon)/\bar{u}_1^*(\varepsilon)$ is decreasing in ε. In addition, by

(22), $\lim_{\varepsilon\to 1}\bar{\gamma}^*_\varepsilon = 1$ and $\lim_{\varepsilon\to 0}\bar{\gamma}^*_\varepsilon = 0$. Therefore, (23) implies that $\lim_{\varepsilon\to 1} u_1^*(\varepsilon)/\bar{u}_1^*(\varepsilon) = 1$ and $\lim_{\varepsilon\to 0} u_1^*(\varepsilon)/\bar{u}_1^*(\varepsilon) = e$. ■

Proof of Proposition 3. The proof of Theorem 1 goes through for any decreasing Pareto frontier ϕ, with the modifications that $v(t,-1)$ now equals $\max_{\tau\geq t} e^{-r(\tau-t)}\phi\left(\underline{\gamma}(\tau)\right)$ rather than $\max_{\tau\geq t} e^{-r(\tau-t)}\left(1-\underline{\gamma}(t)\right)$ (with the analogous modification for $v(t)$), and that the maxmin posture is now given by $\gamma^*(t) = \min\{e^{rt}u_1^*, 1\}$ for a value of u_1^* that may differ from $1/(1-\log\varepsilon)$. Therefore, it suffices to show that, for any $u_1\in[0,1]$, the posture $\gamma(t) = \min\{e^{rt}u_1, 1\}$ corresponds to a weakly higher concession rate $\lambda(t)$, for all t, when the Pareto frontier is $\psi\circ\phi$ than when it is ϕ. Note that the set of times at which $\lambda(t)$ is given by (3) is the same for either Pareto frontier, because $\psi\circ\phi(u) = 1$ if and only if $\phi(u) = 1$. Hence, since ϕ and ψ are concave and thus differentiable almost everywhere, it suffices to show that

$$\frac{r\psi\left(\phi\left(e^{rt}u_1\right)\right) - \psi'\left(\phi\left(e^{rt}u_1\right)\right)\phi'\left(e^{rt}u_1\right)re^{rt}u_1}{1-\psi\left(\phi\left(e^{rt}u_1\right)\right)} \geq \frac{r\phi\left(e^{rt}u_1\right) - \phi'\left(e^{rt}u_1\right)re^{rt}u_1}{1-\phi\left(e^{rt}u_1\right)}$$

for all $t\in\mathbb{R}_+$ and $u_1\in[0,1]$; or, dividing both sides by r and writing u for $e^{rt}u_1$,

$$\frac{\psi\left(\phi(u)\right) - \psi'\left(\phi(u)\right)\phi'(u)\,u}{1-\psi\left(\phi(u)\right)} \geq \frac{\phi(u) - \phi'(u)\,u}{1-\phi(u)}$$

for all $u\in[0,1]$. This inequality may be rearranged as

$$\psi\left(\phi(u)\right) - \phi(u) - \left(\left(1-\phi(u)\right)\psi'\left(\phi(u)\right) - \left(1-\psi\left(\phi(u)\right)\right)\right)\phi'(u)\,u \geq 0. \tag{24}$$

The maintained assumptions on ψ imply that $\psi(x)\geq x$ and $(1-x)\psi'(x)\geq 1-\psi(x)$ for all $x\in[0,1]$, so $\psi(\phi(u)) - \phi(u)\geq 0$ and $(1-\phi(u))\psi'(\phi(u)) - (1-\psi(\phi(u)))\geq 0$. Since $\phi'(u)\leq 0$ and $u\geq 0$, it follows that (24) holds. ■

Proof of Proposition 4. Lemmas 1 through 3 continue to hold, replacing r with r_1 or r_2 as appropriate. In particular, $\lambda(t) = \frac{r_2 v(t) - v'(t)}{1-v(t)}$; and the same argument as in the proof of Theorem 1 implies that the unique maxmin posture γ^* satisfies $\gamma^*(t) = \min\{e^{r_1 t}u_1^*, 1\}$, where u_1^* is the (unique) number such that the time at which $\gamma^*(t)$ reaches 1 equals $\tilde{T}(\gamma^*)$. Thus, given posture γ^*, it follows that $\lambda(t) = \frac{r_2\left(1-e^{r_1 t}u_1^*\right) + r_1 e^{r_1 t}u_1^*}{e^{r_1 t}u_1^*} = r_2\frac{e^{-r_1 t}}{u_1^*} + r_1 - r_2$. Now

$$\exp\left(-\int_0^{\tilde{T}(\gamma^*)}\left(r_2\frac{e^{-r_1 t}}{u_1^*} + r_1 - r_2\right)dt\right) = \exp\left(-\frac{1}{u_1^*}\left(\frac{r_2}{r_1}\right)\left(1-e^{-r_1\tilde{T}(\gamma^*)}\right) + (r_1-r_2)\,\tilde{T}(\gamma^*)\right).$$

Setting this equal to ε and rearranging implies that $\tilde{T}(\gamma^*)$ is given by

$$e^{-r_1\tilde{T}(\gamma^*)} - \frac{r_1}{r_2}u_1^* \log\varepsilon + \left(\frac{r_1}{r_2} - 1\right) u_1^* r_1 \tilde{T}(\gamma^*) = 1. \tag{25}$$

Using the condition that $e^{r_1\tilde{T}(\gamma^*)}u_1^* = 1$, this can be rearranged to yield (11). Finally, there is a unique pair $\left(u_1^*, \tilde{T}(\gamma^*)\right)$ that satisfies both (25) and $e^{r_1\tilde{T}(\gamma^*)}u_1^* = 1$, because the curve in $\left(u_1^*, \tilde{T}(\gamma^*)\right)$ space defined by (25) is upward-sloping, while the curve defined by $e^{r_1\tilde{T}(\gamma^*)}u_1^* = 1$ is downward-sloping. ■

Proof of Corollary 3. As $r_1/r_2 \to 0$, (11) becomes $u_1^*(1 + \log u_1^*) = 1$, which has unique solution $u_1^* = 1$. Therefore, $\lim_{r_1/r_2\to 0} u_1^*(\varepsilon) = 1$.

