Repeated reputational bargaining with deadlines

March 26, 2012

Abstract

We develop a two-sided reputational bargaining model with deadlines, and analyze the implications of linking a reputation for commitment on one bargaining issue, to reputation on future issues. The model is adapted from that of Abreu and Gul [2000], where some agents are committed to achieving a fixed share of any surplus available. Among the conclusions drawn are: the ordering of issues on the bargaining agenda may be of great importance; disagreement may arise even when the probability agents are commitment to demands is arbitrarily small when this is combined with uncertainty about future bargaining conditions; the right-obligation to make the first proposal in bargaining may have significant payoff implications when the reputational cost to backing down from an initial demand is large; long run agents typically have an advantage in bargaining over short run agents, but may in fact optimally moderate their demands.

1 Introduction

The motivation for this work was the 2011 debt ceiling debates in Congress. Under U.S. law the Treasury can only borrow money up to a limit imposed by Congress. The Treasury announced that this limit needed to raised by August 2, or else it would be unable to meet its obligations. Previously raising the debt ceiling had been a non-partisan affair, but Republican legislators who controlled the House of Representatives, used the threat of default as a bargaining chip with which to secure significant spending cuts, without conceding to any tax increases.

In the words of Republican House majority leader John Boehner: “When you look at this final agreement that we came to with the white House, I got 98 percent of what I wanted. I’m pretty happy.” Whether that was quite true or displayed some bravado, certainly the Democrats caved in to many Republican demands, without even securing the closing of tax loopholes, despite a seeming position of strength, having control of both the Senate and the White House.

The Republicans’ success was been widely attributed to an ‘irrational’ unwillingness to compromise which left many believing they would rather see the Treasury default than back down on their political objectives. Certainly the rhetoric of some suggested as much, for instance former Presidential candidate Michelle Bachmann, embraced the prospect of the Treasury not meeting its obligations, claiming it would be “tough love” and not the disaster painted by White House “scare tactics”.

1
The success of Republican tactics, left the likes of Paul Krugman (in the *New York Times*) fuming, and pondering how such an outcome could have come to pass, and its implications for the future: “The G.O.P. has just demonstrated its willingness to risk financial collapse unless it gets everything its most extreme members want. Why expect it to be more reasonable in the next round? In fact, Republicans will surely be emboldened by the way Mr. Obama keeps folding in the face of their threats. He surrendered last December, extending all the Bush tax cuts; he surrendered in the spring when they threatened to shut down the government; and he has now surrendered on a grand scale to raw extortion over the debt ceiling. Maybe its just me, but I see a pattern here.”

Reputational bargaining models, notably Abreu and Gul [2000], have shown the extent to which players’ reputations for commitment can secure higher payoffs in a bargaining context. Rational players of both parties must imitate irrational types behavior, or risk being taken advantage of, and this can explain costly delay in reaching agreement even between agents who are in fact rational and when there are sure gains from trade. When the probability of irrationality is small however, bargaining yields an essentially unique outcome that depends only on payoff fundamentals, the agents’ relative levels of patience.

This paper extends that framework to the context of bargaining with deadlines appropriate for analyzing the debt ceiling debacle, before investigating exactly the repeated game pattern that Krugman identified. When two agents must bargain repeatedly on many issues, reputations carry over from one context into the next. This clearly creates strong incentives not to compromise in early battles as doing so ensures a conciliator’s reputation ensuring that one must concede in future battles. We investigate how this basic force interacts with other features of the bargaining environment to determine both distributional issues and the possibility of inefficiency and disagreement.

The majority of the results in the paper are concerned with situations when the probability of initial commitment is sufficiently small. Such slight perturbations to complete rationality might not always be appropriate if one believes the probability of obstinate agents of particular types (say a “fair” split) is large, but this does allow clear model predictions that do not depend on arbitrary initial parameters, and which illustrate some of the important forces at work. Outcomes may differ markedly from the complete information world.

Our baseline repeated bargaining model features two agents bargaining over two known issues sequentially. Agenda setting power is highlighted to be of key importance in this setting. For small probabilities of irrationality the terms on which *both* agreements are made depend almost entirely by agents’ relative strength on the final bargaining issue (their relative payoffs if no agreement is reached). The result resembles an extreme form of backward induction, where strength at the end of bargaining is directly translated into strength at the beginning. The second notable feature is that while agreement is reached almost immediately on both issues the outcome is still typically inefficient. The reason is that even though agents may value the two pies equally (issues of pure disagreement) the necessity of imitating commitment types does not allow for payoff sets to be expanded to take advantage of agents differing rates of time preference. When dividing one dollar on two subsequent days, agent 1 may value the dollar tomorrow more than agent 2, and so an efficient division would give him more of tomorrow’s dollar and less of today’s, but the bargaining outcome will require a 50/50 division on both days. This is true even when agents can imitate irrational types who make different demands for the first and
second dollar.

A natural concern in repeated bargaining situations is uncertainty about what future bargaining issues will arise. Extending the model to allow for this can show that uncertainty can lead to almost certain disagreement on early bargaining issues, before that uncertainty is resolved. This disagreement exists even for arbitrarily small probabilities of irrationality, in marked contrast to the complete information world. The force at work is the need to preserve a reputation for being an aggressive bargainer, if future bargaining environments turn out to be favorable.

Building on Krugman’s insight that early concessions to the Republicans made their demands even more aggressive on future bargaining issues, we adapt the model to allow for this, and show that such a possibility is typically disadvantageous in bargaining, as it makes an opponent’s threat not to concede on early bargaining issues credible. We also show that in such a model, the order in which agents make their initial demands can matter greatly. Staking out a position, may in some circumstances be highly advantageous, while in others, it would be preferable for an opponent to make the first move.

Finally, we model a situation in which a long run agent bargains against short run agents in consecutive periods. As one would expect in such a scenario, the need to preserve a reputation for being committed to a bargaining position gives the long run agent an advantage in bargaining against early short run agents. However, whether the long run agent takes advantage of that fact depends on a more delicate tradeoff: aggressive demands on a first bargaining issue may be conceded to, but developing a reputation which is too aggressive may leave the agent in a weak position for future bargaining. Whether rational long run agents are initially aggressive then, or make moderate demands when facing all agents, is then determined by the relative size of the surplus available in the two periods.

Many papers have relevance to this work, but probably none more so than Abreu and Gul [2000]. Notable for building on this framework is the paper of Abreu and Pearce [2007], where the small possibility of commitment types, leads to agents writing long term contracts that yield Nash Bargaining with threats payoffs in an infinitely repeated game. Abreu et al. [2012] show that a small possibility of irrational agents who delay making initial offers, combined with asymmetric information about even one agent’s discount factors may lead to significant delay in agreement and non-Coasean outcomes.

A subtly different model of reputational bargaining from Abreu and Gul [2000] has been developed by ?, on which other papers have also built. The work is related to, but different from, bargaining in the presence of unknown valuations. Inderst [2005] makes some progress at integrating the two approaches with a possibly committed seller and a buyer with unknown valuation. For a review of many early results on bargaining with private information see Kennan and Wilson [1993].

Skrzypacz and Toikka [2012] investigates repeated bargaining from a mechanism design perspective when agents have private valuations. The paper shows that tying bargaining from many periods together can allow efficient, individually rational and budget balanced outcomes despite the negative results of Roger B Myerson [1983]. The threat to terminate a future relationship if agents do not pay a tax sufficient to lubricate efficient trade allows for truthful revelation.

Many models of bargaining incorporate deadlines, and private information with deadlines, and many of these are illustrated in Kennan and Wilson. Deadlines are modeled in different ways;
in this paper the deadline arrives stochastically within some fixed interval. Roth et al. [1988] present experimental evidence for a strong deadline effect in bargaining which would seem to be a good fit for a stochastic deadline model with some ‘irrational’ types. Ockenfels and Roth [2006] investigate somewhat comparable deadline effects in ebay auctions.

The result regarding the importance of agenda in sequential bargaining is somewhat related to Fershtman [1990], who shows that when agents value pies differently, and pies divided early in bargaining cannot be eaten until all issues are agreed, it helps to have the issue you value most debated last. This order preference is in fact the exact opposite of the result in this paper, where pies can be eaten immediately after agreement. However, in some sense the force at work is the same, weakness at final stage of bargaining affects bargaining over both issues. In Fershtman a provisional agreement on the first issue about which one agent cares a lot, leaves him in a weak bargaining position on the second issue, due to his impatience to eat the agreed share.

The finding that uncertainty leads to disagreement relates to Christopher Avery [1994]. In an alternating offers game where offers cannot be withdrawn, and the asset being ‘sold’ may change in value, agents may make low offers, which are only accepted when the asset falls in value. Olivier Compte [2004] show that when generous offers increase an opponent’s outside option, gradualism (and delay) in bargaining may result. In this paper generous mutually beneficial offer’s are not made as these would in fact increase an opponent’s option value of waiting until the revelation of uncertainty.

The negative effect on bargaining power created by the inability to commit not to take advantage of an opponent’s revelation of rationality, is related to a great host of results on lack of commitment power. The finding of advantages to a long run agent also has many parallels in the literature, although the more surprising result there is that the long run agent may not take advantage of this in early bargaining.

2 The one period model

We first set up a basic one period bargaining model with deadlines before adapting and extending it in a multi-period setting. Two agents 1 and 2 bargain over the division of surplus whose value is 1. There is some deadline ‘t’, by which time bargaining must be completed or else agents will obtain their disagreement payoffs, \(-D \leq 0\), in addition to obtaining none of the surplus. Payoffs are not time discounted.

The interpretation is that a potentially mutually beneficial agreement (e.g. on deficit reduction) will include a minimum of common ground (avoiding U.S. default), combined with areas where agents disagree (deficit reduction from spending cuts or tax increases). If there is an agreement before the deadline and agent i obtains everything he wanted on issues where the two agents disagree, he obtains a payoff of 1, while j gets 0. The agents’ bargain over who gets most of what they want on issues where they disagree, knowing what would be each other’s first best solution. We assume that there is no transferable utility between agents, which can allow them to obtain payoffs greater than 1 or less than 0, as would certainly seem appropriate in a political bargaining model. Setting the minimum common ground payoff agreement payoff to be zero is simply a normalization.
Adapting this model beyond the political economy context, to other scenarios in which deadlines are applicable, for instance when bargaining over perishable goods or services, the disagreement payoffs may correspond to bounds on the maximum and minimum prices certain goods and services can be sold for; due to legal requirements such as the minimum wage, or company policy forced upon sales agents.

The exact deadline $t$ by which any agreement must be made is uncertain. We assume that it is distributed according to a distribution $G$ on $[0, T]$ with continuous positive density $g$. The ultimate deadline is time $T$, by which point it is certain that no agreement can be made. However, even if a parties attempt to make an agreement at some time $t < T$, it is still possible (probability $G(t)$) that this cannot be implemented. The analogy with the debt ceiling debate is that even after parties had agreed a plan on July 31, the complexities of passing this into law, left uncertainty over whether this could actually be enacted before August 2. Uncertainty was further present in that some analysts also claimed that the Treasury in fact had enough money to last until August 15.

We follow the existing reputational bargaining literature closely. These papers typically have an “open ended” bargaining structure in which there is no time limit before which agreement must be made. Agents are impatient, and so earlier agreements are preferable to later ones. This impatience drives the bargaining in a similar way to our continuous deadline distribution.

Within that reputational bargaining literature there is some variation in the exact bargaining protocol used and methods of reputation formation, notably there are the different approaches of Abreu and Gul [2000] and ?. In this paper, we follow the setup of Abreu and Gul, whose model is in many ways more natural, than Kambe’s, allowing it to be more likely that agents are committed some demands than others, say a 50/50 division of surplus. Moreover, their model allows for significant delay in reaching agreement when the probability of irrational agents is large, while in Kambe there is always an equilibrium with immediate agreement.

The structure of the game is that agent 1 demands a share of the surplus $\alpha_1$ at time 0, whereupon agent 2 can either immediately concede, or make a counteroffer $\alpha_2$, such that $\alpha_1 + \alpha_2 > 1$. Agents can potentially concede to their opponents demand or change their demand, at any time at any time before $T$.

Reputational dynamics comes from an initial prior probability that agents are irrational $z_i$. An irrational type of player $i$ is identified by a demand $\alpha_i \in (0, 1)$; a type $\alpha_i$ always demands $\alpha_i$ and accepts any offer greater or equal to $\alpha_i$, rejecting smaller offers. $C_i$ is the finite set of irrational types, who conditional on irrationality, have probability $\pi(\alpha_i)$.

Given the presence of irrational agent types, following demands $\alpha_i + \alpha_j > 1$ (which are made by irrational players with positive probability), Abreu and Gul show that this game must have a war of attrition structure. The reason is that changing demands after initially demanding $\alpha_i$, reveals to one’s opponent that you are not a committed type of player. If agent $i$ has a reputation for being a committed player but $j$ does not, then reasoning analogous to that of the Coase conjecture requires that in equilibrium, the rational player must immediately concede to the (potentially) irrational player’s demands. They derive a continuous time war of attrition game as the unique limit of discrete time games in which both agents make offers arbitrarily frequently. It is clear that their proof adapts readily to this setting, and given this we focus directly on the continuous time game.
Following much of the notation of Abreu and Gul, a strategy for player 1, $\sigma^1$, is defined by a probability distribution $\mu_1$ on $C^1$ of initial demands by rational agents and a collection of cumulative distributions $F_{\alpha^1}^1$ on $[0, T]$, where $F_{\alpha^1}^1(t)$ is the total probability of player 1 conceding to player 2 by time $t$ following demands $\alpha^1$ and $\alpha^2$. The strategy of not conceding is captured by a concession time of $T$.

A strategy for player 2, $\sigma^2$, requires a collection $\mu_2^2$ on $C^2 \cup \{Q\}$ specifying rational 2’s counter-demand to $\alpha^1$, where $Q$ is immediate acceptance, and $F_{\alpha^1,\alpha^2}^2$ describes 2’s choice of concession time given demands $\alpha^1, \alpha^2$. Given $\mu^1$ and $\mu^2_{\alpha^1}$, the conditional (posterior) probability that agents are irrational after making their demands is given by:

$$
\bar{z}^1(\alpha^1) = \frac{z^1\pi^1(\alpha^1)}{z^1\pi^1(\alpha^1) + (1 - z^1)\mu^1(\alpha^1)}
$$

$$
\bar{z}^2_{\alpha^1}(\alpha^2) = \frac{z^2\pi^2(\alpha^2)}{z^2\pi^2(\alpha^2) + (1 - z^2)\mu^2_{\alpha^1}(\alpha^2)}
$$

Conditional on a war of attrition being started, with demands $\alpha^1, \alpha^2$, agent $i$’s expected utility from conceding at time $t$, given opponents strategy $\sigma^j$, is given by:

$$
U^i(t, \sigma^j|\alpha) = \int_0^t \alpha^i(1 - G(s)) - D^iG(s) dF^j_{\alpha^1,\alpha^2}(s)
+ (1 - F^j_{\alpha^1,\alpha^2}(t))[(1 - G(t))(1 - \alpha^j) - G(t)D^j]
+ \frac{1}{2}(F^j_{\alpha^1,\alpha^2}(t) - F^j_{\alpha^1,\alpha^2}(t-))[(1 - G(t))\frac{1}{2}(\alpha^1 - \alpha^2 - 1) - G(t)D^j]
$$

(1)

Given this rational agent 1’s utility at the start of the war of attrition is:

$$
U^1(\sigma|\alpha) = \frac{1}{1 - \bar{z}^1(\alpha^1)} \int_0^T U_1(y, \sigma_2|\alpha) dF_{\alpha^1,\alpha^2}^1(y)
$$

With a similar expression for $U^2(\sigma|\alpha)$. Finally a rational players’ expected utility is given by:

$$
U^1(\sigma) = \sum_{\alpha^1} \mu^1(\alpha^1)[(1 - z^1)\mu^2_{\alpha^1}(Q) + z^1 \sum_{\alpha^2 \leq \alpha^1} \pi^2(\alpha^2)]
+ \sum_{\alpha^2 > \alpha^1} U^1(\sigma|\alpha)((1 - z^1)\mu^2_{\alpha^1}(\alpha^2) + z^1 \pi^2(\alpha^2))
$$

$$
U^1(\sigma) = \sum_{\alpha^1} [(1 - z^1)\mu^1(\alpha^1) + z^1 \pi^1(\alpha^1)]
\times [(1 - \alpha^1)\mu^2_{\alpha^1}(Q) + \sum_{\alpha^2 > 1 - \alpha^1} U^2(\sigma|\alpha)\mu^2_{\alpha^1}(\alpha^2)]
$$

(2)

The equations in (2) are exact analogues of agents payoffs in Abreu and Gul. The choice of demands in our game is identical to theirs, the only difference is that payoffs $U^2(\sigma|\alpha)$ are determined in a slightly different manner, reflecting different continuation games. In what follows we frequently drop the subscripts of $F^i_{\alpha^1,\alpha^2}$ and other variables, when the context is clearly un-
derstood.

To solve this model, we first consider the game with only one irrational type for each player, such that $\alpha_1 + \alpha_2 > 1$. We sketch the standard arguments which show the structure of a unique sequential equilibrium. Suppose there is an equilibrium. At most one agent can concede with positive mass at any given time $t \in [0, T)$; given the continuous density of deadline failure, waiting $\Delta$ periods longer would secure at least one agent a positive profit bump. Given mass concession by agent $j$ at $t' \in (0, T)$ agent $i$ cannot concede at $t \in (t' - \Delta, t')$, because doing so forgoes agent $i$ a profit bump. But then agent $j$ would not find it optimal to concede at $t'$, preferring to concede at $t' - \Delta$ so there is no mass acceptance at times $(0, T)$. A corollary is that there can also be no time $T^i < T$ at which agent $i$ is revealed to be certainly irrational, but agent $j$ is not (this would initiate mass concession). Furthermore, there must be no interval $(t', t'')$ in which agent $i$ does not concede with positive probability, while agent $j$ does, otherwise $j$ has an incentive to instead concede at $t'$ (and avoid the risk of disagreement). And hence there can be no interval $(t', t'')$ on which agents do not concede, while they do in the interval $[t'', T)$, because agent $i$ who is supposed to concede at $t''$, prefers to instead concede at $t'$ given no mass concession at $t''$. This implies that $F^j$ is continuous on $(0, T)$.

Now suppose there is mass concession by rational agents at period $T$. For this to be optimal there must be concession with positive probability on any interval $(T - \Delta, T)$, if not then agent $i$ has a strict incentive to concede at $T - \Delta$. This certainly implies that there must be positive concession on the interval $(0, T)$. For agents to be willing to concede on an uncountable interval $(0, T^*] \subseteq (0, T]$ they must obtain constant utility on that interval. This ensures that $U^i(t, \sigma^j|\alpha)$ is differentiable with respect to $t$ (with derivative of 0), from which we find concession rates to ensure indifference. Given these to complete the equilibrium we find initial concession with positive probability from one (uniquely determined) agent that makes both agents’ probability of irrationality hit one at the same time $T^*$. A more formal proof can be found in Abreu and Gul.

To find the concession rates then, we differentiate equation 1 with respect to $t$ to give:

$$0 = (\alpha_i + \alpha_j - 1)(1 - G(t))f^j(t) - g(t)(1 - F^j(t))(1 - \alpha_j + D^j)$$

This in turn yields a conditional concession rate for agent $j$:

$$f^j(t) = \frac{g(t)}{1 - G(t)}K^j$$

where:

$$K^j = \frac{1 - \alpha_j + D^j}{\alpha_i + \alpha_j - 1}$$

Which finally gives us a closed form distribution:

$$F^j(t) = 1 - c^j[1 - G(t)]^{K^j}$$

To close the model let $T^j < T$ be the potential exhaustion time of agent $j$ such that:

$$[1 - G(T^j)]^{K^j} = \bar{c}^j$$
And $T^* = \min(T^1, T^2)$. For $T^j = T^*$ we $c^j = 0$, for $T^j > T^*$ we then have:

$$c^j = \frac{\bar{z}^j}{(1 - G(T^*))^{K^j}} = \bar{z}^j(\bar{z}^j)^{-K^j/K^i} < 1$$

And so agent $j$ concedes with positive probability at time 0. Certainly however, at time $T^* < T$ it becomes certain that both agents are irrational, rational agents will have certainly tried to make a deal before this point, although they may not have succeeded (as the effective deadline may already have passed).

While from a pure theory point of view characteristics of this equilibrium are perhaps unsurprising given Abreu and Gul’s model they are nonetheless worth brief discussion, certainly in the context of political bargaining. A rational player $i$’s equilibrium utility at the start of this war of attrition is given by:

$$U^i(\alpha) = (1 - c^j)x^j + c^j(1 - \alpha^i)$$

Therefore an agent can only obtain a payoff higher than her opponent’s offer when there is positive probability of being conceded to at time 0. This concession rate is determined by which agent’s exhaustion times $T^i$. We have $T^j > T^i$, requiring immediate concession by agent $j$ only if $[1 - G(T^j)] < [1 - G(T^i)]$, which requires:

$$\frac{\log \bar{z}^i}{\log \bar{z}^j} < \frac{K^i}{K^j} = \frac{1 - \alpha^i + D^j}{1 - \alpha^j + D^i}$$

A player $i$’s strength (likelihood of winning a concession) in the war of attrition is increasing in the generosity of her own offer: $(1 - \alpha^i)$, her likelihood of irrationality $\bar{z}^i$, and decreasing in her disagreement cost $D^i$, with the converse true for the other player’s factors. The relative disagreement payoffs play a similar role to agents having different discount factors in the “open ended” Abreu and Gul model. An increased cost of disagreement leads agents to be less patient, or more anxious to get an agreement quickly.

Given a war of attrition, concession survivor functions are a convex or concave transformation of the deadline survivor function. If we expect a deadline distribution to be tight with most of the weight around the firm deadline $T$, then rational agents will not concede until very late in the day too. This helps explain why deals typically are cut at the last minute, as in the case of the U.S. debt ceiling, and also why disagreements frequently arise even when there is clearly room for them to benefit both parties. We also note that the given no immediate acceptance the distribution of attempted deals should be closer to the deadline, the higher is disagreement in bargaining positions $(\alpha^i + \alpha^j - 1)$. The larger is disagreement the smaller is $K^i$ and $K^j$, and hence concession is at a slower rate.

Such a reputational theory of bargaining with deadlines draws some support from Figure 1 from the experiment of Roth, Murninghan and Shoemaker (1988). The announced deadline to reach agreement in their experiment was 9 minutes (540 seconds), however the authors note there was almost certainly some delay in the processing power of the computers, meant the true time an agent could agree to another’s offer was uncertain around this time. Half of all agreements occur in the last 30 seconds, and half of those come in the last 5 seconds. Meanwhile there is
substantial fraction of agents for whom the (effective) deadline passed without agreement.

Figure 1: Agreement times in bargaining experiment: Roth, Murnighan and Shoumaker (1988)

The probability of disagreement before the deadline is given by:

\[
\int_0^T g(s)(1 - F^i(s))(1 - F^j(s))ds = \frac{1}{K^j + K^i + 1}\left[1 - (1 - G(T^s))^K^j + K^{j+1}\right]^{\frac{K^j}{K^i}} ds + (1 - G(T^s))^{\frac{K^j}{K^i}}
\]

Where we have assumed that \(T^j \geq T^i\). Clearly if \(D^i = D^j = 0\) this also represents the level of inefficiency.

In the general case, increasing \(D^i\) increases \(K^j\), and so the concession rate of agent i, which through lowering \(T^s\) also raises the immediate concession probability; hence it lowers both \((1 - F^i(s))\) and \((1 - F^j(s))\) for all \(s\) and certainly lowers the total probability of disagreement. On the other hand (assuming \(T^j > T^i\) else the former analysis applies) marginally increasing \(D^j\) increases \(K^j\), raising j’s conditional concession rate, at the expense of lowering the immediate concession probability. The net effect of this is then an increase in \((1 - F^j(s))\) and so an increase in the total probability of disagreement.

By increasing \(D^i\) and \(D^j\) while keeping \(K^j/K^i\) constant, the payoffs to rational players are kept
constant, while the disagreement probability falls (irrational player payoffs increase). Such thinking would seem to be behind the plan to cut the budget deficit as part of the debt ceiling deal. A super committee was set up to find a mutually acceptable deficit cutting plan by a new deadline, without which, cuts would automatically be implemented on defence and medicaid, punishing Democrats and Republicans. The hope was that such punishments would spur agreement, sadly they didn’t.

In this simple setup, non-partisan voters (who care only for the common value outcome) who may not know who ‘should’ have conceded, may then have incentives to strongly punish both sides following any disagreements. Partisan voters (who additionally care for the division of surplus) may on the other hand find it optimal to lower any disagreement costs for their representative in the hope of raising their chance of winning the war of attrition.

Increasing $\bar{z_j}$ does not change $T^*$, but both lowers the immediate concession rate of agent i, and so $(1 - F^j(s))$ for $s < T^*$, and further increases the fraction of agents who never agree so increasing $(1 - F^j(s))$ after $T^*$, and so unambiguously raises the probability of disagreement. Increasing $\bar{z_i}$ has two effects, firstly it results in a lower $T^*$ and higher initial concession probability by player j lowering $(1 - F^j(s))$, but secondly raises $(1 - F^i(s))$ beyond this new $T^*$. Taking explicit derivatives it can be shown that for small $\bar{z_i}$ the net effect is for reduced disagreement, but for large $\bar{z_i}$ the effect is certainly positive. However, certainly increasing $\bar{z_i}$ and $\bar{z_j}$ so that the initial concession probability remains the same, must increase disagreement.

Again non-partisans may thus dislike more intransigent representatives, with the possibility of disaster they bring, but partisans can see electing agents with a greater likelihood of irrational behavior as increasing their chance of winning the war of attrition; a low irrational probability spells (almost) certain defeat. Of course, all of this analysis takes demands as given, an agent with only a mild reputation for stubbornness, but making reasonable demands, may secure higher payoffs than one who is highly likely to be irrational but makes excessive demands. Other things equal, higher demands lower the chance of winning the war of attrition.

Given a unique equilibrium for any given subgame with given demands and probabilities of irrationality, Abreu and Gul show how to solve this model for the multi-type case. They first assume a single irrational type of player 1, and show that a rational player 2 has a unique best response, mixing between imitating particular behavioral types (all types imitated give the same payoff). Given this, one can solve for the best response of player 1, which implies a unique equilibrium distribution of outcomes. The structure of our game is identical to theirs at the demand choice stage with payoffs from choices given by the equations in 2, and so their proof works here as well.

The predictions of this model, in particular which types are imitated and the payoffs obtained are highly dependent on the arbitrary initial specification of irrational types probabilities, are their relative frequencies, $\pi'(\alpha')$. We therefore look at the case where the probability of irrationality is small, indeed vanishingly small, and the type space available is rich in the sense that there is a type who makes a demand arbitrarily close to $\alpha'$ for any demand in $(0,1)$. Such a setting gives clear predictions for outcomes allowing a understand better understanding of the forces at work in the model.

In direct analogy with the limit results of Abreu and Gul, we find that with a sufficiently rich type space as the probability of irrationality becomes small, we have a determinate outcome to
bargaining. More formally:

**Proposition 1.** Fix $\varepsilon > 0$. Let $G_n$ be a sequence of bargaining games in which $z^i_n, z^j_n \to 0$ such that $L > \frac{\varepsilon}{z^i_n} > 1/L$ and for all $x \in (0, 1)$ there exists an irrational type $|\alpha^i - x| < \frac{\varepsilon}{2}$, then agent $i$ can obtain a payoff within $\varepsilon$ of:

$$\alpha^{i*} = \max\{\min\{\frac{D^i - D^j + 1}{2}, 1\}, 0\} \quad (4)$$

The proof of this is not difficult. Notice first that for a given a sequence of bargaining games with $z^i_n \to 0$, an initial demand choices from a finite set, we know that (taking a subsequence if necessary), the probability of those choices must converge to some limit.

Suppose then agent 1 initially demands more than $\alpha^{1*} + \varepsilon$ with positive probability in this limit, any lower demand may be immediately accepted to secure agent 2’s payoff. Given positive limit probability of this choice, when agent 1 makes it, for any counterdemand made by agent 2 there is a bound on the posterior probabilities of irrationality for the two agents, that is for some $L_2$ we have $L_2 > \frac{\varepsilon}{z^i_n}$ sufficiently close to the limit.

Consider then the payoff to agent 2 from mimicking a type within $\frac{\varepsilon}{2}$ of $\alpha^2_{z^i_n}$. This guarantees that agent 2 concedes at a faster rate than agent 1 in the war of attrition ($\frac{L^i}{K^i} > 1$; $\alpha^{i*}$ is calculated so that $\frac{L^i}{K^i} = 1$). Agent 1’s initial mass concession following in a war of attrition is given by:

$$1 - c^1 = 1 - \min\{1, \frac{z^i}{\varepsilon} (\frac{L^i}{K^i} - \frac{L^j}{K^j})\}.$$  

Given the bound on $L_2 > \frac{z^i}{\varepsilon}$, we then have $1 - c^1 \leq 1 - L_3(z^i)^{1 - 1} \frac{L^i}{K^i}$ sufficiently close to the limit for some fixed $L_3$, which can clearly be made arbitrarily close to 1, for small enough $z^1$, ensuring that agent 2’s payoff is greater than $\alpha^2_{z^i_n} - \varepsilon$.

Similarly by demanding arbitrarily close to $\alpha^{1*}$ agent 1 can guarantee that he concedes faster than agent 2 following any counteroffer made with positive probability in the limit, $\frac{L^i}{K^i} > 1$. In either case, the probability of agreement at time 0 must converge to 1.

Notice that this limiting payoff depends exclusively on the disagreement payoffs agents receive, agents who care more about the common ground issue are in a weaker position. Remember both agents would like an agreement on at least the common ground (Disagreement payoffs $-D' \leq 0$). If then the common ground issue was separated from the surplus division issue, the common ground issue would pass uncontroversially. Moreover, in this case (for small probabilities of irrationality) each agent would obtain almost a 50/50 share of the surplus (issues on which agents disagree).

Detaching the common ground issue, must then improve one agent’s total payoff, while harming the other’s. This clearly helps explain why the Republicans were so keen to tie raising the debt ceiling to deficit reduction, while the Democrats wanted to separate the two issues. Disagreement on raising the debt ceiling, would have been far more costly to the party controlling the White House, than to representatives recently elected by the Tea Party, even though both were in fact in favor of the measure necessary to avoid default. Such a model then shows how the power to determine which issues are debated together, can massively affect bargaining outcomes.
3 Repeated bargaining

While the above model has interesting political economy implications in its own right, the model is a straightforward adaptation of Abreu and Gul. Repeated bargaining was impossible in the Abreu Gul setup however, because no set end date can arise for the end of bargaining (irrational agents bargain forever).

The simplest adaptation of the above model, which can nonetheless generate useful insights, extends bargaining to two issues sequentially. In this case the distribution of deadlines times beyond the start of bargaining is given by $G_1$ and $G_2$ distributed on $[0,T]$ and $[T,2T]$. We assume that up to date $T$ agents do not know whether the effective deadline has passed, unless they attempt to make an agreement (by one of them conceding to the other’s demand). The surplus available on issue 1 is 1, while on issue 2 it is $B > 0$. Disagreement payoffs for agent $i$ are $-D'_i$ and $-D''_i$. Agents discount future payoffs at rate $\delta^i$.

Agent 1 demands a share of the surplus $\alpha^1$ at time 0 at the start of bargaining on the first issue, whereupon agent 2 can either immediately accept this offer or make a counteroffer $\alpha^2$ such that $\alpha^1 + \alpha^2 > 1$. Ostensibly these can be revised at any time in $(0,T)$, however, again due to Coasean logic, the only relevant decision for a rational agent is what time to concede. Irrespective of the outcome in bargaining on the first issue, at time $T$, agent 1 makes the demand $\alpha^*_1$, a share of surplus available on the second issue, which agent 2 may accept, or make a counteroffer.

Irrational types are again characterized by demands $\alpha^i \in C^i \subset (0,1)$ who on any issue demand a surplus share of $\alpha^i$ and accept nothing less. Such agents can be committed to always getting their own way, $\alpha^i \approx 1$, splitting the difference on areas of disagreement, $\alpha^i \approx 0.5$, or being generous in the face disagreement, $\alpha^i \approx 0$, but do not modify this behavior on different issues. Given this, if agent $i$ makes a demand $\alpha^i$ on issue one, only to concede to $j$, then she must also concede to $j$ in bargaining on issue 2 if $j$’s reputation remains intact.

This is the krux of the problem faced by Obama, highlighted by Krugman. After initially demanding the highest earners be excluded from an extension of the Bush tax cuts, facing the threat of all expiring (in the middle of a recession) he backed down, clearly demonstrating that he was not committed to demands in the face of threats to the economy. This necessitated concession to the Republican’s demands on a whole string of issues, including the debt ceiling deal.

Following $i$’s acceptance of $j$’s offer on issue 1 then, $j$ can guarantee the demand of his imitated irrational type on issue 2, by the Coase conjecture, however, there is in fact an additional complication to the continuation game. If $j$ subsequently reveals rationality himself, before bargaining on issue 2 we have a continuation game in which both agents are revealed to be rational, which should yield to a Rubinstein [1982] bargaining solution, in which $j$ may do better than his imitated committed type. This possibility was not relevant for one shot bargaining.

We discuss the implications of a model where an agent revealing rationality during bargaining on issue 1, results in their opponent making more aggressive demands on future bargaining issues in a later section, 8. For now we abstract from this concern and assume there is a small possibility that a rational agent $i$ who concedes to agent $j$’s offer $(1 - \alpha^j)$ on issue 1 becomes committed to obtaining the share $\alpha^*_2 = (1 - \alpha^j)$ in bargaining on issue 2. All rational agents will therefore imitate this demand on the second issue, and this will be immediately accepted by any
j agent who initially demanded $\alpha^j$.

We also assume irrational agents with demands $\alpha^2 < (1 - \alpha^1)$, who immediately accept the offer on the first issue, subsequently become committed to the demand $\alpha^2_2 = (1 - \alpha^1)$ for bargaining on issue 2. This is an unimportant assumption, which will not affect the game as the probability of irrationality becomes small but allows us to write simpler equations.

Given this, if the division $(\alpha, 1 - \alpha)$ is provisionally agreed in bargaining issue 1, whether this agreement came before the actual deadline or not (remember that agents do not learn whether the deadline has passed until they try to implement an agreement), this division will be immediately agreed to on issue 2 as well. Bargaining on the two issues then, by the force of reputational concerns, becomes similar to a situation in which there is bargaining over one (bigger) issue. The assumptions also thus ensure that the probability of disagreement on the first issue must always be higher than that on the second issue; if agents were ever to disagree on the second issue, they must have disagreed on the first one.

And so, arguments analogous to those highlighted above, entail a continuous time war of attrition structure, extremely similar to that described above, with no positive mass acceptance except at time 0 (and there by at most one agent), and continuous concession on an interval $(0, T^*)$ as yet undefined. Given this, we continue to allow strategies to be defined by $\mu^1_{\alpha^1, \alpha^2}$ and $F^2_{\alpha^2, \alpha^1}$ where the concession distributions are defined on $[0, 2T]$.