Suppose that $\varepsilon < 1$. For any sequence of relative discount rates $\{r_1/r_2\}_n$, the sequence of corresponding values of $u_1^*(\varepsilon)$ has a convergent subsequence. Suppose that $\{r_1/r_2\}_m \to \infty$ and the corresponding values of $u_1^*(\varepsilon)$ converge to some u_1^*. Then (11) becomes $u_1^* = 0$, because $\varepsilon < 1$ and $u_1^* \leq 1$. Therefore, $\lim_{r_1/r_2\to\infty} u_1^*(\varepsilon) = 0$. ■

Proof of Corollary 4. After such a decrease in r_1/r_2 and ε, the right-hand side of (11) increases if u_1^* is held constant. If $r_1/r_2 \leq 1$, then the left-hand side of (11) is increasing in u_1^* and the right-hand side of (11) is non-increasing in u_1^*. Therefore, if $r_1/r_2 \leq 1$, such a decrease in r_1/r_2 and ε leads to an increase in $u_1^*(\varepsilon)$. ■

Proof of Lemma 4. The fact that $\Omega_2^{RAT}(\gamma) \subseteq \Pi_1^\gamma$ immediately implies that $u_1^{RAT}(\gamma) \geq u_1^*(\gamma) = \min_{t\leq T(\gamma)} e^{-rt}\underline{\gamma}(t)$. Therefore, it suffices to show that $u_1^{RAT}(\gamma) \leq \min_{t\leq T(\gamma)} e^{-rt}\underline{\gamma}(t)$.

Let $\dot{T} \equiv \min \operatorname{argmax}_t e^{-rt}\left(1 - \underline{\gamma}(t)\right)$. Note that $\dot{T}$ is well-defined and finite because $\underline{\gamma}(t)$ is lower semi-continuous and $\lim_{t\to\infty} e^{-rt}\left(1 - \underline{\gamma}(t)\right) = 0$. In addition, $v(t) = e^{-r\left(\dot{T}-t\right)}\left(1 - \underline{\gamma}\left(\dot{T}\right)\right)$ for all $t < \dot{T}$, which implies that $\lambda(t) = p(t) = 0$ for all $t < \dot{T}$. Hence, the mixed strategy π_2^γ coincides with γ for all $t < \dot{T}$.

Let $\dot{\sigma}_2 \in \Sigma_2$ be identical to the γ-offsetting strategy σ_2^γ with the exception that player 2 accepts at date $\left(\dot{T}, -1\right)$ if player 1 follows γ until time $\dot{T}$. Then, under strategy $\dot{\sigma}_2$, player 2 always demands $u_2(t) = 1$ and only accepts player 1's demand if player 1 follows γ until time $\dot{T}$. Since the mixed strategy π_2^γ coincides with γ for all $t \leq \dot{T}$, it follows that $\sup_{\sigma_1} u_1(\sigma_1, \dot{\sigma}_2) = e^{-r\dot{T}}\underline{\gamma}\left(\dot{T}\right) = u_1(\pi_2^\gamma, \dot{\sigma}_2)$, and therefore $\pi_2^\gamma \in \Sigma_1^*(\dot{\sigma}_2)$. In addition, it is clear that $\dot{\sigma}_2 \in \Sigma_2^*(\gamma)$, and furthermore $\gamma \in \Sigma_1^*(\sigma_2^\gamma)$ (by Lemma 3), and $\sigma_2^\gamma \in \Sigma_2^*(\pi_2^\gamma)$ (by Lemma 2). Summarizing, I have established that the arrows in the following diagram may be read as “is a best-response to” :

$$\begin{array}{ccc} \gamma & \rightarrow & \sigma_2^\gamma \\ \uparrow & & \downarrow \\ \dot{\sigma}_2 & \leftarrow & \pi_2^\gamma \end{array}.$$

Therefore, the set $\{\gamma, \pi_2^\gamma\} \times \{\sigma_2^\gamma, \dot{\sigma}_2\}$ has the best-response property given posture γ, which implies that $\{\gamma, \pi_2^\gamma\} \times \{\sigma_2^\gamma, \dot{\sigma}_2\} \subseteq \Omega_2^{RAT}(\gamma)$. Hence, $u_1^{RAT}(\gamma) \leq \sup_{\sigma_1} u_1(\sigma_1, \sigma_2^\gamma) = u_1(\gamma, \sigma_2^\gamma) = \min_{t \leq T(\gamma)} e^{-rt} \underline{\gamma}(t)$. ■

Proof of Theorem 3. Observe that a posture γ in discrete-time bargaining game g induces a “continuous-time posture” $\hat{\gamma}$ (i.e., a map from $\mathbb{R}_+ \rightarrow [0,1]$) according to $\hat{\gamma}(t) = \gamma(\min\{\tau \geq t : \tau \in I_i^g\})$. That is, $\hat{\gamma}$'s time-t demand is simply γ's next demand in g. I henceforth refer to a posture γ in g as also being a continuous-time posture, with the understanding that I mean the posture $\hat{\gamma}$ defined above.