In equilibrium then, j’s utility from conceding at time $t \in (0, T)$ is given by:

$$U^i(t, \sigma^i|\alpha) = \int_0^t (1 + \delta_i B)\alpha^i(1 - G_1(s)) + (\delta_i B\alpha^i - D_i^j)G_1(s)dF^i(s) + (1 - F^i(t))[(1 - G_1(t))(1 + \delta_i B)(1 - \alpha^j) + G_1(t)(\delta_i B(1 - \alpha^j) - D_i^j)]$$

While concession at $t \in (T, 2T]$ gives:

$$U^i(t, \sigma^i|\alpha) = \int_0^T (1 + \delta_i B)\alpha^i(1 - G_1(s)) + (\delta_i B\alpha^i)G_1(s)dF^i(s) - (1 - F^i(T))D_i^1$$

$$+ \delta^i \int_T^t B\alpha^i(1 - G_2(s)) - D_2^1G_2(s)dF^i(s)$$

$$+ \delta^i(1 - F^i(t))[(1 - g_2(t))B(1 - \alpha^j) - G_2(t)D_2^j]$$

Again we begin by focussing on the single type case. Differentiating the first of these yields:

$$\frac{f^i(t)}{1 - F^i(t)} = \frac{g_1(t)}{1 + \delta_i B - G_1(t)}K^j_1$$

While the second gives:

$$\frac{f^i(t)}{1 - F^i(t)} = \frac{g_2(t)}{1 - g_2(t)}K^j_2$$
Where:

\[
K_j^1 = \frac{1 - \alpha^j + D_i^1}{\alpha^j + \alpha^j - 1} \\
K_j^2 = \frac{(1 - \alpha^j) + D_j^2/B}{\alpha^j + \alpha^j - 1}
\]  

(5)

Compared to the one period model above for given demands the concession rate is initially slower, reflecting that being conceded to is substantially more valuable. The ratio of concession rates on \((0, T)\) compared to bargaining on a single issue is

\[
\frac{1 + \delta_i B - G_1(t)}{1 - G_1(t)}
\]

which reflects how much slower agents are to concede given the linkage between the issues, and the larger effective surplus thus disputed. Whereas the concession rate converged to \(\infty\) by time \(T\), it is now uniformly bounded in that interval. However, concession rates can only continue to be positive so long as rational agents exist to concede. Again, there is some time \(T^* < 2T\) such that only irrational agents remain in the war of attrition. The implied distribution functions given above are:

\[
1 - F^j(t) = \begin{cases} 
   c^j (1 - G^j(t))^{K_j^1} & \text{for all } t \leq T, T^* \\
   c^j (1 - G^j(t))^{K_j^2} & \text{for all } t = T, t \leq T^*
\end{cases}
\]

Let \(1 - \hat{F}^j(t)\) be defined by the above equation when \(c^j = 1\), then we can as before define \(T^j\) by

\[
\bar{z}^i = 1 - \hat{F}^j(T^*)
\]

Then \(T^* = \min\{T^1, T^2\}\) and

\[
c^j = \frac{1 - \hat{F}^j(T^j)}{1 - \hat{F}^j(T^*)}
\]

Clearly a higher initial probability of irrationality helps win the war of attrition once again. And again, other things equal having a higher concession rate, results in a smaller exhaustion time, and helps secure initial mass acceptance. Additionally, increased patience, a higher \(\delta^i\), increases the prize on offer to agent \(i\) from the second phase, and therefore induces agent \(j\) to concede more slowly, to the advantage of \(i\), in ensuring her exhaustion time occurs before \(j\)’s.

Equilibrium in the multi-type case then follows from the arguments of Abreu and Gul as outlined before. The most notable contrast to the one shot bargaining game is that concession rate ratios may change midway through the war of attrition. On the first issue agent 1 may have an advantage, while on the second it may be with agent 2. This variation arises because the amount each agent values the common ground outcome can vary \((D_1^1 \neq \frac{D_2}{B})\) while their bargaining posture over their share of the surplus issues must remain the same. Agents are unable to moderate their demands to recognize weakness on particular issues; to do so would reveal rationality (we relax this assumption later).

Abstracting from differences in discount rates, during bargaining over the first (respectively
second) issue, concession rates follow the ratio:

\[
\frac{K_i}{K_j} = \frac{1 - \alpha_i + D_i^j}{1 - \alpha_j + D_j^i}
\]

\[
\frac{K_i^*}{K_j^*} = \frac{(1 - \alpha_i)B + D_j^i}{(1 - \alpha_j)B + D_i^j}
\]

The changing concession rates can imply that the time 0 concession probability is no longer monotonic in a type’s demand, for a fixed probability of irrationality. In particular in contrast to the single type case, for a given demand by agent 1, a higher demand by agent 2 may elicit more mass concession at time 0. It may then pay to be greedy.

To see this, first consider when \(\alpha_i + \alpha_j - 1\) is small. In this case both agents concede at a high rate. In this case the war of attrition is likely to be over quickly and to grant initial mass acceptance to the agent, taken to be agent i, who concedes faster during phase 1 (before T), \(\frac{K_i}{K_i^*} > 1\). However, secondly, notice that the concession rate of agent i, which is proportional to \(K_i\), is decreasing in agent j’s initial demand. So agent j, by demanding more causes agent i to concede at a slower rate to ensure indifference. Intuitively as the difference between j’s demand and i’s offer grows \((\alpha_i + \alpha_j - 1)\), being conceded to becomes much more valuable than conceding and so the rate of concession must decline to make j indifferent between conceding at two different points.

Of course it is also true that \(K_i\) is decreasing in j’s demand and indeed \(\frac{K_i}{K_i^*}\) can only increase. However, this is not all that matters as the higher demand may push the war of attrition into phase 2 (after time T), where j may have an advantage \(\frac{K_j}{K_j^*} < 1\). This advantage, may be of a sufficient size that \(T^i > T^j\) and so the more aggressive agent j secures initial mass acceptance in the war of attrition.

For a numerical example of this, consider initially the types \(\alpha_1 = \alpha_2 = \frac{5}{9}\) with prior probability of irrationality \(z^1 = z^2 = \frac{1}{16}\). Further let \(D_1^1 = D_2^1 = 0\) and \(\delta_1 = \delta_2 = B = 1\) and finally \(D_2^1 = \frac{2}{9}, D_2^2 = 0\). This is then the game in which the surplus size is the same in each period and agents are patient. There is no common ground on the first issue, but agent 1 cares deeply about securing the common ground outcome on issue 2, while agent 2 is indifferent. In this case \(K_i^* = K_i^2 = 4\), which is all we need to know to see that the war of attrition is completed in phase 1, with neither agent conceding with positive probability at time zero, as \(T^i = T^j = T\).

Now consider a similar case, except that while we keep \(\alpha_1 = \frac{5}{9}\), we increase \(\alpha_2 = \frac{8}{9}\). Although agent 2 is now at a disadvantage in stage 1, it forces the war of attrition into phase 2, in which agent 2 has a considerable advantage, despite his higher demand. Solving for the relevant variables we find \(K_1^i = K_2^1 = 1\), while \(K_2^2 = \frac{1}{2}\) and \(K_2^3 = 2\). Solving implicitly for \(T^1, T^2\) we have \(1 - G_2(T^1) = \frac{1}{8}\) while \(1 - G_2(T^2) = (\frac{1}{2})^\frac{1}{3} \approx \frac{1}{2}\). So \(T^1 > T^2\), entailing initial mass concession by 1, and a higher payoff for 2 than in the previous case, despite a more aggressive demand.

In a repeated bargaining scenario even for a fixed probabilities of irrationality, the likelihood of President Obama conceding may then have increased the more outrageous the bargaining position of Republican representatives.
The above possibility was created by the fact that for a relatively large probability of prior irrationality, concession rates may entail that the war of attrition does not spend much time in phase 2, while a larger demand, actually ensured a large fraction of rational agents were still present in phase 2. As the probability of irrationality decreases however, much of the war of attrition must take place in phase 2 in any case. In fact, agents’ relative strength in phase 2, \( \frac{k_i^2}{k_j^2} \), will (almost) entirely determine the bargaining outcome for small prior probabilities of irrationality.

The reason for this is that concession rates on issue 1 are bounded for any given demands (given the presence of some amount of money to be bargained over tomorrow), and so the conditional probability of irrationality given no concession up to time \( T \) (and no time 0 concession) is \( \tilde{z}^i (\frac{1+\delta_j B}{\delta_j B})^{k_i^1} \), which is arbitrarily close to zero for small \( \tilde{z}^i \). Moreover, given bounds on \( \tilde{z}^i \) we have bounds on the ratios of conditional probabilities at time \( T \) as well. But, then we are back in the one shot setting for bargaining at time \( T \), where we know that the agent who reaches his exhaustion time first is entirely determined by \( \frac{k_i^2}{k_j^2} \), for such small probabilities of irrationality.

Stronger claims can clearly be made in the multiple type case. We are led to strong equilibrium payoff bounds, as the probability of irrationality becomes small, which depend entirely on the bargainers inherent strength on issue 2 (how much they care about the common ground outcome). Specifically:

**Proposition 2.** Agent \( i \) can obtain payoff in the rich type limiting equilibrium of \( i \):

\[
V^{i*} = (1 + \delta B)\alpha^i 2
\]

where:

\[
\alpha^i 2 = \max\{\min\{\frac{B + D_j^i - D_i^2}{2B}, 0\}, 1\}
\]

The proof is very much similar to that of the earlier payoff bound and thus not presented here.

As discussed above however, the result follows from the fact that for any given incompatible demands, the continuous concession rates on \((0,T)\) can create only a bounded difference in agents posterior probabilities of irrationality at time \( T \).

At first glance the result is somewhat remarkable. Notice that this is completely independent of how important the common ground payoffs to the two parties is on the first issue, and the respective discount rates. Certainly then, for such slight perturbation settings agents’ should then seek to have issues on which they are in an inherently strong position debated last (an issue on which they do not care much about the outcome, but their opponent does), attesting to the great importance of agenda setting power, determining the order in which issues are debated.

To see the true perversity of the result, consider an issue on which in the first period, agent \( i \) has cares much more than agent \( j \) about securing the common ground result \( D_j^1 >> D_i^1 \), but has an advantage on the second issue \( D_j^2 \geq D_i^2 + B \) then in this case, in the limiting case of complete

---

1 By this we mean: Fix \( \varepsilon > 0 \). Let \( G_n \) be a sequence of bargaining games in which \( \tilde{z}_{i_n}, \tilde{z}_{j_n} \rightarrow 0 \) such that \( L > \tilde{z}_{i_n}^2, \tilde{z}_{j_n}^2 > 1/L \) and for all \( x \in (0, 1) \) there exists an irrational type \( |\alpha' - x| < \varepsilon \), then agent \( i \) can obtain a payoff within \( \varepsilon \) of...
rationality we should expect agent 1 to obtain all the surplus available on both bargaining issues, no matter how small the surplus available on the second issue is (we can have $B << 1$).

For larger prior probabilities of irrationality, agenda setting power will clearly still be important in determining outcomes, but it is unclear whether it is advantageous to have a particular issue debated first or second. Indeed for fixed but small probability of irrationality, taking $B \to 0$ the outcome must ultimately come to be entirely determined by strength on the first bargaining issue, with payoffs in the rich type equilibria converging to $\alpha_i^* = \max\{\min\{\frac{1+D_j^* - D_i^*}{2}, 0\}, 1\}$.

As explained above, our assumptions on reputations can be viewed as forcing bargaining over two issues into bargaining over one (bigger) issue. We have also argued that differences in disagreement payoffs can be viewed as analogous to differences in discount factors in the in the open ended Abreu Gul setup. Given this, in some sense our repeated bargaining game is similar to single issue open ended bargaining, where agents discount factors change at a predetermined time $T$ during the bargaining process. A similar result easily translates to that setup, in that only agents discount factors from $T$ onwards will determine the outcome of bargaining, no matter how large $T$ is, as the probability of irrationality becomes arbitrarily small.

Despite the fact that the finding that only the tail end of bargaining matters in such two sided reputational settings could be derived elsewhere and is perhaps no surprise to those familiar with Abreu Gul type models, we believe it deserves further comment.

Intuitively, the result resembles an extreme form of backward induction. Whoever is strongest at the end of bargaining will in fact be strong at the start. This certainly has a flavor of the chain-store paradox, although in fact reputational types actually give this result, rather than destroy it; subgame perfect payoffs from a Rubinstein type bargaining game will very much depend on the parameters involved in bargaining over the first issue.

Indeed, the result is very much bound up with the reputational types present, for which such a finding might in some sense be considered typical. For instance, consider the simplest case of one-sided uncertainty of commitment in open ended bargaining. A standard proof of the Coase conjecture in this setting first shows that there are at most a finite number of periods $T$ by which the agent who is known to be uncommitted must concede even when expecting all the surplus in bargaining from an uncommitted opponent. Given this the known uncommitted agent must have a low continuation value close to $T$, while the potentially committed agent must have a high one. But given this the uncommitted agent must concede to her opponent’s demand considerably earlier than time $T$; strength at the end of bargaining is translated into strength at the start. With two sided reputations, and no restriction on the times at which agents can speak, it should not then come as a surprise that strength at the end of bargaining matters disproportionately.

A separate but interesting feature of the result is that even though agreement must occur with probability approaching 1 on both issues, the outcome is not necessarily efficient in the limit. The reason is that in any equilibrium, the surplus sharing rule is the same in both periods. However, if $\delta_i' \neq \delta_j'$ then there are clearly more efficient equilibria where the agent $i$, for whom $\delta_i' < \delta_j'$ takes a larger share of the surplus on issue 1, and a smaller share on issue 2.

This may not seem surprising given the type space laid out in which agent’s demands on different issue are effectively forced to be the same. However, in fact the result is more inherent to

17
the need to build (and maintain) a reputation. In the next section, we make a minor extension to
the above model by allowing for more sophisticated types who do make different demands on
different issues but find this inefficiency still holds.

In the section following that, 5, we find that when there is uncertainty about bargaining condi-
tions on future issues, inefficiency may occur even when there is no difference in agents time
preference; there is disagreement on early bargaining issues even in the limiting case of com-
plete rationality.

4 Sophisticated irrational types

Previously we have supposed that agents built reputations by demanding that they get their own
way on a fraction $\alpha$ of all issues on which they disagree with their opponent (the surplus). For
agent $i$ to demand that he gets completely his own way on issue 1 (where he is in a strong
position) but demand only 20% of the surplus on issue 2 (where he is in a weak position),
might seem to show that he isn’t really an agent committed to any particular demands at all, as
such behavior seems to show too much sophistication/rationality. Agents who *always* insist on
going their own way, or *always* insist on splitting the difference between the parties positions,
just seem more realistic behaviorally.

However, there would seem to be some occasions where agents, particularly political parties,
do pick their battles more wisely, in particular moderating demands on issues where they are
weak, in order not to be forced to back down on a position they claimed to be committed to.
Intuitively, this is because of a fear that demonstrating a willingness to back down will only
compromise them in bargaining on other issues, where they may otherwise be strong.

In this section, we briefly analyze the possibility of sophisticated committed types who com-
mit to different demands on different issues. We show that, in the limiting case of complete
rationality, it leads to a strong degree of independence of outcomes between bargaining issues.
Moreover, although both issues are agreed almost without delay in the limit, the result remains
inefficient when agents rates of time preferences differ. In some sense this is entirely because
bargaining on the two issues becomes independent, while efficiency would require a deal which
involved a dependency between outcomes.

To save space, we do not present the model in its full formality here, however, the analysis
should nonetheless be sufficient for explanatory purposes. Bargaining protocol is as follows:
agent 1 at time 0 makes a demand $\alpha^1_1$ of the surplus on issue 1, and simultaneously, a demand
$\alpha^1_2$ on bargaining issue 2. Whereupon agent 2 can either immediately accept this demand in full,
accept the demand on issue 1 and make a counterdemand $\alpha^2_2$ on issue 2, accept the demand on
issue 2 and make a counterdemand $\alpha^2_1$ on issue 1, or make a counterdemand on both issues $\alpha^2_1$, $\alpha^2_2$. Provisional agreement on issue 2 may be reached before the distribution of deadlines has
begun, although the agreement will not be implemented until time $T$ and so payoffs from that
agreement remain discounted by $\delta^i$.

Irrational types are now indexed by $\gamma \in C^i$. Type $\gamma'$ demand at least $\alpha^{i_1}_1(\gamma') \in (0, 1)$, $\alpha^{i_2}_2(\gamma') \in (0, 1)$ on issues 1 and 2 respectively and accept nothing less. The probability of being of type $\gamma'$
conditional on being irrational is given by: $\pi_{\gamma'}(\gamma')$. Given such types, Coasean arguments then
ensure that a rational agent who mimics the demand of type $\gamma^i$, only to accept an opponents offer $(1 - \alpha^i_j(\gamma^j))$ on issue 1, reveals rationality, and thus must concede immediately to j’s demand on bargaining issue 2 as well.

Given this, a strategy for player 1, $\sigma^1$, is defined by a probability distribution $\mu^1(\gamma^1)$, and $\mu^2(\gamma^2)$ on $C^i$ of initial demands by rational agents and a collection of cumulative distributions $F^i_{\gamma^1,\gamma^2}$ on $[0, 2T]$, where $F^i_{\gamma^1,\gamma^2}(t)$ is the total probability of player i conceding to player 2 (on all issues such that $\gamma^i(\alpha^i_k) + \gamma^j(\alpha^j_k) > 1$) by time t following demands $\gamma^i$ and $\gamma^j$.

Given initial demands, we have associated conditional probabilities of irrationality. If initial demands are compatible on either issue 1 or issue 2, with payoffs on those issues secured, then equilibrium in the subgame for bargaining on the remaining issue follows a structure much as in section 5. The only novel case is when demands are incompatible on both issues, $(\alpha^i_1 + \alpha^j_1 - 1 > 1$ and $\alpha^i_2 + \alpha^j_2 - 1 > 1$.

Much as in section 8, arguments ensure that mass concession can only occur by at most one agent at time 0, and must otherwise be continuous for both agents on some interval $(0, T^*)$. Moreover, concession rates that make agent i indifferent to conceding at different points on $(0,T)$ are:

$$\frac{f^j(t)}{1 - F^j(t)} = \frac{g_1(t)}{1 + \delta^i B' - G_1(t)} K^j_1$$

While indifference on $(T,2T)$ requires:

$$\frac{f^j(t)}{1 - F^j(t)} = \frac{g_2(t)}{1 - g_2(t)} K^j_2$$

Where:

$$K^j_1 = \frac{1 - \alpha^j_1 + D^j_i}{\alpha^i_1 + \alpha^j_1 - 1}$$

$$K^j_2 = \frac{(1 - \alpha^j_2) + D^j_2/B}{\alpha^i_2 + \alpha^j_2 - 1}$$

$$B' = B \frac{\alpha^i_2 + \alpha^j_2 - 1}{\alpha^i_1 + \alpha^j_1 - 1} > 0$$

As in section we can then calculate potential exhaustion times $T^i$, and set $T^* = \min\{T^i, T^j\}$ before calculating mass concession at time zero to ensure both agents reach conditional probability 1 of irrationality at time $T^*$.

The utility of agents in the multi-type case is given by a formula that is in all essentials the same as that of equation 2, and the same arguments for the existence and uniqueness of the equilibrium go through.

Given such a setting, with a rich enough type space, and a small probability of rationality, we then have payoffs determined by the following result.
Proposition 3. Agent $i$ can obtain payoff in the rich type limiting equilibrium of:

$$V^{i*} = \alpha^{i*}_1 + \delta^i B \alpha^{i*}_2$$

where for $B_1 = 1$, $B_2 = B$ we have:

$$\alpha^{i*}_k = \max\{\min\left\{\frac{B_k + D^j_k - D^i_k}{2B_k}, 0\right\}, 1\}$$

Again the proof of this is very much the same as those analyzed previously.