However, γ may *not* be a posture in the continuous-time bargaining game of Section 3, because it may be discontinuous at a non-integer time. To avoid this problem, I now introduce a modified version of the continuous-time bargaining game of Section 3. Formally, let the continuous-time bargaining game g^{cts} be defined as in Section 3, with the following modifications: Most importantly, omit the requirement that player i's demand path $u_i^t : [t, t+1) \rightarrow [0,1]$ (which is still chosen at integer times t) is continuous. Second, specify that the payoffs if player i accepts player j's offer at date $(t, -1)$ are $(e^{-rt}(1 - \liminf_{\tau \uparrow t} u_j(\tau)), e^{-rt} \liminf_{\tau \uparrow t} u_j(\tau))$ (because $\lim_{\tau \uparrow t} u_j(\tau)$ may now fail to exist). Third, add a fourth date, $(t, 2)$ to each instant of time t. At date $(t, 2)$, each player i announces *accept* or *reject*, and, if player i accepts player j's offer at date $(t, 2)$, the game ends with payoffs $(e^{-rt}(1 - \liminf_{\tau \downarrow t} u_j(\tau)), e^{-rt} \liminf_{\tau \downarrow t} u_j(\tau))$. Adding the date $(t, 2)$ ensures that each player has a well-defined best-response to her belief, even though $u_j(t)$ may now fail to be right-continuous. One can check that the analysis of Sections 3 and 4, including Lemmas 1 through 3 and Theorem 1, continue to apply to the game g^{cts}, with the exception that in g^{cts} the maxmin posture γ^* is not in fact unique; however, every maxmin posture corresponds to the continuation value function $v^*(t)$ (by the same argument as in Step 1 of the proof of Theorem 1).[29] Because of this, for the remainder of the proof I

[29] The reason I did not use the game g^{cts} in Sections 3 and 4 is that it is difficult to interpret the assumption

slightly abuse notation by writing $u_1^*(\gamma)$ for player 1's maxmin payoff given posture γ *in the game* g^{cts}, rather than in the model of Section 3. Importantly, $u_1^*(\gamma)$ equals player 1's maxmin payoff given γ in both g^{cts} and in the model of Section 3 when γ is a posture in the model of Section 3, but $u_1^*(\gamma)$ is well-defined for all $\gamma : \mathbb{R}_+ \to [0,1]$. Similarly, I write $u_1^*(v)$ for player 1's maxmin payoff given continuation value function $v : \mathbb{R}_+ \to [0,1]$. This is well-defined because $u_1^*(\gamma) = \min_{t \le T} e^{-rt}\underline{\gamma}(t)$ by Lemma 3, T depends on γ only through v (by Lemma 1), and it can be easily verified that $\min_{t\le T} e^{-rt}\underline{\gamma}(t) = \min_{t \le T} e^{-rt}(1 - v(t))$ (and thus depends on γ only through v). A similar argument, which I omit, implies that one may write $u_1^{*,g_n}(v^{g_n})$ for player 1's maxmin payoff given continuation value function v^{g_n} in discrete-time bargaining game g_n.

With this notation, I may state the following lemma, from which Theorem 3 follows:

Lemma 5 *Let $\{g_n\}$ be a sequence of discrete-time bargaining games converging to continuous time. There exists a sequence of postures $\{\gamma^{g_n\prime}\}$ with $\gamma^{g_n\prime}$ a posture in g_n and $\lim_{n\to\infty} u_1^{*,g_n}(\gamma^{g_n\prime}) \ge u_1^*$. In addition, for any sequence of functions $\{v^{g_n}\}$ such that v^{g_n} is a continuation value function in g_n and $\lim_{n\to\infty} v^{g_n}(t)$ exists for all $t \in \mathbb{R}_+$, it follows that $\lim_{n\to\infty} u_1^{*,g_n}(v^{g_n})$ exists and equals $\lim_{n\to\infty} u_1^*(v^{g_n})$.*

Proof. I first introduce some additional notation. Let Σ_i^g be the set of player i's strategies in g^{cts} with the property that player i's demand only changes at times $t \in I_i^g$, player i only accepts player j's offer at times $t \in I_j^g$, and player i's action at time t only depends on past play at times $\tau \in I_i^g \cup I_j^g$. One can equivalently view Σ_i^g as player i's strategy set in g itself. Thus, any belief π_2 in g may also be viewed as a belief in g^{cts} (with $\mathrm{supp}(\pi_2) \subseteq \Sigma_i^g$).

Let $\gamma^{g_n\prime}$ be given by $\gamma^{g_n\prime}(t) = (n/(n+1))\gamma^*(\max\{\tau \le t : \tau \in I_1^{g_n}\})$ for all $t \in \mathbb{R}_+$, with the convention that $\max\{\tau < t : \tau \in I_1^{g_n}\} \equiv 0$ if the set $\{\tau < t : \tau \in I_1^{g_n}\}$ is empty. I first claim that $\lim_{n\to\infty} u_1^*(\gamma^{g_n\prime}) \ge u_1^*$.[30] To show this, I first establish that $\tilde{T}(\gamma^{g_n\prime}) \le \min\{\tau > T^1 : \tau \in I_1^{g_n}\}$ for all n, where T^1 is defined as in the proof of Theorem 1. Since

that player i can accept the demand $\liminf_{\tau \downarrow t} u_j(\tau)$ at time t, since the demand $u_j(\tau)$ has not yet been made at time t for all $\tau > t$. Thus, I view the game g^{cts} as a technical construct for analyzing the limit of discrete-time games, and not as an appealing model of continuous-time bargaining in its own right.

[30] Theorem 1 implies that $\lim_{n\to\infty} u_1^*(\gamma^{g_n\prime}) \le u_1^*$, so this inequality must hold with equality. But only the inequality is needed for the proof.