As in section , it is clear that these payoffs are not necessarily efficient. If $\delta^i < \delta^j$ then by having agent $i$ receiving more of the surplus on issue 1, and agent $j$ receiving more of the surplus on issue 2, both agents can strictly increase their payoffs above the bound outlined above. However, in fact, such higher payoffs cannot be achieved, even with our sophisticated committed types.

Suppose agent $\delta^2 < \delta^1$, then for an efficient outcome agent 1 should demand a share of the surplus on issue 2 strictly larger than $\alpha^{1*}_1$ in the limit, while demanding strictly less than $\alpha^{1*}_1$ on issue 1. However, if agent 1 makes the demand $\alpha^{2}_1 = \alpha^{1*}_1 + \epsilon$ with positive probability in the limit then any demand $\alpha^{2}_1 < \alpha^{2*}_1 + \epsilon$ will ensure that $K^2_2 > K^1_1$, and so agent 1 must concede to agent 2 at time 0 on both issues, with probability approaching 1, irrespective of how aggressive agent 2’s demand on issue 1 is. Given this, agent 1 will never make such a demand in the limit.

Equally, if $\delta^2 > \delta^1$ then agent 1 should demand a strictly smaller share of the surplus on issue 2 but a larger one on issue 1, say $\alpha^{1}_1 = \alpha^{1*}_1 + \epsilon$. But in this case, agent 2 can simply accept the strictly larger share of the surplus on issue 2, diffusing that issue, and so a counterdemand such that $\alpha^{2}_1 < \alpha^{2*}_1 + \epsilon$, will ensure acceptance with probability 1 at time 0 in the limit. Given this, agent 1 will never make such a demand.

Inefficiency here is created, despite the ability to imitate types who moderate their demands on different issues. The reasoning is that a reputation for being committed in one period, still remains tied to a reputation in another period, even though the demands involved are different. And so making mutually beneficial proposals by agent 1 that take advantage of the difference in discount factors, simply allow 2 to take advantage of any generosity on issues where 1 is in a strong position, while leaving incentives for 2 to be even more aggressive on issues where 1 is in a weak position.

5 Uncertainty and disagreement

This section highlights the possibility for inefficiency through disagreement even for arbitrarily small possibilities that agents are committed to their demands. Uncertainty about the future bargaining environment provides incentives for agents to initially make aggressive, mutually incompatible demands, and only back down from them when the resolution of future uncertainty is unfavorable.

Intuitively developing an aggressive reputation for demanding that he gets things (mostly) his own way, is needed for the agent to take advantage of possible future scenarios, in which he is in
a strong position (his opponent faces high disagreement costs). If the revelation of uncertainty is unfavorable, and the future bargaining issue is one which leaves the agent in a weak position, he retains the option to concede, and so the incentive to moderate his demand on account of this situation is slight.

Given incompatible extreme demands agents have incentives to wait, not conceding at all during bargaining on early (known) issues, because waiting may (given a favorable revelation of uncertainty), lead one’s opponent to concede with positive mass, while concession on early issues sacrifices a reputation for being committed, necessitating immediate concession on future bargaining issues, even when in an inherently strong position.

We focus on the case when there are only two possible issues which may be debated at time $T$. On the first issue the surplus available is again 1, with disagreement payoffs $-D_i^1$, and with a deadline distributed on $[0,T]$. However, at time $T$, issue 2a will be debated with probability $p_a$, and surplus $B_a$ while issue 2b (surplus $B_b$) is debated with probability $(1 - p_a)$. Disagreement payoffs are given by $-D_i^2_a$ and $-D_i^2_b$. Let $(B_a p_a + B_b(1 - p_a)) = B$. Deadlines are distributed $[T,2T]$.

We adopt the assumptions on committed agents given in section 8, in particular a given type demands $\alpha^i$ of the surplus on all issues on which are to be debated. Standard arguments make it immediately clear that at most one agent can concede with positive probability in equilibrium at time 0, and similarly, at most one can concede with positive probability following any particular revelation of uncertainty at time $T$, while there must be no mass concession at any time other than 0 and $T$. Furthermore, if either agent concedes with positive probability in $(0,T)$ then both must concede continuously on the entire interval $(0,T)$. Similarly, if either concedes with positive probability on $(T,2T)$, then both must concede continuously on the entire interval $(T,T_2\omega)$, where $\omega \in \{a, b\}$ and $T_2\omega$ depends on the revelation of uncertainty.

We let strategies be described by the collections $F^{j}_{\alpha^i,\alpha^j,\omega^j}$ (where $F^{j}_{\alpha^i,\alpha^j,a}(s) = F^{j}_{\alpha^i,\alpha^j,b}(s)$ for $s < T$), and initial choices $\mu_1$ and $\mu_2$ as described previously.

We again initially focus on the case of a single type of each agent (or the subgame following initial demands) and prior probabilities of irrationality $\overline{\pi}$. Then given initial demands $\alpha^i + \alpha^j > 1$ we have agent i’s utility from conceding at time $t < T$ given by:

$$U^i(t, \alpha^j|\alpha) = \int_0^t (1 + \delta_i B \alpha^j (1 - G_1(s)) + (\delta_i \alpha^i - D_i^1)G_1(s)dF^j(s)$$

$$+ (1 - F^j(t))[(1 - G_1(t))(1 + \delta_i B)(1 - \alpha^i) + G_1(t)(\delta_i B (1 - \alpha^i) - D_i^1)]$$
While concession at $t \in (T, 2T]$, conditional on $\omega$ gives:

$$
U^j(t, \sigma^j|\alpha, \omega) = \int_0^T (1 + \delta_i B_{i\omega})\alpha^j(1 - G_1(s)) + (\delta_i B_{\omega_j}\alpha^j)G_1(s)\,dF_{i\omega}^j(s) \\
- (1 - F_{i\omega}^j(T))D_i^j \\
+ \delta^j \int_T^t B_{i\omega_j}\alpha^j(1 - G_{2\omega}(s)) - D_i^2 G_{2\omega}(s)\,dF_{i\omega}^j(s) \\
+ \delta^j (1 - F_{i\omega}^j(t))(1 - G_{2\omega}(t))B_{\omega_j}(1 - \alpha^j) - G_{2\omega}(t)D_i^2]
$$

Conditional mass concession of agent $j$ at time $T$ after the revelation of uncertainty is given by:

$$
\frac{F_{i\omega}^j(T) - \lim_{t \to T} F_{i\omega}^j(t)}{1 - \lim_{t \to T} F_{i\omega}^j(t)} = 1 - c_{2\omega}
$$

This means that in equilibrium $1 - c_{2\omega}$ is the initial concession in a one shot (second issue) bargaining environment with posterior probability of irrationality $\bar{z}^j = \frac{\omega}{1 - \lim_{t \to T} F_{i\omega}^j(t)}$, and revelation of uncertainty $\omega$.

To help analyze the equilibrium structure we first focus on the subgame following the deadline of issue 1, and the revelation of uncertainty. For any probabilities of irrationality, fixed bargaining positions and disagreement payoffs the outcome of stage 2 is uniquely determined by the one shot bargaining single type case analyzed in section . Let the outcome of these bargaining games be given by $V_{2\omega}^j(y)$, and let $V_2^j(y) = p_\alpha V_{2\omega}^j(y) + (1 - p_\alpha)V_{2\beta}^j(y)$ for any given probabilities of irrationality $\gamma$.

**Type 1 equilibrium:**

Now clearly if $\delta^j V_2^j(\bar{z}) - D_i^j \geq (1 + \delta B)(1 - \alpha^j)$ for both agents then there are incentives to wait out the entire first stage, sacrificing any surplus available there and incurring the cost of disagreement. Clearly this can be an equilibrium outcome when for given initial priors, different agents win the war of attrition for a given prior, in scenarios a and b.

**Type 2 equilibrium:**

When $\delta^j V_2^j(\bar{z}) - D_i^j > (1 + \delta B)(1 - \alpha^j)$, but $\delta^j V_2^j(\bar{z}) - D_i^j < (1 + \delta B)(1 - \alpha^j)$, then there is an equilibrium similar to type 1, in which there is initial mass concession by agent $j$, followed by waiting by both agents waiting on $(0, T)$. Notice that in this case $\bar{z}^j = \frac{\bar{z}}{c_i}$, and so we clearly have $V_2^j(\bar{z})$ increasing in $(1 - c^j)$, while $V_2^j(\bar{z})$ is decreasing.

So long as $\delta^j [(1 - \bar{z}^j)\alpha^j + \bar{z}(1 - \alpha^j) > (1 + \delta B)(1 - \alpha^j)$, by continuously increasing initial mass concession by agent $j$, we must eventually obtain $\delta^j V_2^j(\bar{z}) - D_i^j = (1 + \delta B)(1 - \alpha^j)$. If at this $c^j$, we still have $\delta^j V_2^j(\bar{z}) - D_i^j \geq (1 + \delta B)(1 - \alpha^j)$ then clearly we have an equilibrium.

**Type 3 equilibrium:**

When the probability of rationality is relatively large then there is an equilibrium at the other extreme, when the war of attrition is entirely over before time $T$. This case effectively follows an equilibrium similar to that in the previous section when the probability of irrationality was large. In this case each agent concedes from 0 to $T^j_1 \leq T$ at a rate to make her opponent indifferent.
between conceding at time \( t \) or \( (t + \Delta) \), that is:

\[
\frac{f^j(t)}{1 - F^j(t)} = \frac{g_i(t)}{1 + \delta B - G_i(t)}K^j_1
\]

(8)

Where \( K^j_1 \) is given in equation 5. Initial concession rates are again derived to ensure that the two agents hit probability 1 of irrationality at the same time \( T^*_1 \).

**Type 4 equilibrium:**

There is only possibility left for an equilibrium which doesn’t follow structures: 1, 2 or 3. The structure of this equilibrium requires continuous concession on some interval \((0, T^*_1] \) with \( T^*_1 \in (0, T] \) before no-concession on the interval \((T^*_1, T) \), and then concession again on some interval \([T, T^*_2) \), depending on the revelation of uncertainty. At time \( T^*_1 \) both agents must be indifferent between conceding immediately or waiting until time \( T \), and seeing if the revelation of uncertainty favors them (hoping to benefit from mass concession at that point). Letting \( \overline{z}(t) \) be agent i’s equilibrium level of irrationality at time \( t < T \), the time \( T^*_1 \) must satisfy the following equation for both agents.

\[
(1 - G(T^*_1))(D^i_1 + (1 - \alpha^i)) = \delta^i[V^i_2(\overline{z}(T^*_1)) - B(1 - \alpha^j)]
\]

(9)

**Lemma 1.** An equilibrium of type of type 1-4 exists for any given prior probabilities of irrationality, and initial demands, moreover it is unique.

**Proof.** See appendix.

Given the existence of a unique equilibrium in this subgame, an equilibrium for the full multi-type case can again be derived in a similar manner to that of Abreu and Gul.

To characterize the equilibrium in the limiting case of complete rationality in the subgame following demands \( \alpha^i + \alpha^j > 1 \), we first notice that we cannot have an equilibrium of type 3. This is for the same reason that the war of attrition always endured until the second phase of bargaining in section 1; the rate of concession is bounded during the bargaining over issue 1 and so is insufficient to exhaust a mass of rational agents that approaches 1.

We distinguish two cases, the first in which agent i’s concession rates dominate that of agent j \( (K^i_{2a} > K^j_{2a} \) and \( K^i_{2b} > K^j_{2b} \)) and the second of separate advantages \( (K^i_{2a} > K^j_{2a} \) and \( K^i_{2b} < K^j_{2b} \)).

Conditional on reaching stage 2 therefore with arbitrarily small conditional probabilities of irrationality, whose ratio is bounded, we have i’s continuation payoff in the case of dominance approaching \( V^i_j = \alpha^j B \), and j’s approaching \( V^j_j = (1 - \alpha^j)B \). In the case of separate advantages this is without loss of generality \( V^i_j = p_aB_a(1 - \alpha^j) + (1 - p_a)B_b(1 - \alpha^j) \) and \( V^j_j = p_aB_a(1 - \alpha^i) + p_aB_b\alpha^j \).

In the case of separate advantages it may be that both agents have an incentive to not concede before time \( T \) even if their opponent doesn’t, \( \delta^i[V^i_j - B(1 - \alpha^j)] > -D^i_j \), hoping that the revelation of uncertainty will be favorable at date \( T \). In which case we have a type 1 equilibrium with delay in the limit.\(^2\) Assuming this is not the case then we define \( T^*_1 \) to be the minimum time at which

\(^2\)In what follows we ignore the non-generic limiting case of \( \delta^i[V^i_j - B(1 - \alpha^j)] = -D^i_j \)

23
agent \(i\) is indifferent to immediate concession to \(j\) or waiting until period \(T\) to receive \(\delta^i V_j\). More precisely let it be defined as:

\[
\arg\min_{T^i \in [0,T]} [(1 - G_i(T^i))(1 - \alpha^i) + D^i + \delta^i (B(1 - \alpha^i) - V^i_j)]
\]

The cost to waiting is \((1 - G_i(T^i))(1 - \alpha^i) + D^i\), while the benefit is \(\delta^i [V^i_j - B(1 - \alpha^i)]\). We have already dealt with the case where \(T^i = T^i_j = 0\), in which there is a type 1 equilibrium in the limit. In the case where i’s concession rates dominate agent j’s then we have \(T^i_j = T\). Given this, we can characterize the limit equilibrium.

**Lemma 2.** For fixed irrational types, such that \(\alpha^i + \alpha^j > 1\), \(T^i_j > T^j_i\), then agent \(j\) must concede to \(i\) with probability approaching 1 at time 0, if the probabilities of irrationality converge to zero at the same rate.

**Proof.** See appendix. \(\square\)

We now prove by example that uncertainty about future disagreement payoffs can lead to inefficiency in the limit, assuming a rich type space.

Consider a particular game where \(D^i_j = D^2_i = 0\), \(B_a = B_b = 6\), \(D^1_a = D^2_a = 6\), \(D^1_b = D^2_b = 0\), while \(\delta^i = \delta^j = 1\) and \(p_a = 1/2\). In this case the first issue is pure surplus, while the second issue is high surplus, but is bound up with a common ground issue about which at most one of the agents cares a lot about. The weight towards the second bargaining issue may seem large, but simplifies the analysis.

Consider first the setting where agents have exactly one type, \(\alpha^1 = 1 - \epsilon\), and \(\alpha^2 \in (\epsilon, 1 - \epsilon]\). Prior probabilities of irrationality are given by \(z^j\). Then should the game proceed to debating issue 2 with such prior probabilities of irrationality, with uncertainty revealed as \(\omega = b\), then by the analysis of the section 5 we have \(T^i_b < T^j_b\) so long as:

\[
\frac{\log z^2_b}{\log z^1} < \frac{K^1_{2b}}{K^2_{2b}} \Rightarrow \frac{1 - \alpha^1 + D^2_b/B_b}{1 - \alpha^2 + D^1_b/B_b} = 1 + \epsilon
\]

On issue b agent 2 cares a lot about the common ground outcome, and so the fraction on the right hand side is always strictly greater than 1. Given this, if the probabilities of irrationality go to zero at the same rate, then agent 1 must win the war of attrition on issue b. Moreover the concession probability is given by:

\[
1 - \zeta_{2b}^2 = 1 - \frac{\bar{z}_2}{(1 - G(T^*))} = 1 - \bar{z}^2(\bar{z}^1) \frac{K_{2b}^2}{K_{1b}^2}
\]

Which must converge to 1. Equally however, in this case any demand by \(\alpha^2\) will secure agent 2 victory with probability approaching 1 on issue a. For given \(\alpha^2\) and sufficiently small \(\epsilon\) and \(\bar{z}^j\) we then have:

\[
\delta^j V^j_2(\bar{z}) - D^j_2 \approx \alpha^2 B_a p_a = 3\alpha^2 > (1 + \delta^2 B)(1 - \alpha^1) \approx 0
\]

And so agent 2 will not concede before time \(T\) (unless agent 1 concedes with mass approaching...
we need \( \alpha \) to be so we need either:

We consider whether it is possible that the continuation equilibrium is not of type 1. For this \( \alpha \), then we need \( \alpha \) with immediate acceptance after a countero.

Consider then agent 2’s best response. By imitating his maximal type 1, this contradicts our established need for \( \alpha \) since the lower bound on his payoff this is the case the maximum payoff of almost 3 is imitating her highest possible type.

In conclusion, in the limit, the unique equilibrium has both agents always selecting \( \alpha^i = 1 - \epsilon \), there is always disagreement on the first issue debated (no concession), and at least 1/7 of the
possible surplus is wasted (including the common ground $1/13$ of the possible payoff is wasted). This illustrates that committing to a bargaining posture when you may be in either a strong or weaker position in the future, can give incentives to make extreme demands. If you are in a strong position tomorrow, you want to be committed to a large initial demand, if it turns out you are in a weak position, you can concede to the other agent anyway. Moderating your demand to allow a mutually beneficial efficient solution, merely causes your opponent to accept your moderate demand tomorrow when it turns out he is in a weak position, but continue with his extreme demand, which you must accept, if you are in a weak position.

This possibility of inefficiency through delay for arbitrarily small probabilities of irrationality is in contrast to reputational bargaining results without uncertainty, whether the open ended setup of Abreu and Gul or the model of the previous section. It is somewhat similar to the results of Abreu et al. [2012], where initial asymmetric information about discount factors can lead to inefficiency in the limit, although the mechanism is very different.

Viewing differences in disagreement payoffs as differences in discount factors in the open ended Abreu and Gul bargaining, the comparable result is for uncertainty about agents discount factors after time $T$ to cause inefficiency and delay in agreement. The results of this section should again easily generalize to that setting.

Despite the above discussion, it is clearly not true that there must always be delay in the limit whenever there is uncertainty about future bargaining parameters. We do not seek to explicitly characterize for which parameters delay will occur in the limit, or agents payoffs, primarily because we have two many free variables to make the exhaustive exercise worthwhile, and believe the above example sufficiently illustrates the point we seek to make.\(^3\) However, it is certainly worth noticing that bargaining parameters during stage 1, such as $D^i_1$ and $\delta^i$, clearly matter for the limiting equilibrium outcome, as they affect agent’s incentives to wait, following incompatible initial demands. This is in contrast to the finding in section 8. The characterization of the equilibrium in the subgame after demands $\alpha^i, \alpha^j$ are made with positive probability in the limit, lemma 2 allows one to solve for the multi-type limiting equilibrium for any given parameters, with sufficient patience.

\section{The weakness of non-stationary demands and first mover (dis)advantage}

In a repeated bargaining situation, when an agent backs down on a previously held demand he reveals that he is not a commitment type, and so he must concede to future opponent demands if that opponent still maintains a reputation, by the Coase conjecture. Previously, we assumed that some rational agents, who concede to their opponents demand $\alpha^i$ become committed to obtaining $(1-\alpha^i)$ on future bargaining issues, which entails a continuation game with immediate agreement on the second issue with surplus shared according to $(\alpha^i, 1-\alpha^i)$. However, one might expect an agent who has shown himself uncommitted to a previous demand, may face even more aggressive demands in the future. Opponents, both rational and committed, may want to take advantage of this revelation of rationality.

\(^3\)We have 10 free variables: $(\delta^i, D^i_1, D^i_2a, D^i_2b)$ for each agent as well as the surplus shares $B_a, B_b$
A particular case of interest is when agent j’s complete information Rubinstein bargaining payoff on issue 2 (from a limiting discrete time game) is higher than his committed type’s demand $\alpha_j^i$. Even if committed agents are always attached to stationary demands, if agent i reveals rationality conceding to j’s demand on issue 1, then by revealing rationality himself at the start of bargaining on issue 2, j can secure the higher Rubinstein continuation payoff, which should be instantly agreed.

Another possibility, is that some commitment types do change their demands on future bargaining issues, when they realize they are faced with a weak opponent. And this will of course lead rational types to mimic this behavior, which does not revealing rationality, but secures immediate concession from the uncommitted opponent. This indeed seems to be Krugman’s fear regarding the debt ceiling deal; having conceded to Republican demands, he fears they will only be encouraged to make even more outrageous demands in the future.

In this section, we setup a model to capture the above features. We favor simple assumptions that lead to transparent solutions, and argue that these simplifications should not matter, at least in the limiting case of complete rationality.

Unsurprisingly we find that bargaining is affected by this change as this creates incentives for agents to delay agreement (fearing concession, and revelation of rationality, will induce more aggressive demands in future). Despite these incentives, in a rich type space, we show that agreement on all bargaining issues occurs immediately as the probability of rationality converges to zero, and so for equal rates of time preference, there is no inefficiency.

Furthermore, we show that the possibility that player i will raise demands in the future, will in fact benefit player j, increasing his payoff. The reason is that the threat of higher future demands makes agent j much more reluctant to concede first, and so strengthens his bargaining position in any war of attrition.

Moreover, when both agents threaten to increase future demands following an opponent’s revelation of rationality, we find that it can matter greatly which agent makes the initial demand. The reason for this is an asymmetry between the two agents, when agent 1 makes an initial demand, he may immediately face high costs to backing down (revealing rationality), while agent 2 can accept an initial demand from agent 1 without revealing rationality (he may simply be a committed agent with a lower reservation value). In different circumstances, this may favor making a demand first or second. The importance of which agent makes an initial demand is in marked contrast to the standard repeated bargaining game analyzed in section 8, or other reputational bargaining models in the style of Abreu and Gul, of which we are aware.

The game structure remains the same as in section , however, we now make different assumptions on irrational types. Irrational types of agent 1 are again indexed by $\alpha_1 \in C_1$. A type $\alpha_1^i$ makes an initial demand $\alpha_1^i$ on bargaining issue 1. If this offer is immediately accepted (without a counter-offer being made), then this type makes a similar demand $\alpha_1^i$ on the second bargaining issue. If however, a counteroffer is made such that $\alpha_2^i + \alpha_1^i > 1$ but agent 2 subsequently does concede before time T (revealing weakness/rationality) agent 1 makes a demand on the second issue $\phi^i(\alpha_1^i) \geq \alpha_1^i$. If agent 2 does not concede on issue 1, then such a type will again demand $\alpha_1^i$ on bargaining issue 2.

Irrational types of agent 2 are similarly indexed by $\alpha_2^i \in C_2$. If agent 1 makes a demand $\alpha_1^i$ and agent 2’s irrational type is such that $\alpha_2^i \leq 1 - \alpha_1^i$, then he immediately accepts the offer
and subsequently becomes committed to the demand \( \alpha_{2'} = (1 - \alpha_1) \) on future bargaining issues. If however \( \alpha^2 > 1 - \alpha_1 \), he makes the counterdemand \( \alpha^2 \). If agent 1 accepts 2’s counterdemand before time \( T \) however (revealing weakness/rationality), then a committed agent 2 raises his demand on the second issue to \( \phi^2(\alpha^2) \geq \alpha^2 \). If time \( T \) is reached without agent 1 revealing rationality, such a committed agent 2 will counterdemand \( \alpha^2 \) on the second issue as well. Conditional on irrationality, agent i’s probability of being of type \( \alpha_i \) is again \( \pi^i(\alpha_i) \).

The justification for these types is primarily as a simplification of more complicated models. The key dynamic captured is that backing down on a demand you initially made reveals a lack of commitment, which opponents may take advantage of in the future. For instance when seeking to model the case where a rational agent i reverts to her Rubinstein bargaining demand on issue 2, \( \phi^i(\alpha') \), we make the simplification that all types of agent i revert to the demand \( \phi^i(\alpha') \geq \alpha' \). This will be a good approximation when the probabilities of insistent types is small.

Focussing on the subgame after agent 2 has made a counteroffer such that \( \alpha_i + \alpha_j > 1 \), our ‘standard arguments’ ensure that at most one agent can concede with positive probability in equilibrium at time 0, and similarly at most one can concede with mass at time \( T \), while there must be no mass concession at any time other than 0 and \( T \). If either agent concedes with positive probability in \((0,T)\), then both must concede continuously on an interval \((0,T_1)\). Similarly if either agent concedes with positive probability in \((T,2T)\), then both must concede continuously on an entire interval \((T,2T)\). Strategies can thus be described by \( \mu_1, F_{\alpha_1,\alpha_2}, \) and \( \mu_2, F_{\alpha_1,\alpha_2} \), where the concession distributions are defined on \([0,2T]\) as before.

The action of this new model is clearly all before time \( T \). If neither agent has conceded before time \( T \) the continuation game must follow much the same structure as in section 5. Agent i’s utility from conceding at \( t \in (0,T) \) is given by:

\[
U^i(t,\sigma^j|\alpha) = \int_0^t \alpha^j(1 - G_1(s)) - D^i_1G_1(s) + \delta^i B \phi^i(\alpha^i)dF^i_j(s)
+ (1 - F^i(t))\left[\delta^i B(1 - \phi^i(\alpha^i)) + (1 - G_1(t))(1 - \alpha^j) - G_1(t)D^i_j\right]
\]

To calculate the concession rate of j to i that makes rational agent i indifferent between conceding at \( t' \) or \( t' + \Delta \), for \( t' < T \), we need:

\[
\frac{f^i(t)}{1 - F^i(t)} = \frac{g(t)}{1 - G(t) + \delta^i B} K^i_1
\]

Where:

\[
K^i_1 = \frac{1 - \alpha^j + D^i_1}{\alpha^i + \alpha^j - 1}
\]

\[
B' = B \frac{\phi^i(\alpha^i) + \phi^j(\alpha^j) - 1}{\alpha^i + \alpha^j - 1} \geq B
\]

When characterizing the structure of the equilibrium for this model, we assume that we do not have the non-generic case of \( (1 - \alpha^j) + D^i_1 = \delta^i B(\phi^i(\alpha^i) - \alpha^j) \) which could allow for non-uniqueness in equilibrium outcomes, and would requires the description of an entirely new equilibrium type, for little explanatory gain. For various later results we make even further
assumptions about the generic nature of the game’s parameters. We also assume \( \phi'(<\alpha_1) > \alpha \) for at least one agent, or else the equilibrium follows that of section 2.

Type 1 equilibrium

Again we let \( V_j(y) \) be the payoff agent i will obtain in the subgame where issue 2 is debated alone with probabilities of irrationality \( y \). (If \( y = \bar{y} \) these probabilities are the initial priors). Now suppose we have \((1 – \alpha') + \delta' B(1 – \phi'(<\alpha_1)) \leq –D_i' + \delta' V_j'(\bar{z}) \) for both agents, then we have an equilibrium in which both agents wait until time T rather than concede immediately. This is possible because while for one agent we must have \( V_j'(\bar{z}) = B(1 – \alpha') \) we can have \( \phi'(<\alpha_1) > \alpha \). At least one agent waits even though he is certain to lose the ensuing war of attrition at stage 2, because the terms of that war which will be less disadvantageous conditional on waiting.

Type 2 equilibrium

Suppose \((1 – \alpha') + \delta' B(1 – \phi'(<\alpha_1)) \leq –D_i' + \delta' B(1 – \phi'(<\alpha_1)) > –D_i' + \delta' V_j'(\bar{z}) \), so agent i would prefer to wait until time T rather than concede immediately (even if she loses the war of attrition upon getting there), but agent j would prefer to concede immediately rather than wait to receive \( V_j'(\bar{z}) \geq B(1 – \alpha') \). Then there is an equilibrium in which there is positive initial concession by agent j, followed by no concession on \((0, T)\), followed by a war of attrition, with mass concession at time T by agent i.

To find the level of initial concession for agent j, notice that \( \bar{z}_j' \) and hence \( V_j'(\bar{z}_j) \) is continuously increasing in j’s initial concession probability. Then we merely need to solve:

\[
\argmin_{c_j \in [1 – \bar{z}_j, 1]} [(1 – \alpha') + D_i' + \delta' B(1 – \phi'(<\alpha_1)) – \delta' V_j'(|\bar{z}_j)|]
\]

If the minimizer is the lower bound on \( c_j = 1 – \bar{z}_j \), then all rational j agents immediately exit, otherwise the minimum of 0 is achieved and some rational agents are induced to wait until time T, for the prospect of concession by agent i there. Agent i may wait even when certain that he faces an irrational opponent, in order to avoid that irrational opponent taking advantage of i’s revelation of rationality.

Agent j in this situation concedes with positive mass at time 0, and so can obtain the payoff of at most: \((1 – \alpha') + \delta' B(1 – \phi'(<\alpha_1)) \). For parameters of this equilibrium to arise, we certainly needed \( \phi'(<\alpha_1) > \alpha \), and so we have an early indication of how agent j’s threat to raise his demand after the revelation of i’s rationality hurts j in equilibrium.

Type 3 equilibrium

If there is no type 1 or 2 equilibrium, then we must have: \((1 – \alpha') + \delta' B(1 – \phi'(<\alpha_1)) \geq –D_i' + \delta' V_j'(\bar{z}) \) for some agent j, and furthermore for agent i we have \((1 – \alpha') + \delta' B(1 – \phi'(<\alpha_1)) \geq –D_i' + \delta' B(1 – \alpha') \), then in any equilibrium we must have positive concession by both agents on the interval \((0, T)\).

Rearranging the above inequalities we have: \((1 – \alpha') + D_i' > \delta' B(\phi'(<\alpha_1) – \alpha) \). More generally,

---

4To see the ‘problem’ in the non-generic case, we notice that agent i is indifferent between conceding at time 0 and waiting until the second issue at time T to concede only to concede immediately there. Suppose then we also have: \((1 – \alpha') + D_i' < \delta' B(\phi'(<\alpha_1) – \alpha') \), then irrespective of the conditional probability of rationality of agent i, j will certainly wait until time T. But in which case, agent i could concede at time 0 with probability ranging from 0 to \( (1 – \bar{z}_j) \), and otherwise wait until time T.
For given initial demands, and probabilities of rationality, an equilibrium of the Lemma 3.

committed opponent, seeking to benefit from a more generous o
above model exists and is unique.

If however \( T = 1 \) at time \( T \) the latter case involves agent i waiting to concede between \( (T, T) \) equals 1 at time \( T \) equals 1 at time \( T \) means concession at time zero by agent j so that his conditional probability 1 of irrationality that both agents reach conditional probability 1 of irrationality at the same time \( T \) by agent i at time \( T \) that makes agent j indi
he loses the war of attrition at time \( T \). In this case either we either find feasible mass concession to concede, while agent j prefers to quit immediately rather than to wait until time \( T \) (assuming

\[ (1 - G_1(T))(D_1 + (1 - \alpha^j)) = \delta^i B(\phi^i(\alpha^i) - \alpha^i) \]  

(11)

\( T_{1m} \) is then the time agent i would be indifferent to conceding at time \( T_{1m} \) or waiting until \( T \) given that agent j does not concede to him henceforth. We then define \( T_{1m}^* = \min\{T_{1m}, T_{1m}^j\} > 0 \). Notice that \( T_{1m}^* < T \), so long as for one agent we have \( \phi^i(\alpha^i) - \alpha^i > 0 \).

Given the continuous concession rates on \((0,t)\) defined above, we also define a potential stage 1 exhaustion time for agent i, given by: \( T_{1e}^i \):

\[ arg\min_{T_{1e}^i \in (0, T)} \left( \frac{\delta^i B' + (1 - G_1(T_{1e}^i))}{1 + \delta^i B^i} \right)^{k^i} - z^i \]

And similarly define \( T_{1e}^* = \min\{T_{1e}^i, T_{1e}^j\} > 0 \). Now suppose \( T_{1e}^* \leq T_{1m}^* < T \), then we have an equilibrium in which both agents continuously concede on \((0, T_{1e}^*)\), and initial concession is made by the agent i if \( T_{1e}^i > T_{1e}^j \), so that the two agents’ probability of irrationality reach 1 at the same moment. We call this a type 3 equilibrium. No agent has an incentive to wait until time \( T \) rather than concede on this interval.

**Type 4 equilibrium**

When type 1, 2 or 3 equilibrium do not exist then we must have \( 0 < T_{1m}^* < T_{1e}^* \). In this case, there must be continuous concession on \((0, T_{1m}^*)\) in equilibrium, at which point neither agent has exhausted his stock of rational agents.

If \( T_{1m}^* = T_{1m}^i \) then we have an equilibrium in which there can be no positive mass concession at time \( T \). Initial concession (at time 0) is then calculated to make each agent’s ultimate exhaustion time, in bargaining game 2 equal, \( T_2^i = T_2^* \).

If however \( T_{1m}^* > T_{1m}^i \), then a rational agent i is indifferent between waiting until time \( T \) in order to concede, while agent j prefers to quit immediately rather than to wait until time \( T \) (assuming he loses the war of attrition at time \( T \)). In this case either we either find feasible mass concession by agent i at time \( T \) that makes agent j indifferent between conceding at \( T_{1m}^* \) and \( T^* \), or we have mass concession at time zero by agent j so that his conditional probability 1 of irrationality equals 1 at time \( T_{1m}^i \). In the former case, we must then ensure mass concession at time zero so that both agents reach conditional probability 1 of irrationality at the same time \( T_{1m}^* > T \). The latter case involves agent i waiting to concede between \( (T_{1m}^*, T) \) even though certainly facing an committed opponent, seeking to benefit from a more generous offer at time \( T \).

**Lemma 3.** For given initial demands, and probabilities of rationality, an equilibrium of the above model exists and is unique.

---

5If we do not have this condition the equilibrium is exactly identical to section
Proof. See Appendix.

Given a unique equilibrium once a demand and counterdemand have been made, this case can then again be solved for the multi-type model following Abreu and Gul. Again, each option chosen with positive probability, must have an equal payoff.

One notable feature of such an equilibrium is if agent 1 makes a demand of $\alpha^i$, agent 2 may accept this demand with positive probability in equilibrium, so receiving $(1 + \delta^i B)(1 - \alpha^i)$, but also make a counterdemand $\alpha^j > (1 - \alpha^i)$ with positive probability, and agent 1 following that counterdemand may concedes with positive probability. In some sense then for a given demand by agent 1, both agents concede with positive probability at time 0. This may occur because if there is no time 0 mass acceptance, then agent 2’s continuation payoff falls below $(1 + \delta^2 B)(1 - \alpha^i)$, because of the possibility of more aggressive future demands should he concede (having made the counteroffer).

We now characterize the equilibrium in the subgame following demands $\alpha^i + \alpha^j > 1$ when the probability of irrationality converges to zero for both agents at the same rate. Suppose $K_2^i > K_1^i$, and there is no concession on $[0, T)$ player i’s payoff converges to $\delta^i V^i_2(\bar{z}) = \delta^i B \alpha^i$, while j’s payoff must be $\delta^j B(1 - \alpha^i)$. So if we have $(1 - \alpha^i) + D_1^i \leq \delta^i B(\phi^i(\alpha^i) - \alpha^i)$ and $(1 - \alpha^j) + D_1^j < \delta^j B(\phi^j(\alpha^j) + \alpha^j - 1)$ then we have a type 1 equilibrium in the limit, which exhibits delay. Payoffs converge to $\delta^j B \alpha^j$ and $\delta^i B(1 - \alpha^i)$.

Still assuming $K_2^i > K_1^j$, suppose we have $(1 - \alpha^i) + D_1^i \leq \delta^i B(\phi^i(\alpha^i) - \alpha^j)$ but $(1 - \alpha^i) + D_1^i > \delta^j B(\phi^j(\alpha^i) - \alpha^j)$, then we must have a type 2 equilibrium. Furthermore, in this type 2 equilibrium agent j must concede to i at time 0 with probability approaching 1. To see this, notice that for $(1 - \alpha^j) + \delta^j B(1 - \phi^j(\alpha^j)) = -D_1^j + \delta^j V^j_2(\bar{z}_j)$ we need initial concession by agent j, so that $\bar{z}_j = \bar{z}_j^i$ converges to 0 at a slower rate than $\bar{z}_i^i$ (agent i does not concede). For this to be the case we must certainly have $c_j \to 0$. And so agent i’s payoff converges to $\alpha^i + \delta^j B \phi^j(\alpha^j)$, while agent j’s converges to $(1 - \alpha^j) + \delta^j B(1 - \phi^j(\alpha^j))$.

If alternatively $K_2^j > K_1^i$ and $(1 - \alpha^j) + D_1^j \leq \delta^j B(\phi^j(\alpha^j) - \alpha^j)$ but $(1 - \alpha^i) + B(1 - \phi^i(\alpha^i)) \geq -D_1^i + \delta^j B \alpha^i$. Agent j then has no incentive to wait until time T even if certain to lose, while agent i has an incentive to wait even if certain to lose. In this case clearly agent j must concede to i with probability 1 at time zero.