γ^* (and thus $\gamma^{g_n\prime}$) are non-decreasing, $\sup_{\tau \geq t} e^{-r(\tau - t)} \left(1 - \gamma^{g_n\prime}(\tau)\right) = 1 - \gamma^{g_n\prime}(t)$. Therefore, by Lemma 1, $\tilde{T}\left(\gamma^{g_n\prime}\right)$ satisfies

$$\exp\left(-\int_0^{\tilde{T}(\gamma^{g_n\prime})} \frac{r\left(\frac{n+1}{n} - \gamma^*\left(\max\left\{\tau \leq t : \tau \in I_1^{g_n}\right\}\right)\right)}{\gamma^*\left(\max\left\{\tau \leq t : \tau \in I_1^{g_n}\right\}\right)} dt\right) \prod_{t \in I_1^{g_n} \cap \left[0, \tilde{T}(\gamma^{g_n\prime})\right)} \frac{\gamma^*\left(\max\left\{\tau < t : \tau \in I_1^{g_n}\right\}\right)}{\gamma^*(t)} \geq \varepsilon. \tag{26}$$

Now

$$\begin{aligned}
& \exp\left(-\int_0^{\tilde{T}(\gamma^{g_n\prime})} \frac{r\left(\frac{n+1}{n} - \gamma^*\left(\max\left\{\tau \leq t : \tau \in I_1^{g_n}\right\}\right)\right)}{\gamma^*\left(\max\left\{\tau \leq t : \tau \in I_1^{g_n}\right\}\right)} dt\right) \prod_{t \in I_1^{g_n} \cap \left[0, \tilde{T}(\gamma^{g_n\prime})\right)} \frac{\gamma^*\left(\max\left\{\tau < t : \tau \in I_1^{g_n}\right\}\right)}{\gamma^*(t)} \\
\leq\ & \exp\left(-\int_0^{\tilde{T}(\gamma^{g_n\prime})} \frac{r\left(1 - \gamma^*(t)\right)}{\gamma^*(t)} dt\right) \frac{\gamma^*(0)}{\gamma^*\left(\max\left\{\tau < \tilde{T}\left(\gamma^{g_n\prime}\right) : \tau \in I_1^{g_n}\right\}\right)} \\
\leq\ & \exp\left(-\int_0^{\max\left\{\tau < \tilde{T}(\gamma^{g_n\prime}) : \tau \in I_1^{g_n}\right\}} \frac{r\left(1 - \gamma^*(t)\right) + \gamma^{*\prime}(t)}{\gamma^*(t)}\right).
\end{aligned} \tag{27}$$

If $\tilde{T}\left(\gamma^{g_n\prime}\right) > \min\left\{\tau > T^1 : \tau \in I_1^{g_n}\right\}$, then $\max\left\{\tau < \tilde{T}\left(\gamma^{g_n\prime}\right) : \tau \in I_1^{g_n}\right\} > T^1$ and therefore (27) is less than ε, which contradicts (26). Hence, $\tilde{T}\left(\gamma^{g_n\prime}\right) \leq \min\left\{\tau > T^1 : \tau \in I_1^{g_n}\right\}$ for all n. In addition, $\gamma^{g_n\prime}(t)$ is non-decreasing and $\gamma^{g_n\prime}(t) < 1$ for all t, which implies that $T\left(\gamma^{g_n\prime}\right) = \tilde{T}\left(\gamma^{g_n\prime}\right)$. Hence, by Lemma 3, $u_1^*\left(\gamma^{g_n\prime}\right) = \min_{t \leq \tilde{T}(\gamma^{g_n\prime})} e^{-rt} \gamma^{g_n\prime}(t)$. Since $\tilde{T}\left(\gamma^{g_n\prime}\right) \leq \min\left\{\tau > T^1 : \tau \in I_1^{g_n}\right\}$ for all n, and $\{g_n\}$ converges to continuous time, $\lim_{n\to\infty} \tilde{T}\left(\gamma^{g_n\prime}\right) \leq T^1$. In addition, $\lim_{n\to\infty} \sup_{t \in \mathbb{R}_+} \left|\gamma^{g_n\prime}(t) - \gamma^*(t)\right| = 0$, so it follows that $\lim_{n\to\infty} u_1^*\left(\gamma^{g_n\prime}\right) \geq \min_{t \leq T^1} e^{-rt}\gamma^*(t) = u_1^*$.

Next, I claim that $u_1^{*,g_n}\left(\gamma^{g_n}\right) \geq u_1^*\left(\gamma^{g_n}\right)$ for *any* posture γ^{g_n} in discrete-time bargaining game g^n. To see this, note that if $\operatorname{supp}(\pi_2) \subseteq \Sigma_1^{g_n}$ and $\sigma_2 \in \Sigma_2^{*,g_n}(\pi_2)$, then $\sigma_2 \in \Sigma_2^*(\pi_2)$ as well (i.e., there is no benefit to responding to a strategy in $\Delta\left(\Sigma_1^{g_n}\right)$ with a strategy outside of $\Sigma_2^{g_n}$). Therefore, if $\pi_1 \in \Pi_1^{\gamma^{g_n}, g_n}$ (i.e., if π_1 is consistent with knowledge of rationality in g_n), then $\pi_1 \in \Pi_1^{\gamma^{g_n}, g^{cts}}$; that is, $\Pi_1^{\gamma^{g_n}, g_n} \subseteq \Pi_1^{\gamma^{g_n}, g^{cts}}$. Now

$$\begin{aligned}
u_1^{*,g_n}\left(\gamma^{g_n}\right) &= \sup_{\sigma_1 \in \Sigma_1^{g_n}} \inf_{\pi_1 \in \Pi_1^{\gamma^{g_n}, g_n}} u_1(\sigma_1, \pi_1) \\
&\geq \sup_{\sigma_1 \in \Sigma_1^{g_n}} \inf_{\pi_1 \in \Pi_1^{\gamma^{g_n}, g^{cts}}} u_1(\sigma_1, \pi_1) \\
&= u_1\left(\gamma^{g_n}, \sigma^{\gamma^{g_n}}\right) \\
&= u_1^*\left(\gamma^{g_n}\right),
\end{aligned}$$

where $\sigma^{\gamma^{g_n}}$ is defined in Definition 6, and the second line follows because $\Pi_1^{\gamma^{g_n}, g_n} \subseteq \Pi_1^{\gamma^{g_n}, g^{cts}}$; the third line follows because $u_1\left(\gamma^{g_n}, \sigma^{\gamma^{g_n}}\right) = \sup_{\sigma_1 \in \Sigma_1^{g^{cts}}} \inf_{\pi_1 \in \Pi_1^{\gamma^{g_n}, g^{cts}}} u_1(\sigma_1, \pi_1)$ by Lemma 3, and $\gamma^{g_n} \in \Sigma_1^{g_n} \subseteq \Sigma_1^{g^{cts}}$; and the fourth line follows by Lemma 3.