Clearly, we cannot have a type 3 equilibrium in the limit, for given concession rates potential exhaustion times $T_{1e}$ on issue 1 must converge to, and hit time T, for any given demands and so $T_{1e}^i > T_{1m}^i$.

For parameters which do not fit those of a type 1 or 2 equilibrium then, we have a type 4 equilibrium. To characterize this equilibrium in the limit first notice that $T_{1m}^i$ does not change as the probabilities of irrationality do. First we consider the case when $T_{1m}^i = T_{1m}^i$. In this case we clearly have everything determined by rates of concession on issue 2. If $K_2^i > K_2^j$ then with probability approaching 1 agent i wins the war of attrition at time 0.

Now consider if $T_{1m}^j > T_{1m}^i$ and:

\[(1 - G(T_{1m}^i))(D_1^j + (1 - \alpha^i)) \geq \delta^j B(\alpha^j + \phi^j(\alpha^j) - 1)\] (12)
In this case, given the proscribed concession rates on \((0, T^*_{1m}]\), then in order that agent \(j\) reaches probability time \(T^*_{1m}\) with probability of irrationality equal to 1, we must have concession by \(j\) at time 0 with probability approaching 1.

Suppose the inequality is reversed, then the war of attrition certainly enters time 2 (for sufficiently low initial prior probability of irrationality). Notice then that the conditional concession at time \(T\) by agent \(i\), \(c^i_2\), is fixed by parameters (it does not depend on the probabilities of irrationality), see equation 15, in which case, suppose that \(K^i_2 > K^j_2\) then we have: 
\[
1 - c^i_2 \geq 1 - M^z_2(\bar{z}^j)^{-K^i_2/K^j_2} \text{ which converges to zero.} 
\]
Obviously we have the reverse result if \(K^j_2 < K^i_2\).

Clearly there are incentives in the above model following given demands and counterdemands, for neither agent to concede on the first bargaining issue, if doing so induces a more aggressive future demand. However, in a multi-type model, agents have incentives to make demands which can be accepted immediately and avoid such delay, which will effectively burn money for little purpose. The result below shows that in fact, this effect means that for a rich enough type space, there should be immediate agreement in the limit. In this proof we make a technical assumption on the genericity of the parameters of our finite type space.\(^6\)

**Lemma 4.** For generic parameters, agents must reach agreement on both issues with probability approaching 1 in the rich type limiting equilibrium.

**Proof.** See Appendix. \(\square\)

The reasoning behind the proof is that as argued above, inefficiency can only come through type 1 equilibrium in limit, and in such an equilibrium agent 1 must lose the war of attrition that only begins at time \(T\). But for this to be so we need \(K^j_2 > K^j_1\) (otherwise agent 2 would have optimally conceded at time 0, for a payoff of \((1 + \delta^2)B(1 - \alpha^1)\) instead of \(B(1 - \alpha^1))\). However, by making a less aggressive demand, agent 1 could however have ensured that \(K^j_2 < K^j_1\), and guaranteed himself a higher limiting payoff even if this did require waiting.

**The cost of aggression towards uncommitted agents**

We now give an example of how the possibility of agent \(i\) raising his demand after the revelation of \(j\)’s rationality, can harm agent \(i\), and benefit agent \(j\), in the limiting case of complete rationality.

We initially consider the case where only one agent increases his demand after the revelation of rationality in his opponent. This setup can then model the case where if agent \(i\) reveals rationality, then agent \(j\) by revealing rationality himself can obtain his Rubinstein bargaining payoff on issue 2. Clearly for any given incompatible demands (\(\alpha^i + \alpha^j > 1\)), this could be the case for at most one player.

Consider then a game with \(D^i_1 = D^i_2 = 0\) and \(\delta^i = 1\) and \(B = 2\), so there is bargaining over a dollar today and two dollars tomorrow no discounting between, and no common ground issues.

\(^6\)Without this one needs to worry about boundary cases and make additional assumptions on the form of the \(\phi^i\) functions; arbitrary \(\phi^i\) functions can otherwise essentially eliminate the richness of the type space and potentially allow for delay in the limit. We assume for all irrational types we do not have \(K^1_2 = K^2_2\), and for agent 2 in particular there are no types \(\alpha^2, \alpha^2\) such that \(\alpha^2 + \delta^2B\phi^2(\alpha^2) = \delta^2Ba^2\).
However, suppose the Rubinstein bargaining payoff on issue 2 is arbitrarily close to 2, for player 2 (because player 2 makes offers much more frequently in the game whose continuous time limit we approximate). We then model this as $\phi^i(\alpha^j) = 1 - \varepsilon$, while $\phi^i(\alpha^1) = \alpha^1$.

The analysis of section showed that if $\phi^i(\alpha^j) = \alpha^i$ for both agents, in the limit both agents guarantee almost half of the surplus available. Suppose now that agent 1 makes the offer of $\alpha^1 >> 0.5$ with positive probability in the limit. For agent 2 accepting this offer secures a payoff of $3(1 - \alpha^1)$. In the game of section, agent 2 could make the counteroffer $\alpha^2 \approx \alpha^1$, while maintaining $K_2^2 > K_1^2$ to secure immediate acceptance with almost probability 1, for a payoff of almost $3\alpha^1$.

Now, however any counterdemand will lead to the following calculation for agent 1, by conceding immediately he obtains $1 - \alpha^2 + 2(1 - \phi^i(\alpha^2)) = 1 - \alpha^2 + 2\varepsilon$. However, if agent 1 has not conceded prior to time 2, agent 2 cannot be convinced that agent 1 is uncommitted, and so must demand $\alpha^2$ again on issue 2 to preserve his own reputation for irrationality. Given this, waiting until time $T$, secures agent 1 an expected payoff of at least $2(1 - \alpha^2)$. If no agent concedes until time $T$ then, agent 1 will obtain the payoff of $2(1 - \alpha^2)$ which is greater than $1 - \alpha^2 + 2\varepsilon$ for $\alpha^2 < 1 - 2\varepsilon$.

For such counterdemands then, a rational agent 1 cannot be induced to concede, for fear of what will happen if he does. In this case we must be in an equilibrium of type equilibrium 1 or 2. If agent 2 makes a counterdemand to $\alpha^1$ such that $\alpha^2 > \alpha^1$ then $K_2^2 < K_1^2$ and the payoff to waiting is $2(1 - \alpha^1)$, which is certainly less than the payoff to agent payoff to immediate concession $3(1 - \alpha^1)$. However, by making a demand an arbitrarily small amount below $\alpha^1$ she obtains the payoff from waiting $2\alpha^1$.

So now we come back to the optimal initial demand of agent 1. When demanding $\alpha^1$ he knows that agent 2 can guarantee herself a payoff of almost $2\alpha^1$. In order then to induce agent 2 not to make a counterdemand agent 1 must ensure that the payoff to acceptance is greater than the payoff to waiting, and so: $3(1 - \alpha^1) \geq 2\alpha^1$, which means $\alpha^1 \leq 3/5$. But then making a demand arbitrarily close (from below) to $\alpha^1 = 3/5$ agent 1 can secure a payoff of almost $9/5$, while demanding more can secure a payoff of at most $2(1 - \alpha^1) \leq 4/5$. So ultimately in equilibrium there is (almost) no delay, but agent 1 obtains noticeably more than half of the surplus available.

Player 2’s threat to demand his Rubinstein bargaining payoff, following 1’s revelation of rationality, makes it very costly for agent 1 to back down on his initial demand, indeed making his commitment not to concede on the first bargaining issue credible. Given a demand by agent 1, agent 2 realizes a counteroffer cannot induce 1’s acceptance before time $T$, and so faces the problem of either accepting the 1’s offer or seeing the money available on issue 1 burnt. Agent 1 makes a demand sufficiently generous that 2 has no incentive to see the money burnt.

If alternatively $\phi^i(\alpha^1) = 1 - \varepsilon$ and $\phi^2(\alpha^2) = \alpha^2$, then we can show that agent 1 initially makes the demand $\alpha^1 \approx 3/5$, with agent 2 counterdemanding $\alpha^2 \approx 3/5$ with agent 1 conceding immediately to this counterdemand with probability approaching 1.

The benefit agent j obtained from the threat by i to be more aggressive in future bargaining should j reveal rationality, is in fact much more general than the specific example above. Indeed it is immediately clear that agent j can never be harmed by such aggressive behavior from i:

**Lemma 5.** Suppose $\phi^i(\alpha^j) = \alpha^i$, for all $\alpha^j$, then agent i’s payoff is bounded above by $(1+\delta^i B)\alpha^j$
concede immediately; we cannot have \((1 - \alpha') (1 + \delta B) \leq -D_i + (1 - \alpha') \delta B\). This means there cannot be a type 2 equilibrium in which agent \(i\) receives mass concession at time zero or a limiting type 1 equilibrium following such a demand. Moreover, if there is a type 4 equilibria, we have \(T_{1m}^i = T > T_{1n}^j\), but this combined with \(K_2^j > K_2^i\), by the characterization above ensures that agent \(i\) must concede at time zero with probability approaching 1.

In fact in the appendix we can show something stronger, in that agent \(j\) will strictly benefit from \(i\)'s potential aggression towards agents revealed as rational, with \(i\) facing a comparable cost.

**Lemma 6.** Suppose \(\phi^i(\alpha^i) = \alpha^i\), for all \(\alpha^i\), but for \(\alpha^i \in [\alpha_{2n}^i - \varepsilon, \alpha_{2n}^i + \varepsilon]\) we have \(\phi^i(\alpha^i) > \alpha_{2n}^i + \varepsilon\), then \(\exists \nu > 0\) such that agent \(j\)'s payoff in the limit is bounded below by \((1 + \delta B)(\alpha_{2n}^i + \nu)\), with agent \(i\)'s payoff bounded above by \((1 + \delta B)(\alpha_{2n}^i - \nu)\), in the rich type limiting equilibrium.

The proof relies on showing that agent \(j\) can make a demand slightly larger than \(\alpha_{2n}^i\), so that even if agent \(i\)'s demand is such that \(\alpha^i\) so that \(K_1^i > K_1^j\), and this results in a type 4 equilibrium, agent \(i\) must concede with probability approaching 1 in the limit.

**Why it matters who speaks first**

In situations where both agents have types who make more aggressive demands following an opponent revealing a lack of commitment \(\phi^i(\alpha^i) > \alpha^i\), it can matter greatly, which agent has the right/obligation, to make an initial demand. In some cases an agent would prefer to be first to make a demand, and on other occasions, he would prefer to wait for his opponent to speak first. While it is difficult to characterize which predominates for arbitrary model parameters given too much freedom in the relationship between \(\phi^i(\alpha^i)\) and \(\alpha^i\), we illustrate such forces at work with the help of two examples, one of which favors the first mover, and another the second mover.

An example to demonstrate first mover advantage is extremely similar to that given above with \(D_k^i = 0, \delta^i = 1\) and \(B = 2\), however, now we assume that \(\phi^i(\alpha^i) = \max C^i = 1 - \varepsilon \approx 1\) for both agents. Conditional on an initial demand \(\alpha^i < 1 - 2\varepsilon\), and a counterdemand \(\alpha^2 < 1 - 2\varepsilon\), both agents will wait rather than concede during bargaining over issue 1 for fear that doing so will induce a more aggressive demand in the future. In this case almost any demand/counterdemand will result in the surplus available on issue 1 being burnt. And so even when \(K_2^j > K_2^i\), agent 2 will receive a payoff at most \(2\alpha\) in the limit.

However, immediate acceptance of agent 1’s initial demand does not reveal a lack of commitment by either agent (agent 2 may be a committed agent with a lower reservation value), and so will result in the payoffs of \((3\alpha^1, 3(1 - \alpha^1))\). Given, this agent 1 can take advantage of the threat to burn money by demanding \(\alpha^1 \approx 3 / 5\), which will be accepted with probability approaching 1 in the limit, despite the fundamentals of the situation being entirely symmetric.

There is clearly an intuitive mechanism at work in this example that has a flavor of Stackelberg leadership. When there is a strong reputational cost of backing down, threats to demand a large share of the surplus on all bargaining issues and wait if you do not get it, become credible, and
so even if an opponent would have a similar credible threat not to back down following any counterdemand, that counter-threat will merely result in money being burnt. Given this, staking out a position first has an advantage, as it lets others accept this without losing reputation, and can be made just generous enough that there is no incentive for agent 2 to burn the money.

However, it is not true that there is a first mover advantage in more general settings, rather in some circumstances it may be beneficial to make a demand second.

Consider another example where $\delta^i = 1, B=2, D_0^1 = 0, D_1^1 = 0, \text{but } D_2^2 = 0.4$. In this case we have $\alpha_2^{1*} = 0.6$. Also let $\phi^i(\alpha^i) = 0.75$ for $\alpha^i < 0.75$, and $\phi^i(\alpha^i) = \alpha^i$ otherwise, for both players. In the limiting equilibrium, agent 1 demands $\alpha_1^1 \approx 0.68$ which is the maximum demand to induce agent 2 not to make a counterdemand (which would burn money). This results a payoff for agent 1 of 2.04.

However, if we reverse the situation as to which agent is in the weaker inherent position in bargaining on issue 2, so $D_2^2 = 0$, but $D_1^2 = 0.4$, and so $\alpha_2^{1*} = 0.4$, things are slightly different. Most notably, the more generous an offer made by agent 1 (for $\alpha^1 < 0.5$), the greater the incentives for agent 2 to wait until time $T$ rather than accepting the offer immediately as (assuming no prior acceptance) this more generous offer will be on the table at time $T$ as well and the surplus on offer on issue 2 is larger. A more generous offer can simultaneously decreases the incentive for agent 1 to wait until $T$ as well; when $K_2^1 > K_2^2$ the payoff to waiting is $2\alpha^1$.

The equilibrium involves agent 1 demanding $\alpha_1^1 \approx \alpha_1^{1*} + \frac{1}{30} = 0.43$, and agent 2 counterdemanding $\alpha_2^2 \approx \alpha_2^{2*} + \frac{1}{30} = 0.63$, to which agent 1 concedes with probability approaching 1. In the second bargaining situation agent 2 then demands $\phi^2(\alpha_2^2) = 0.75$, which is immediately accepted. This gives agent 2 a payoff of approximately 2.13 > 2.04.

Any more aggressive initial demand by agent 1 $\alpha_1^{1*} + \epsilon$ will result in counterdemand of almost $\alpha_2^{2*} + \epsilon$ while ensuring that $K_2^1 > K_2^2$, which will require agent 1 to concede straight away, for an even lower payoff. Any more generous offer will, due to the logic outlined in the previous paragraph, increase agent 2’s incentive to wait while lowering them for agent 1. And so it elicits an even more aggressive counterdemand from agent 2, which even though $K_2^1 > K_2^2$, leads to agent 1 conceding immediately.

The reason agent 1 cannot obtain as high a payoff as agent 2 in this scenario, is in some sense precisely the same mechanism that allowed agent 1 a first mover advantage in the previous example. Moving first, agent 1 stakes out a position, from which it is costly to back down, while giving the option to agent 2 to concede without revealing rationality. However, when agent 1 is in a strong position, she would like to use that position of advantage to force agent 2 to reveal rationality, but cannot. By contrast moving second from an inherently strong position, agent 2 can force a rational agent 1 to back down from almost any position he takes up, leaving room for more aggressive demand in the future.

The potential for second mover advantage, may create a simple mechanism for inefficiency and delay in agreement in an extension to the above model. If agents are not required to make offers immediately, and some committed types do delay making an offer, rational agents will seek to imitate this, hoping their opponent will announce first; in the meantime there is a risk of the effective deadline on issue 1 passing without agents even announcing their demands. The technology of such a proposed extension resembles that of Abreu, Pearce and Stachetti (2011); in their model only one agent waits, and inefficiency is created by patient agents attempting to
distinguish themselves from the less patient.

The example also shows that when agent $i$ does has aggressive types $\phi^i(\alpha^i) > \alpha^i$, it can benefit agent $j$ to have types such that $\phi^j(\alpha^j) > \alpha^j$ as well. Suppose for the above game we instead had $\phi^2(\alpha^2) = \alpha^2$, then agent 1 would demand arbitrarily close to 0.48, to which 2 will optimally counterdemand 0.96, for a payoff of only 2.04.

7 A one sided future

In this section we show how bargaining unfolds when only one of the agents has a reputation to defend for a later period. The scenario can correspond to a political economy context where say, an institution like the IMF, has to bargain with many agents over the terms of its lending and support program. However, this setting might in particular be thought to correspond well to a situation in which a seller bargains with one buyer over the price of a perishable good today, after which a second separate buyer, having observed that outcome, attempts to buy a second good the following day.

The results of such a model suggest that, outcomes split two ways depending on the relative level of surplus available in both periods. The long term agent always has greater incentives not to concede to the short term agent because revealing a lack of commitment is very costly as it allows future short term agents to mimic the most aggressive commitment types and be accepted immediately, by the Coase conjecture. This entails in the limiting case of complete rationality, that agent 1 can make an extremely aggressive demand on the first bargaining issue (potentially almost the entire surplus) and this will be accepted with probability approaching 1. However, taking up an aggressive bargaining posture, even without revealing rationality during bargaining on issue 1, puts the long term agent in a weak position in bargaining on issue 2, allowing the second short run agent to make an aggressive counterdemand at date 2, which the long run agent will have to accept.

Given these competing influences, if the surplus available tomorrow is small relative to today, bargaining follows the above pattern with the long term agent adopting an aggressive initial approach, and being conceded to immediately on the first issue only to concede to an aggressive counterdemand tomorrow. If on the other hand the surplus available tomorrow is large relative to today, the long term agent moderates his initial demand (which is again immediately accepted), so as to be in a strong bargaining position tomorrow. Bargaining is then concluded almost immediately tomorrow as well, at a level close to that proposed by the long term agent.

The situation where a long term bargainer faces many future short term opponents is best compared to the situation in which most of the surplus is available on issue 1, and so the long term agent adopts an aggressive bargaining posture. All short term bargainers except the final one, have greater incentives to concede immediately than the long term agent, and so an aggressive initial position will be conceded to immediately with probability approaching one, by all but the final agent.