Combining the above claims, it follows that $\lim_{n\to\infty} u_1^{*,g_n}\left(\gamma^{g_n\prime}\right) \geq \lim_{n\to\infty} u_1^*\left(\gamma^{g_n\prime}\right) \geq u_1^*(\gamma)$. This proves the first part of the lemma.

For the second part of the lemma, fix a sequence of continuation value functions $\{v^{g_n}\}$ (with v^{g_n} a continuation value function in discrete-time game g_n) converging pointwise to some function $v : \mathbb{R}_+ \to [0,1]$. I have already shown that $u_1^{*,g_n}(\gamma^{g_n}) \geq u_1^*(\gamma^{g_n})$ for any posture γ^{g_n} in game g_n, or equivalently $u_1^{*,g_n}(v^{g_n}) \geq u_1^*(v^{g_n})$. This immediately implies that $\lim_{n\to\infty} u_1^{*,g_n}(v^{g_n}) \geq \limsup_{n\to\infty} u_1^*(v^{g_n})$ for every convergent subsequence of $\{u_1^{*,g_n}(v^{g_n})\}$. Hence, I must show that $\lim_{n\to\infty} u_1^{*,g_n}(v^{g_n}) \leq \liminf_{n\to\infty} u_1^*(v^{g_n})$ for every convergent subsequence of $\{u_1^{*,g_n}(v^{g_n})\}$. I establish this inequality by assuming that there exists $\eta > 0$ such that $\lim_{n\to\infty} u_1^{*,g_n}(v^{g_n}) > \liminf_{n\to\infty} u_1^*(v^{g_n}) + \eta$ for some convergent subsequence of $\{u_1^{*,g_n}(v^{g_n})\}$ and then deriving a contradiction.

Let $t_{g_n}^{next}(i) = \min\{\tau > t : \tau \in I_i^{g_n}\}$ be the time of player i's next demand at t. Given continuation value function v^{g_n}, fix any corresponding posture γ^{g_n}, let $\widetilde{\gamma^{g_n}}^n$ be defined as follows: First, $\widetilde{\gamma^{g_n}}^n$ demands $\widetilde{\gamma^{g_n}}^n(h^t) = \gamma^{g_n}(h^t)$ for all $t \in I_1^{g_n}$. Second, $\widetilde{\gamma^{g_n}}^n$ accepts player 2's demand at time $t \in I_2^{g_n}$ with probability

$$\hat{p}^n(t) \equiv \min\left\{\frac{p^n(t)}{\chi^n(t)}, 1\right\},$$

where

$$p^n(t) \equiv \max_{\substack{\tau < t: \\ \tau \in I_1^{g_n}, \tau_{g_n}^{next}(2) = t}} \frac{e^{r(t-\tau)} v^{g_n}(\tau) - v^{g_n}(t)}{1 - v^{g_n}(t)}$$

if $\left\{\tau < t : \tau \in I_1^{g_n}, \tau_{g_n}^{next}(2) = t\right\}$ is non-empty and $v^{g_n}(\tau) < 1$ for all time τ in this set, and $p^n(t) \equiv 0$ otherwise; and

$$\chi^n(t) \equiv \max\left\{\frac{\Pi_{\tau < t : \tau \in I_2^{g_n}}(1 - p^n(\tau)) - \varepsilon}{\Pi_{\tau < t : \tau \in I_2^{g_n}}(1 - p^n(\tau))}, 0\right\}.$$

Let $\tilde{T}^n$ be the supremum over times t at which $\chi^n\left(t_{g_n}^{next}(2)\right)\hat{p}^n\left(t_{g_n}^{next}(2)\right) = p^n\left(t_{g_n}^{next}(2)\right)$, and let

$$T^n \equiv \sup \operatorname*{argmax}_{\substack{t \geq \tilde{T}^n: \\ t \in I_1^{g_n}}} e^{-rt} v^{g_n}(t).$$

By an argument similar to the proof of Lemma 2, if $\gamma^{g_n}(t) < \eta$ for some $t \leq T^n$, then there exists a belief $\pi_2 \in \Delta(\Sigma_1^{g_n})$ and strategy $\sigma_2 \in \Sigma_2^{g_n}$ such that $\pi_2(\gamma^{g_n}) \geq \varepsilon$, $\sigma_2 \in \Sigma_2^{*,g_n}(\pi_2)$, and the demand $\gamma^{g_n}(t)$ is accepted under strategy profile (γ^{g_n}, σ_2). In particular, $u_1^{g_n}(\gamma^{g_n}, \sigma_2) < \eta$. Thus, by the hypothesis that $\lim_{n\to\infty} u_1^{*,g_n}(v^{g_n}) > \liminf_{n\to\infty} u_1^*(v^{g_n}) + \eta$, there must exist $N > 0$ such that $\gamma^{g_n}(t) \geq \eta$ for all $t \leq T^n$ and all $n > N$, and hence $v^{g_n}(t) \leq 1 - \eta$ for all $t \leq T^n$ and all $n > N$.