We model this scenario, as agent 1 being the long run agent, agent 2 being the first short run opponent, and agent 3 the second short run opponent. Bargaining over issue 1 is over a surplus of value 1, and if an agreement is not reached before the (uncertain) deadline in addition to no
surplus being created, disagreement payoffs are \(-D_1^1\), and \(-D_2^2\). Bargaining over issue 2 is over a surplus of value B, with disagreement payoffs are \(-D_1^2\), and \(-D_2^2\). Agent 1 discounts future payoffs by \(\delta^1\).

At the start of bargaining on issue 1, agent 1 makes a demand \(\alpha^1\), which agent 2 can either accept immediately or make a counterdemand \(\alpha^2\). Irrespective of the outcome of bargaining on issue 1, at time T agent 1 demands a share of the surplus on the second bargaining issue, which agent 2 can either accept or counterdemand \(\alpha^3\).

Committed types are again characterized by demands \(\alpha^i \in C^i \subset (0, 1)\) who on any issue demand a surplus share of \(\alpha^i\) and accept nothing less. Conditional on irrationality, agent i’s probability of being of type \(\alpha^i\) is again \(\pi(\alpha^i)\). Coasean logic again ensures this game has a continuous time war of attrition structure following demands \(\alpha^i + \alpha^j > 1\). Moreover concession by agent 1 on bargaining issue 1 ensures that agent 1 must concede immediately at time T to any demand made by agent 3. To simplify the analysis in an unimportant way, we assume that should agent 1 reveal rationality at time 0, all committed agent 3s demand \(\max C^3 = 1 - \epsilon\), the demand a rational agent 3 would always mimic after 1 reveals a lack of commitment. We make the technical assumption \(\max C^i + \min C^j > 1\), so that the most aggressive type of agent i is incompatible with all of agent j’s types.

A strategy for player 1 is then defined by a probability distribution \(\mu_1\) on \(C^1\) of initial demands by rational 1 agent’s and a collection of cumulative distributions \(F^{1^i,\alpha^i}_{\alpha^i,\alpha^i,\alpha^i}\) on \([0, 2T]\), where \(F^{1^i,\alpha^i}_{\alpha^i,\alpha^i,\alpha^i}(t) = F^{1^i}_{\alpha^i,\alpha^i,\alpha^i}(t)\) for \(t < T\). A strategy for player 2, requires a distribution \(\mu^2_{\alpha^1}\) on \(C^2 \cup \{Q\}\) and cumulative distributions on \(F^{2^i}_{\alpha^1,\alpha^2}\) is on \([0, T]\). A strategy for player 3 requires a collection \(\mu^3_{\alpha^1,\alpha^2}\) on \(C^2 \cup \{Q\}\), where \(\mu^3\) describes agent 3’s initial demand at time T given agent 1’s demand and his reputation for irrationality, \(\bar{z}^1\) at time \(T^-\). Conditional probabilities of irrationality after agents’ initial decisions are calculated in the obvious way.

We analyze this game backwards. Given that agent 1 has not revealed rationality prior to time T and given demand to \(\alpha^1\) and \(\alpha^3\), with conditional probabilities of irrationality \(\bar{z}^2\) and \(\bar{z}^3\), the game must proceeds uniquely according to the logic of section 5. This in turn entails a unique best response for agent 3 over \(C^3 \cup \{Q\}\). This in turn enables us to calculate a unique continuation payoff for agent 1 for any given conditional probability of irrationality \(V^1_2(\alpha^1, \bar{z}^2)\). Moreover, given the fact that following any war of attrition started with probability of irrationality, \(\bar{z}^2\) and \(\bar{z}^3\), we have that rational agent’s payoff in that war of attrition are continuously increasing in that \(\bar{z}^2\) it follows that \(V^1_2(\alpha^1, \bar{z}^2) < B\) must be continuously increasing in \(\bar{z}^2\) as well.

Given this, we are ready to analyze the first stage war of attrition. Given demands \(\alpha^1, \alpha^2\), we have agent 2’s payoff from concession on \((0, T)\) given by:

\[
U^2(t, \sigma^1|\alpha) = \int_0^t \alpha^2(1 - G_1(s)) - D_2^2G_1(s)dF^1(s) + (1 - F^1(t))[(1 - G_1(t))(1 - \alpha^1) - G_1(t)D_1^1]
\]
To keep agent 2 indifferent between conceding at time $t$ and $t + \Delta$ then we need:

$$\frac{f^1(t)}{1 - F^1(t)} = \frac{g_1(t)}{1 - G_1(t)} K_1^1$$

Where:

$$K_1^1 = \frac{1 - \alpha^1 + D^2_1}{\alpha^1 + \alpha^2 - 1}$$

Agent 1’s utility from conceding at $t \in (0, T)$ is given by:

$$U^1(t, \alpha^2|\alpha) = \int_0^t \alpha^1 (1 - G_1(s)) - D^1_1 G_1(s) + \delta^1 V^1_2(\alpha^1, \bar{z}^1(t))dF^1(s)$$

$$+ (1 - F^2(t))[(1 - G_2(t))(1 - \alpha^2) - G_1(t)D^1_1 + \delta^1 B(1 - \text{maxC}^3)]$$

For agent 1 to be kept indifferent between conceding at time $t$ and $t + \Delta$ then we need:

$$\frac{f^2(t)}{1 - F^2(t)} = \frac{g_1(t)}{1 - G_1(t) + \delta^1 B'(t)} K_1^2$$

Where:

$$K_1^2 = \frac{1 - \alpha^2 + D^1_1}{\alpha^1 + \alpha^2 - 1}$$

$$B'(t) = \frac{V^1_2(\alpha^1, \bar{z}^1(t)) - B(1 - \text{maxC}^3)}{\alpha^1 + \alpha^2 - 1}$$

Our usual arguments show that at most one agent can concede at time 0, and otherwise concession must be on a continuous interval of $(0, T)$, while the continuation game after $T$ is uniquely determined above.

Given the threat to agent 1 of facing the most aggressive possible counterdemand from agent 3, if he concedes before time $T$, the structure of the equilibrium is in many ways similar to the case analyzed in section 8. Here again revelation of rationality before time $T$, results in aggressive future demands.

**Type 1 equilibrium**

If $\delta^1 V^1_2(\bar{z}^1) - D^1_1 \geq (1 - \alpha^2) + B(1 - \text{maxC}^3)$, then agent 1 has an incentive not to concede during $[0,T)$, even when certain to be facing an irrational agent 2. Given this, all rational type 2 agents must concede to agent 1’s demand at time 0.

**Type 2 equilibrium**

If $\delta^1 V^1_2(\bar{z}^1) - D^1_1 < (1 - \alpha^2) + B(1 - \text{maxC}^3)$ then agent 1 cannot be induced to wait faced with an agent 2 who is certainly irrational. In this case, we must have continuous concession by both agents on some interval on some interval of $(0,T)$. We start describing such equilibria, by first defining a potential exhaustion time for agent 1, $T^1_{1e}$. Given the continuous rate of concession required to make agent 2 indifferent when conceding at time $t$ or $t + \Delta$, we can readily define $\bar{z}^1 = [1 - G_1(T^1_{1e})]K_1^1$. Assuming no initial mass concession by agent 1 then, we have $\bar{z}^1(t) = \frac{\bar{z}^1}{[1 - G_1(T^1_{1e})]K_1^1}$ for $t < T^1_{1e}$ and $\bar{z}^1(t) = 1$ for $t \geq T^1_{1e}$ and so we can explicitly calculate...
For equilibrium in case of $T$ while the right hand side is continuously increasing, this is uniquely defined. In fact, given that by agent 2 in order that 2’s conditional probability of irrationality hits 1 at time $t$, we define $T^*_{1e}$ to concede at $T^*_{1e}$. The case of $\delta$ is certainly irrational. We therefore need to adjust agent 2’s initial mass concession, in order that his conditional probability of irrationality hits 1 at time $T^*_{1e}$. We describe both of the above cases as type 2 equilibria even though they are slightly different from each other.

Type 3 equilibria

In the case of $T^*_{2e} < T^*_{1e}$, and $T^*_{2e} < T^*_{1e}$, equilibrium requires finding a level of concession by agent 1 at time zero so that at $T^*_{1} > T^*_{1e}$, either both agents hit probability 1 of irrationality together at time $T^*_1$, or at $T^*_1$ agent 1 is indifferent between conceding immediately at $T^*_1$, or waiting until time $T$. The reason that $T^*_1 > T^*_2$ is that mass concession by agent 1 increases $V^*_2(\tilde{z}^*_1(t))$ (at least for $\tilde{z}^*_1(t) \approx 1$) and so slows down agent 2’s rate of concession.

Lemma 7. For any initial probabilities of irrationality, and initial demands, an equilibrium of one of the types described above exists and is unique.

Proof. See appendix.

Given that for any initial parameters the equilibrium in this subgame is unique, a unique distribution of equilibrium outcomes in the multi-type case, can again be shown by the arguments of Abreu and Gul.

To analyze the limiting equilibrium, in a given subgame as the probability of irrationality for agents 1, 2 and 3 converge to zero at the same rate, we first need to define a characterizing level of demand on issue 1, $\hat{\alpha}^1$. In pursuit of this, we first define:

$$\varphi(\alpha^1, \alpha^2) = \frac{1 - \alpha^1 + D^1}{1 - \max C^3 + D^1} + \frac{1 - \alpha^1 + D^1}{1 - \alpha^2 + D^1}$$
Notice that $\varphi(\alpha^1, \alpha^2) = \frac{k_1^1}{k_2^1} + \frac{k_1^2}{k_2^1}$ when agent 1 demands $\alpha^1$, agent 2 demands $\alpha^2$ and agent 3 demands $\max C^3$. $\hat{\alpha}^1 \in (0, 1)$ is defined as the solution to $\varphi(\alpha^1, 1 - \alpha^1) = 1$, with the exceptions that if $\varphi(\max C_1^1, 1 - \max C^1_1) \leq 1$ then $\hat{\alpha}^1 = 1$ and if $\varphi(1 - \max C^2_1, \max C^2_1) \geq 1$, then $\hat{\alpha}^1 = 0.$

Given this we can show:

**Lemma 8.** For fixed demand by agent 1, $\alpha^1 < \hat{\alpha}^1$ then for a sufficiently rich type space for agent 2 and 3, agent 2 will concede to agent 1’s demand with probability approaching 1 in the limiting equilibrium, with agent 1 making the concession if the inequality is reversed.

**Proof.** See appendix. $\square$

The proof revolves around showing that for $\alpha^1 < \hat{\alpha}^1$, agent 1’s reputation must endogenously grow by the end of bargaining on issue 1, to be sufficiently large that if agent 3 made the counterdemand $\alpha^3 = \max C^3$, it would be agent 3, rather than agent 1 who would concede in the continuation game. This is precisely determined by whether or not $\frac{k_1^3}{k_2^3} + \frac{k_1^2}{k_2^3}$ is greater or less than 1. Given this agent 1 has a strict incentive not to concede to 2 sufficiently close to the deadline $T$ for bargaining on issue 1, and so agent 2 must concede with sufficient initial probability to exhaust all rational agents by that point.

Given immediate concession of agent 2 to agent 1, agent 1’s conditional probability of irrationality does not change (with probability almost 1), and so if we have $\alpha^1 \leq \alpha^{1*}_2$, then agent 3 must concede to this demand in bargaining over issue 2 with probability approaching 1, by our standard limiting arguments. On the other hand if $\alpha^1 = \alpha^{1*}_2 + \epsilon$, then for any counterdemand by agent 3 such that $\alpha^3 < \alpha^{3*}_2 + \epsilon$, agent 1 must concede with probability approaching 1 in the limit.

Given all this, for a rich type space we have that in equilibrium agent 1 by demanding arbitrarily close to $\alpha^{1*}_2$ can earn $(1 + \delta^1)B\alpha^{1*}_2$ in the limit. For $\alpha^1 = \alpha^{1*}_2 + \epsilon \leq \hat{\alpha}^1$ he earns profits of approximately $\alpha^{1*}_2 + \epsilon + \delta^1 B\max(\alpha^{1*}_2, \epsilon, 0)$. Finally demanding strictly more than $\hat{\alpha}^1$ he earns strictly less than $\hat{\alpha}^1$. In full multi-type model then, we summarize the outcome as follows:

**Proposition 4.** If $\delta^1 B < \max\{1, \frac{\alpha^1}{\alpha^{1*}_2} - 1\}$, then in the rich type limiting equilibrium agent 1 makes an initial demand (arbitrarily close to) $\hat{\alpha}^1$, resulting in payoffs $(\hat{\alpha}^1, 1 - \hat{\alpha}^1, B)$, while if $\delta^1 B \geq \max\{1, \frac{\alpha^1}{\alpha^{1*}_2} - 1\}$ agent 1 makes a demand (arbitrarily close to) $\alpha^{1*}_2$ for payoffs ($(1 + \delta^1 B)\alpha^{1*}_2, 1 - \alpha^{1*}_2, B(1 - \alpha^{1*}_2))$.

This result immediately follows from the preceding analysis. If there is more surplus available today than tomorrow, agent 1 will make an aggressive demand today, seeking mass concession by the agent 2 thanks to her strong position, and accept the need to concede to the most aggressive type of agent 2 tomorrow however if there is more surplus available tomorrow, then agent 2 will moderate her demand today in order to not find herself in a weak bargaining position tomorrow.

\footnote{To see that $\hat{\alpha}^1$ is thus well defined notice that except in the two boundary cases highlighted it is simply solved as a quadratic equation, and to see there is a unique solution in $(0,1)$ notice that $\varphi(\alpha^1, 1 - \alpha^1)$ is continuously (strictly) decreasing in $\alpha^1$ but $\varphi(1 - \max C^2_1, \max C^2_1) < 1$ and $\varphi(\max C^1_1, 1 - \max C^1_1) > 1.$}
The fact that agent 1 may in equilibrium initially make a demand to which agent 2 concedes (immediately) but to which agent 3 responds with his most aggressive demand, to which 1 then concedes (immediately) is at first sight somewhat surprising. This occurs because late in the war of attrition, agent 1 will have endogenously built up a reputation that is worth defending, giving him strong position; the incentives are created by the fact that concession will lead to facing the most aggressive type of agent 3 in the future, instead of the second most aggressive type. Given this, as has been discussed at length elsewhere, strength even at the very end of bargaining translates in to strength at the start of bargaining.

Despite the outcome being surprising then, the forces at work are extremely plausible. That having a reputation to defend for being a committed bargainer, gives an agent an advantage is also plausible, if not surprising. Moreover, the more general conclusion is a also very intuitive. That is, that a long run agent will moderate his demands when there is a large amount of surplus available in bargaining at a later date, in order to have a strong reputation at that date, and make more aggressive demands when the surplus available tomorrow is small, and so there is less worth to a strong reputation.

To see what all this means in a simple example, let $D^i_k = 0$ so that all issues are then pure surplus division. Also let $\max C^i = 1 - \varepsilon \approx 1$ for all agents. Then $\varphi(\max C^1, 1 - \max C^1) = \frac{\varepsilon}{\delta^1 + \frac{1}{1 - \varepsilon}} > 1$ and so by the above analysis, any demand demand made by agent 1, will be accepted with probability approaching 1 in the limit. However, given $\alpha^i_2 = 1/2$, we have that if $\delta^1 B > 1$ agent 1 will demand arbitrarily close to $a^1 = 1/2$ in the limit and have this demand accepted in both periods with probability close to 1. While if $\delta^1 B > 1$, then agent 1 will demand almost the entire surplus on bargaining issue 1, and be accepted with probability approaching 1, before agent 3 counterdemands almost the entire surplus on issue 2 and has this proposal accepted with probability approaching 1. Clearly, agent 2 and 3 have a great interest in the size of $\delta^1 B$.

The model can clearly be extended with relative ease to more than 2 periods of short run players. This will be equivalent to decreasing $\delta^1 B$ in the above model. The same logic as above will ensure that agent 1 has lower incentives to concede than short run agents, in all periods prior to the final one, making them comparable to the first period above. However, by increasing the number of periods prior to the final one, we make the payoff to this period larger relative to the last bargaining issue.

For instance, with three short run bargainers (2,3,4) let $D^i_k = 0$ and $\max C^i = 1 - \varepsilon \approx 1$. Then given that we must certainly have $\varphi(\max C^1, 1 - \max C^1) = \frac{\varepsilon}{\delta^1 + \frac{1}{1 - \varepsilon}} > 1$, in the subgame involving bargaining with agent 3, agent 3 must concede to any demand of agent 1 with probability approaching 1 in the limit; certainly then, agent 2 must do likewise. But now agent 1 must tradeoff getting $1 - \varepsilon$ of the surplus when bargaining against the first two agents and $\varepsilon$ against the last, or getting $1/2$ of the surplus against all three.

8 Conclusion

We have analyzed a repeated bargaining model in which agents have reputations for being committed, which carry over from one issue to another. This led to many interesting conclusions, such as the importance of the agenda, and the possibility of disagreement for arbitrarily small
probabilities of commitment, when combined with uncertainty about future bargaining conditions. The conclusions of such a model can help shed light on legislative bargaining but also many other situations beyond the political economy context.

Regarding extensions and future work: the models of section and the related model in section , could be extended to allow questions to be asked about replacement of agents in a repeated bargaining context. Intuitively, if an agent may be replaced (with exogenous probability) in a repeated bargaining setting, this will lower that agent’s effective cost of conceding, while increasing that of his opponent (the cost of facing an aggressive initial demand from a potentially aggressive replacement). Both of these should be expected to hurt the agent’s bargaining strength. However, if replacement is endogenous, coming only after backing down to an opponent, this may increase the agent’s bargaining strength, by increasing the effective cost of concession.

Modeling the implications of agents being able to somewhat recover a reputation for commitment, for instance through unobserved replacement, would also be a worthwhile avenue of inquiry, although it is not obvious that the technique used in this paper of characterizing outcomes as the probability of rationality becomes (extremely) small is the best way to analyze that.

Proof of Lemma 1

We prove the existence of such an equilibrium, when type 1, 2 and 3 equilibria do not prove that an equilibrium exists in all cases. In fact we can view a type 4 equilibrium as generalizing a type 2 equilibrium when \(T^*_1 = 0\), however in this case one agent may have a strict incentive to wait until time \(T\), rather than conceding at \(0^*\).