Let π_2^n assign probability ε to γ^{g_n} and probability $1-\varepsilon$ to $\widetilde{\gamma^{g_n}}^n$, and fix $\sigma_2^n \in \Sigma_2^{*,g_n}(\pi_2^n)$ with the property that σ_2^n always demands 1 and rejects player 1's demand at any history at which player 1 has deviated from γ^{g_n} (which is possible because π_2^n assigns probability 0 to such histories, except for terminal histories), as well as at any history at which player 2 is indifferent between accepting and rejecting player 1's demand, given belief π_2^n. Note that γ^{g_n} is a best-response to σ_2^n in g_n. This implies that $u_1^{*,g_n}(\gamma^{g_n}) \leq u_1^{g_n}(\gamma^{g_n}, \sigma_2^n)$ for all n. Thus, to show that $\lim_{n\to\infty} u_1^{*,g_n}(v^{g_n}) \leq \liminf_{n\to\infty} u_1^*(v^{g_n}) + \eta$ (the desired contradiction), it suffices to show that $\lim_{n\to\infty} u_1^{g_n}(\gamma^{g_n}, \sigma_2^n) \leq \liminf_{n\to\infty} u_1^*(v^{g_n}) + \eta$.

Observe that $p^n(t)$ satisfies

$$\exp(-r(t-\tau))\left(p^n(t)(1) + (1-p^n(t))\, v^{g_n}(t)\right) \geq v^{g_n}(\tau)$$

for all $\tau \leq t$ such that $\tau \in I_1^{g_n}$ and $\tau_{g_n}^{next}(2) = t$. Hence, it is optimal for player 2 to reject player 1's demand γ at any time τ at which $\chi^n\left(\tau_{g_n}^{next}(2)\right)\hat{p}^n\left(\tau_{g_n}^{next}(2)\right) = p^n\left(\tau_{g_n}^{next}(2)\right)$, given belief π_2^n. Therefore, $u_1^{g_n}(\gamma^{g_n}, \sigma_2^n) = \min_{t\leq T^n} e^{-rt}(1 - v^{g_n}(t))$. Now $u_1^*(v^{g_n}) = \min_{t\leq T(v^{g_n})} e^{-rt}(1 - v^{g_n}(t))$, and $\lim_{n\to\infty} \tilde{T}(v^{g_n}) = \tilde{T}(v)$. Hence, showing that $\lim_{n\to\infty} \tilde{T}^n = \tilde{T}(v) \equiv \tilde{T}$ would imply that $\lim_{n\to\infty} u_1^{g_n}(\gamma^{g_n}, \sigma_2^n) = \liminf_{n\to\infty} u_1^*(v^{g_n})$, yielding the desired contradiction.

To see that $\lim_{n\to\infty} \tilde{T}^n = \tilde{T}$, first fix $t_0 \leq \tilde{T}$ and note that for all $\delta > 0$ there exists $N' > 0$ such that, for all $t \leq t_0$ and all $n \geq N'$, if $g_n(t) = 2$ then $\min\left\{\tau \leq t : \tau \in I_1^{g_n}, \tau_{g_n}^{next}(2) = t\right\} \geq t - \delta$ (if this set is non-empty). Next, since both $e^{-rt}v(t)$ and $e^{-rt}v^{g_n}(t)$ are non-increasing (by the same argument that showed that $e^{-rt}v(t)$ is non-increasing) and $v^{g_n}(t) \to v(t)$ for all $t \in \mathbb{R}_+$, it follows that for all $\delta' > 0$ there exists $\delta > 0$ such that $t \leq t_0$ and $\tau \in [t-\delta, t]$ implies that $\left|e^{r(\tau-t)}v^{g_n}(\tau) - v(t,-1)\right| < \delta'$. Since $1 - v(t) \geq \eta$ for all $t \leq \tilde{T}$, combining these observations and letting S be the (countable) set of discontinuity points of $v(t)$, for all $\delta' > 0$ there exists N'' such that if $t = s_{g_n}^{next}(2)$ for some $s \in S \cap [0, t_0]$, and $n \geq N''$, then $\left|p^n(t) - \frac{v(t,-1)-v(t)}{1-v(t)}\right| < \delta'$.[31] Hence,

$$\lim_{n\to\infty} \prod_{s\in S\cap[0,t_0]} \left(1 - p^n\left(s_{g_n}^{next}(2)\right)\right) = \prod_{s\in S\cap[0,t_0]} (1 - p(s)) \tag{28}$$

for all $t_0 \leq \tilde{T}$.

Finally, I establish that, whenever v is continuous on an interval $[t_0, t_\infty]$ with $t_\infty \leq \tilde{T}$,

$$\lim_{n\to\infty} \prod_{t\in I_2^{g_n}\cap[t_0,t_\infty]} (1 - p^n(t)) = \exp\left(-\int_{t_0}^{t_\infty} \frac{rv(t) - v'(t)}{1 - v(t)} dt\right) = \exp\left(-\int_{t_0}^{t_\infty} \lambda(t)\, dt\right). \tag{29}$$

I will prove this fact by showing that the limit as $n \to \infty$ of a first-order approximation of the logarithm of $\prod_{t\in I_2^{g_n}\cap[t_0,t_\infty]} (1 - p^n(t))$ equals $-\int_{t_0}^{t_\infty} \frac{rv(t)-v'(t)}{1-v(t)}$.

Let $\left\{t_{1,g_n}, t_{2,g_n}, \ldots, t_{K(n),g_n}\right\} = \{t \in [t_0, t_\infty] : p^n(t) > 0\}$, with $t_{k,g_n} < t_{k+1,g_n}$ for all $k \in \{1, \ldots, K(n) - 1\}$ and all $n \in \mathbb{N}$; note that $K(n)$ is finite because $I_2^{g_n} \cap [t_0, t_\infty]$ is finite, and that in addition $t_{k,g_n}^{next}(1) < t_{k+1,g_n}$ for all k (where $t_{k,g_n}^{next}(1) \equiv t_{k,g_n,g_n}^{next}(1)$ to avoid redundant notation). Furthermore, since $e^{-r\tau}v^{g_n}(\tau)$ is non-increasing,

$$t_{k,g_n}^{next}(1) \in \operatorname*{argmax}_{\substack{\tau<t_{k+1,g_n}: \\ \tau\in I_1^{g_n}, \tau_{g_n}^{next}(2)=t_{k+1,g_n}}} e^{r\left(t_{k+1,g_n}-\tau\right)} v^{g_n}(\tau)$$

for all $k \in \{0, 1, \ldots, K(n) - 1\}$. Therefore,

$$\begin{aligned}
&\prod_{k=1}^{K(n)} (1 - p^n(t)) \\
= &\prod_{k=1}^{K(n)} \min_{\substack{\tau<t_{k,g_n}: \\ \tau\in I_1^{g_n}, \tau_{g_n}^{next}(2)=t_{k,g_n}}} \frac{1 - e^{r\left(t_{k,g_n}-\tau\right)} v^{g_n}(\tau)}{1 - v^{g_n}\left(t_{k,g_n}\right)} \\
= &\prod_{k=1}^{K(n)} \min_{\substack{\tau<t_{k,g_n}: \\ \tau\in I_1^{g_n}, \tau_{g_n}^{next}(2)=t_{k,g_n}}} \frac{1 - e^{r\left(t_{k,g_n}-\tau\right)} v^{g_n}(\tau)}{1 - e^{-r\left(t_{k,g_n}^{next}(1)-t_k\right)} v^{g_n}\left(t_{k,g_n}^{next}(1)\right)} \\
= &\left(\prod_{k=1}^{K(n)-1} \frac{1 - e^{r\left(t_{k+1,g_n}-t_{k,g_n}^{next}(1)\right)} v^{g_n}\left(t_{k,g_n}^{next}(1)\right)}{1 - e^{-r\left(t_{k,g_n}^{next}(1)-t_{k,g_n}\right)} v^{g_n}\left(t_{k,g_n}^{next}(1)\right)}\right) \frac{1 - e^{r\left(t_{1,g_n}-t_{0,g_n}^{next}(1)\right)} v^{g_n}\left(t_{0,g_n}^{next}(1)\right)}{1 - e^{-r\left(t_{K(n),g_n}^{next}(1)-t_{K(n),g_n}\right)} v^{g_n}\left(t_{K(n),g_n}^{next}(1)\right)}
\end{aligned} \tag{30}$$

[31] S is countable because $e^{-rt}v(t)$ is non-increasing, and monotone functions have at most countably many discontinuity points. Unlike in Section 4, S need not be a subset of $\mathbb{N}$ here.

Next, taking a first-order Taylor approximation of $\log\left(1-e^{rx}v^{g_n}(t)\right)$ at $x=0$ yields

$$\log\left(1-e^{rx}v^{g_n}(t)\right)=\log\left(1-v^{g_n}(t)\right)-\frac{rxv^{g_n}(t)}{1-v^{g_n}(t)}+O\left(x^2\right).$$

Therefore, a first-order approximation of the logarithm of (30) equals

$$\begin{aligned}&\left(\sum_{k=1}^{K(n)-1}-\left(t_{k+1,g_n}-t_{k,g_n}\right)\frac{rv^{g_n}\left(t_{k,g_n}^{next}(1)\right)}{1-v^{g_n}\left(t_{k,g_n}^{next}(1)\right)}\right)\\&+\log\left(1-e^{r\left(t_{1,g_n}-t_{0,g_n}^{next}(1)\right)}v^{g_n}\left(t_{0,g_n}^{next}(1)\right)\right)-\log\left(1-e^{-r\left(t_{K(n),g_n}^{next}(1)-t_{K(n),g_n}\right)}v^{g_n}\left(t_{K(n),g_n}^{next}(1)\right)\right).\end{aligned}$$

I now show that

$$\lim_{n\to\infty}\sum_{k=1}^{K(n)-1}-\left(t_{k+1,g_n}-t_{k,g_n}\right)\frac{rv^{g_n}\left(t_{k,g_n}^{next}(1)\right)}{1-v^{g_n}\left(t_{k,g_n}^{next}(1)\right)}=-\int_{t_0}^{t_\infty}\frac{rv(t)}{1-v(t)}dt \tag{31}$$

and

$$\begin{aligned}&\lim_{n\to\infty}\left(\log\left(1-e^{r\left(t_{1,g_n}-t_{0,g_n}^{next}(1)\right)}v^{g_n}\left(t_{0,g_n}^{next}(1)\right)\right)-\log\left(1-e^{-r\left(t_{K(n),g_n}^{next}(1)-t_{K(n),g_n}\right)}v^{g_n}\left(t_{K(n),g_n}^{next}(1)\right)\right)\right)\\=&\int_{t_0}^{t_\infty}\frac{v'(t)}{1-v(t)}dt,\end{aligned}\tag{32}$$

which completes the proof of (29). Equation (32) is immediate, because, since v is continuous on $[t_0,t_\infty]$, both the left- and right-hand sides equal

$$\log\left(1-v\left(t_0\right)\right)-\log\left(1-v\left(t_\infty\right)\right).$$

To establish (31), let

$$f^n(t)\equiv\exp\left(-r\left(\frac{1+\eta}{\eta}\right)t\right)\frac{rv^{g_n}(t)}{1-v^{g_n}(t)}$$

and

$$f(t)\equiv\exp\left(-r\left(\frac{1+\eta}{\eta}\right)t\right)\frac{rv(t)}{1-v(t)}.$$

For all $n>N$, it can be verified that both $f^n(t)$ and $f(t)$ are non-increasing on the interval $[t_0,t_\infty]$, using the facts that $e^{-rt}v^{g_n}(t)$ and $e^{-rt}v(t)$ are non-increasing, and that $v^{g_n}(t)\le 1-\eta$ for all $n>N$ and $t\le t_\infty\le\tilde{T}$. Fix $\zeta>0$ and $m\in\mathbb{N}$. Because $v^{g_n}(t)\to v(t)$ for all $t\in\mathbb{R}_+$, there exists $N'''\ge N$ such that, for all $n>N'''$, $|f^n(t)-f(t)|<\zeta$ for all t in the set

$$\left\{t_0,\frac{(m-1)t_0+t_\infty}{m},\frac{(m-2)t_0+2t_\infty}{m},\ldots,t_\infty\right\}.$$

Since both f^n and f are non-increasing on $[t_0, t_\infty]$, this implies that

$$|f^n(t) - f(t)| < \zeta + \max_{k \in \{1,\ldots,K(n)-1\}} \left(f\left(\frac{(m-k)\,t_0 + kt_\infty}{m}\right) - f\left(\frac{(m-k-1)\,t_0 + (k+1)\,t_\infty}{m}\right)\right)$$

for all $t \in [t_0, t_\infty]$. Since f is continuous on $[t_0, t_\infty]$, taking $m \to \infty$ implies that $|f^n(t) - f(t)| < 2\zeta$ for all $t \in [t_0, t_\infty]$, and therefore $\left|\frac{rv^{g_n}(t)}{1-v^{g_n}(t)} - \frac{rv(t)}{1-v(t)}\right| \leq 2\zeta \exp\left(r\left(\frac{1+\eta}{\eta}\right)t_\infty\right)$ for all $t \in [t_0, t_\infty]$. Hence,

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=1}^{K(n)-1} -\left(t_{k+1,g_n} - t_{k,g_n}\right) \frac{rv^{g_n}\left(t_{k,g_n}^{next}(1)\right)}{1 - v^{g_n}\left(t_{k,g_n}^{next}(1)\right)} &= \lim_{n\to\infty} \sum_{k=1}^{K(n)-1} -\left(t_{k+1,g_n} - t_{k,g_n}\right) \frac{rv\left(t_{k,g_n}^{next}(1)\right)}{1 - v\left(t_{k,g_n}^{next}(1)\right)} \\
&= \lim_{n\to\infty} \sum_{k=1}^{K(n)-1} -\left(t_{k+1,g_n} - t_{k,g_n}\right) \frac{rv\left(t_{k,g_n}\right)}{1 - v\left(t_{k,g_n}\right)} \\
&= -\int_{t_0}^{t_\infty} f(t)\,dt,
\end{aligned}$$

where the first equality follows because $\sum_{k=1}^{K(n)-1}\left(t_{k+1,g_n} - t_{k,g_n}\right) \leq t_\infty - t_0$ for all $n \in \mathbb{N}$, the second follows because $t_{k,g_n}^{next}(1) \in \left[t_{k,g_n}, t_{k+1,g_n}\right]$ and v is continuous on $[t_0, t_\infty]$, and the third follows by definition of the (Riemann) integral.

Combining (28) and (29), it follows that

$$\lim_{n\to\infty} \prod_{s \in I_2^{g_n} \cap [0,t]} \left(1 - p^n(s)\right) = \exp\left(-\int_0^t \lambda(s)\,ds\right) \prod_{s \in S \cap [0,t]} \left(1 - p(s)\right)$$

for all $t \leq \tilde{T}$. This implies that $\lim_{n\to\infty} \tilde{T}^n = \tilde{T}$, completing the proof of the lemma. ■

I now complete the proof of Theorem 3.

Let $\{g_n\}$ be a sequence of discrete-time bargaining games converging to continuous time. Recall that $u_1^{*,g_n} = \sup_{\gamma^{g_n}} u_1^{*,g_n}(\gamma^{g_n})$. Thus, there exists a sequence of postures $\{\gamma^{g_n}\}$, with γ^{g_n} a posture in g_n, such that $\lim_{n\to\infty} |u_1^{*,g_n} - u_1^{*,g_n}(\gamma^{g_n})| = 0$. Let $\{v^{g_n}\}$ be the corresponding sequence of continuation value functions. Because $e^{-rt}v^{g_n}(t)$ is non-increasing and the space of monotone functions from $\mathbb{R}_+ \to [0,1]$ is sequentially compact (by Helly's selection theorem or footnote 25), this sequence has a convergent subsequence $\{v^{g_k}\}$ converging to some v on $\mathbb{R}_+$.

I claim that $v = v^*$. Toward a contradiction, suppose not. Since v^* is the unique maxmin continuation payoff function in g^{cts}, there exists $\eta > 0$ such that $u_1^* > \lim_{k\to\infty} u_1^*(v^{g_k}) + \eta$. By Lemma 5, $\lim_{k\to\infty} u_1^{*,g_k}(v^{g_k}) = \lim_{k\to\infty} u_1^*(v^{g_k})$. Finally, again by Lemma 5, there exists an alternative sequence of postures $\{\gamma^{g_k\prime}\}$ such that $\lim_{k\to\infty} u_1^{*,g_k}(\gamma^{g_k\prime}) \geq u_1^*$. Combining these observations implies that there exists $K > 0$ such that, for all $k \geq K$,

$$u_1^{*,g_k}(\gamma^{g_k\prime}) > u_1^* - \eta/3 > u_1^*(v^{g_k}) + 2\eta/3 > u_1^{*,g_k}(v^{g_k}) + \eta/3,$$

which contradicts the fact that $\lim_{k\to\infty} |u_1^{*,g_k} - u_1^{*,g_k}(\gamma^{g_k})| = 0$. Therefore, $v = v^*$. Since this argument applies to any convergent subsequence of $\{v^{g_n}\}$, and every subsequence of $\{v^{g_n}\}$ has a convergent sub-subsequence, this implies that $v^{g_n} \to v^*$ pointwise.

A similar contradiction argument shows that $\lim_{k\to\infty} u_1^{*,g_k}(v^{g_k}) = u_1^*$, for any convergent subsequence $\{v^{g_k}\} \subseteq \{v^{g_n}\}$. Since $\lim_{k\to\infty} |u_1^{*,g_k} - u_1^{*,g_k}(\gamma^{g_k})| = 0$, it follows that $u_1^{*,g_k} \to u_1^*$. And, since this argument applies to any convergent subsequence of $\{v^{g_n}\}$, this implies that $u_1^{*,g_n} \to u_1^*$. ∎