Excluding one non-generic case, we can see the equilibrium must have some interval of non-concession \((T^*_1, T)\). Given that if there is no type 3 equilibrium then using the continuous concession rates of equation 8 we cannot continuously exhaust the store of rational agents by time \(T\), without having mass concession by both agents at time 0 (which is ruled out). Now suppose there is continuous concession up to time \(T\) let:

\[
K^j_{2a} = \frac{(1 - \alpha^j) + D^j_{2a}/B_{\omega}}{\alpha^j + \alpha^j - 1}
\]

Now notice that in the non-generic case with:

\[
\frac{K^j_{2a}}{K^j_{2b}} = \frac{K^j_{2a}}{K^j_{2b}}
\]

We have \(c^j_a = c^j_b = \min[1, \frac{1}{(1-G(T^*_1))\bar{z}^j_2}] = \bar{z}^j_2(\bar{z}^j_2)^{-K^j_{2a}/K^j_{2b}}\) (if this is so equilibrium can be solved in a similar way to the model of section ), but in any other case, it is impossible to have \(c^j_a = c^j_b = 1\) for both agents. In which case, at least one agent, say \(i\), must concede with positive mass in the subgame at time \(T\) should there be an ongoing war of attrition. This means that at least one agent must obtain a profit bump \(\delta^j(V^j_2(\bar{z}^j_2) - (1 - \alpha^j)B) > 0\) by waiting until at least time \(T\), instead of conceding at time \(T^*\).

For us to have an equilibrium then, there must be some non-empty gap: \((T^*_1, T)\) in which agents do not concede. Given this gap we must have that in equilibrium one agent must make positive mass concession at time \(T\) once it is revealed that issue \(a\) is to be debated, while the other must make mass concession on issue \(b\), for there to be incentives for both agents to have incentives to wait rather than concede at \(T^*_1\).

Indeed, conditional on not conceding up until time \(T^*_1 < T\), the expected cost of waiting until time \(T\) for agent \(i\) is:

\[
(1 - G(T^*_1))(D^j_1 + (1 - \alpha^j)) > 0,
\]

and so there must be an equal positive benefit of waiting \(\delta^j[V^j_2(\bar{z}^j_2) - B(1 - \alpha^j)]\), which requires \(V^j_2(\bar{z}^j_2) > B(1 - \alpha^j)\). Equality of costs and benefits is obtained when:

\[
(1 - G(T^*_1))(D^j_1 + (1 - \alpha^j)) = \delta^j[V^j_2(\bar{z}^j_2) - B(1 - \alpha^j)] = \delta^j(1 - c^j_{2a})p_\omega B_\omega(\alpha^j + \alpha^j - 1)
\]

(13)
Where the final equality assumes agent $i$ is conceded to in equilibrium with positive mass on issue $\omega$.

Now ignoring incentive compatibility in stage 1, we have that $c'_{j,\omega}$ is a deterministic function of $\hat{z}_2$, which is itself completely determined by parameters $c', c''$ and $T'$. We have: $c'_{j,\omega} = \min\{1, \frac{\delta' V_j^2(\zeta) - D_j' - \gamma}{\delta' B(1 - \delta')}\}$, where $K'_{j,\omega}$ is given in equation 5. So equilibrium involves using the free variables $T'_i$ and $c', c''$ (where $(1 - c')(1 - c'') = 0$) to solve equation 6, for both agents.

If there is no type 1 equilibrium we must have: $\delta' V_j^2(\zeta) - D_j' < (1 + \delta' B)(1 - \alpha')$ for some agent, let us suppose it is agent $i$. Let $\hat{z}(t) = \frac{\zeta(t)}{1 - \gamma}$, now certainly for any $T'_i$ such that $(1 - G(T'_i))(D_i' + (1 - \alpha')) < \delta' B(1 - \hat{z}(t))(\alpha' + \alpha' - 1)$, you can find a $c', c''$ that solves the equation 13 for agent $i$.

To see this, let $c'' = (1 - c'')(1 - c''') = c' - c'$, for fixed $(1 - G(T'_i))(D_i' + (1 - \alpha'))$ you have $V_j^2(\hat{z}_2)$ is continuously and monotonically increasing in $c'$, and eventually converging to $B[\alpha' - \hat{z}(t)]$, while if $c'$ falls sufficiently it converges to $\delta' B(1 - \alpha')$. This that there is a unique $c'$ where equilibrium obtains. Moreover, that there is no type 3 equilibrium we have $\hat{z}(T) < 1$, the condition, ensuring that there is certainly an interval $(T', T)$ on which we have $c'(T'_i)$ well defined and continuous in $T'_i$.

We then need to find a $T'_i$ to solve the following equation for agent $j$:

$$ (1 - G(T'_i))(D_i' + (1 - \alpha')) = \delta' [V_j^2(\hat{z}_2(T'_i, c'(T'_i))) - B(1 - \alpha')] $$

Clearly both sides of this equation are continuous in $T'_i$. Furthermore, as $T'_i \to T$ we have that $c'(T'_i)$ must ensure that $V_j^2(\hat{z}_2(T'_i, c'(T'_i))) \to 0$ and so $c'(T'_i)$ must fall, and given the non-generic assumption above we do not have $c'_{2a} = c_2$, and so agent $i$ begins to lose one of the wars of attrition (on say issue a) before the other, and so we must eventually have $\hat{V}_j^2(\hat{z}(T'_i), c'(T'_i)) > B(1 - \alpha')$ for $T'_i$ arbitrarily close to $T$. In this turn means $(1 - G(T'_i))(D_i' + (1 - \alpha')) < \delta' [V_j^2(\hat{z}_2(T'_i, c'(T'_i))) - B(1 - \alpha')]$.

By contrast suppose we reduce $T'_i$ toward $T'$, a point at which we have $(1 - G(T'))(D_i' + (1 - \alpha')) = \delta' B(1 - \hat{z}(T'))(\alpha' + \alpha' - 1)$, at which $c'(T')$ cannot be defined, then as $T'_i \to T'$ we must have that $c'(T'_i)$ is rising so that $V_j^2(\hat{z}_2(T'_i, c'(T'_i)) \to B[\alpha' - \hat{z}(T') + \hat{z}(T')(1 - \alpha')]$, but clearly in this case we have $V_j^2(\hat{z}_2(T'_i, c'(T'_i)) \to B(1 - \alpha')$ and so: $(1 - G(T'_i))(D_i' + (1 - \alpha')) > \delta'[V_j^2(\hat{z}_2(T'_i, c'(T'_i))) - B(1 - \alpha')] \to 0$. But in which case, again by the intermediate value theorem we have equality in equation 13 for agent $j$ at some point.

If alternatively there is no $T'$ with the above properties, then we continuously lower $T'_i$ from $T$ to 0. At $T'_i = 0$ we must still have $\delta' V_j^2(\hat{z}_2(0, c'(0))) - D_i' = (1 - \alpha')) + \delta' B(1 - \alpha')$. If at this point there has still been no $T'_i$ such that $(1 - G(T'_i))(D_i' + (1 - \alpha')) = \delta'[V_j^2(\hat{z}_2(T'_i, c'(T'_i))) - B(1 - \alpha')]$, then we have: $(1 - \delta') B(1 - \alpha') < \delta'[V_j^2(\hat{z}_2(0, c'(0))) - D_i']$.

In which case, conditional on initial mass acceptance $c'(0)$, we have a situation of initial mass concession such that agent $j$ strictly prefers to wait on the interval $(0, T)$, while is indifferent to doing so or immediately conceding. Given that we are excluding the case where there is a type 1 equilibrium and thus initially had $\delta' V_j^2(\hat{z}_2(0, c'(0))) - D_i' < (1 + \delta') B(1 - \alpha')$, then in order for $c'(0)$ to make $\delta'[V_j^2(\hat{z}_2(0, c'(0))) - D_i'] = (1 - \alpha')) + \delta' B(1 - \alpha')$ we must have $c'(0) > 0$, and so it is agent $i$ that conceders at time 0, and the initial mass concession is incentive compatible. But in this case, we must have found an equilibrium of type 2.

We now prove that the equilibrium found above is unique.

To prove this first suppose there is a type 1 equilibrium, and so $\delta' V_j^2(\hat{z}_2) - D_i' \geq (1 + \delta') B(1 - \alpha')$ for both agents. Can there be an equilibrium of type 3? In a type 3 equilibrium all rational agents concede by time $T'_i \leq T'$. But in which case, at least one of the agents, say i, must obtain a payoff of $(1 + \delta') B(1 - \alpha')$ in this equilibrium. Consider then the case when agent $i$ chooses a concession time of $T'_i$. In this scenario, he must obtain an expected payoff of $(1 + \delta') B(1 - \alpha')$, but all rational agent $j$s will have conceded to him at times strictly before they would have in the type 1 equilibrium, from which his payoff was $\delta' V_j^2(\hat{z}_2) - D_i'$, but in this case we must have a contradiction: $(1 + \delta') B(1 - \alpha') > \delta' V_j^2(\hat{z}_2) - D_i' \geq (1 + \delta') B(1 - \alpha')$.

Suppose there is a type 1 equilibrium, can there also be a type 4 equilibrium (of which we treat type 2 as a special case)? First notice that if we have $\delta' V_j^2(\hat{z}_2) - D_i' \geq (1 + \delta') B(1 - \alpha')$, it must be that at prior probabilities, $\hat{z}$, agent $i$ must win the war of attrition on one of the issues debated at time $T$, say issue $a_i$ (while agent $j$ wins on the
other issue) otherwise we would have $V'(z) = B(1 - \omega^t)$. But in this case, we have $c_d^i = \frac{\delta^t(z)}{c_d^i(z)} - K_d^i/K_d^i < 1$ and $c_d^j = \frac{\delta^t(z)}{c_d^j(z)} - K_d^j/K_d^j < 1$, which in turn implies $\frac{K_d^b}{K_d^b} < \frac{K_d^i}{K_d^i}$. This implies that in any equilibrium it must be on issue $a$ that agent $i$ wins with positive probability.

Secondly, notice that in the considered case one agent (agent $i$) must obtain the equilibrium payoff of $(1 + \delta B)(1 - \alpha^t)$ and concede with positive probability in the interval $[0, T)$. Then focusing solely on issue $a$, reaches time $T$ with a strictly higher conditional probability of irrationality. But in which case we must have a strictly lower exhaustion time on issue $a$, $T_{2a}$, than in the type 1 equilibrium $T_{2a}^i$.

Let $T^\omega(\omega)$ be time $T^\omega$ after the revelation of uncertainty $\omega$ and let $\hat{z}(t)$ be i’s posterior probability of irrationality at time $t$. Then given that $T_{2a} < T_{2a}^i$, we must have we must have $\hat{z}(T_{2a}^i) < \hat{z}(T_{2a})$ in order to reach probability 1 by time $T_{2a}$ and $T_{2a}^i$ respectively. But given this, we must have that $F^i_d(t) > F^i_d(t)$ for all $t \geq T$ and $F^i(t) \geq F^i(t)$ for all $t < T$. But this says that agents always concede faster in the type 4 equilibrium than in the type 2 equilibrium, which means that agent $i$ must be strictly better off in the type 4 equilibrium. But this contradicts $\delta^t V'(z) - D_i^a \geq (1 + \delta B)(1 - \alpha^t)$. (For instance consider the expected payoff of an agent in both equilibria who at exactly $T_{2a}$, after favorable revelation of uncertainty.)

Now suppose there is a type 3 equilibrium. Clearly there cannot be more than one. Is it possible that there is also a type 4 equilibrium? In this type 3 equilibrium at least one of the agents (say agent $i$) obtains an equilibrium payoff $(1 + \delta B)(1 - \alpha^t)$. Let $T_1^i$ be the first stopping time in the type 4 equilibrium and the exhaustion time in the type 3 equilibrium be $T_1^i$. Clearly we must have $T_1^i < T_1^i$ or else all given the same rates of concession on $[0, T_1^i]$ agent $j$’s rational types will be exhausted by $T_1^i$. 

Now consider the subgame immediately after initial concession in each equilibrium (time $0^*$), the conditional continuation payoffs must be $(1 + \delta B)(1 - \alpha^t)$ for both players in the type 3 equilibrium, and at least one of the players in the type 4 equilibrium. The posterior probabilities of irrationality at $0^*$ must be at least weakly greater than the initial priors. Consider the distribution functions of concession times in this subgame $F^i(t)$, $F^j(t)$ so that $F^i(0^*) = F^j(0^*) = 0$ (i.e., $F^i(t) = \frac{F^i(t)}{F^i(t)}$). Given that in the subgames of both equilibria agents must concede at a continuous rate on $[0, T_1^i)$, and agents concede at the same rate, we must have $F^i_d(t) = F^i_d(t)$, for all $t \leq T_1^i$ and $F^i_d(t) < F^i_d(t)$ for all $t > T_1^i$. But in which case, if agent $i$ concedes in equilibrium 3 at a time $T_1^i$ (to obtain an expected payoff of $(1 + \delta B)(1 - \alpha^t)$) must have a strictly higher expected payoff than agent i’s expected continuation payoff in the type 4 at time $0^*$ when he intends to wait until at least time $T$. But this yields a contradiction as the latter must be at least $(1 + \delta B)(1 - \alpha^t)$.

Suppose there are two type 4 equilibria. Clearly these must have different $T_1^i$. Let the first of these be given by $T_1^i$ the second by $T_1^i$ (assume $0 \leq T_1^i < T_1^i < T$). At least one agent, say $i$, obtains a payoff of $(1 + \delta B)(1 - \alpha^t)$ in the equilibrium with $T_1^i$, and wins on issue $a$, while $j$’s continuation payoff at $0^*$ must be similar. At least one agent must likewise obtain this payoffs in the equilibrium with $T_1^i$, suppose that is still agent $i$.

Now notice that as $T_1^i < T_1^i$ agent $i$’s immediate concession probability does not change and so he must arrive at time $T$ with a lower probability of irrationality in the equilibrium with $T_1^i$. But in that case we must have $T_1^i > T_i$. But in which case we must have $\hat{z}(T_1^i) < \hat{z}(T_1^i)$; we need a lower probability of $i$’s irrationality, given concession at the same rate thereafter. But in order to make agent $j$ at least as well off conceding at time $T_1^i$ compared to at time $T_1^i$ agent $i$’s conditional probability of concession at time $T_1^i$ must increase in the $T_1^i$ equilibrium. And therefore, we must have $\hat{z}(T_1^i) < \hat{z}(T_1^i)$ as well. But if this is so we must also have $T_1^i > T_1^i$, requiring $\hat{z}(T_1^i) < \hat{z}(T_1^i)$, which in turn means that $F^i_d(t) < F^i_d(t)$ for all $t > T_1^i$. This leads to a contradiction (consider the expected payoff to the agent $i$ in both equilibria when he concedes at time $T_1^i$).

Suppose then that agent $j$ obtains a payoff of $(1 + \delta B)(1 - \alpha^t)$ in the equilibrium with $T_1^i$ (allowing i’s payoff to be larger in this equilibrium if j concedes with positive probability at time 0). In which case as i’s initial concession probability decreases in the $T_1^i$ equilibrium (compared to the $T_1^i$ equilibrium), and we have that i must reach $T$ with a lower probability of irrationality, requiring that $T_1^i > T_1^i$. By a similar argument to that above implies $\hat{z}(T_1^i) < \hat{z}(T_1^i)$, $\hat{z}(T_1^i) < \hat{z}(T)$ and $T_1^i > T_1^i$. But this means that conditional on not conceding at time 0, i faces a lower distribution of concession times in the equilibrium with $T_1^i$ ($F^i_d(t) < F^i_d(t)$ for $t > T_1^i$). But in which case, given that the continuation payoffs for agent j at time 0 are the same in both equilibria, we have a contradiction (consider the payoff to agent $j$ when he concedes at time $T_1^i$). And we are done.

Proof of Lemma 2
For fixed $\bar{\epsilon} > 0$, for agent $i$ to be indifferent to conceding in the limit between conceding at time $T_{1i}' + \hat{\epsilon}$ or waiting for the revelation of uncertainty at time $T$, his continuation payoff from waiting until time $T$ must be smaller than $V_i' - \hat{\epsilon}z_i$, for some fixed $\hat{\epsilon}z_i > 0$. For this to be so, it must be that ratio of posterior probabilities of irrationality at $T$, $\frac{e}{z_i}$ converges to zero in the limit. If this is to occur, we clearly need that agent $j$ concedes to $i$ at time 0 with probability approaching 1.

For fixed $\bar{\epsilon} > 0$, for agent $j$ to be indifferent to conceding in the limit between conceding at time $T_{1i}' - \hat{\epsilon}$, or waiting for the revelation of uncertainty at time $T$, his continuation payoff from waiting until time $T$ must be larger than $V_j' - \bar{\epsilon}z_j$, for some fixed $\bar{\epsilon}z_j > 0$. Again, for this to be so, it must be that ratio of posterior probabilities of irrationality at $T$, $\frac{\bar{\epsilon}z_i}{\bar{\epsilon}z_j}$ converges to zero in the limit. If this is to occur, we clearly need that agent $j$ concedes to $i$ at time 0 with probability approaching 1.

But now as $\bar{\epsilon}$ was arbitrary, we know that in the limit if, $T_{1i}' > T_{1j}'$ then certainly agent $i$ wins the war of attrition at time 0 with probability approaching 1.

**Proof of Lemma 3**

To prove the existence of an equilibrium in this model, we show, like the previous proof that if no type 1, 2, or 3 equilibrium exist, then a type 4 equilibrium is certain to exist. Note that in this case we must have $T_{1i}' \in (0, T_{1i}')$. The case of $T_{1i}' = T_{1i}'$ is easily dealt with above.

Suppose then $T_{1i}' > T_{1i}'$. Let $z_i = \frac{\epsilon}{(1 - c_i'(R))c_i'}$ be agent $i$'s conditional probability of irrationality, given no concession at time 0, and continuous concession on $(0, T_{1i}']$. Suppose we have:

$$
(1 - G(T_{1i}'))(D_1^i + (1 - \alpha')) \geq \delta^i B[(1 - z_i^2)\alpha' + z_i^2(1 - \alpha') - (1 - \phi(\alpha'))]
$$

$$\geq \delta^i B[\alpha' + \phi(\alpha') - 1 - z_i^2(\alpha' + \alpha'- 1)]
$$

Then agent $j$ would prefer to concede at time $T_{1i}'$ rather than wait until $T$ even if all remaining rational $i$ players conceded to him at that point. In which case we increase agent $j$’s initial concession (at time 0) so that his posterior probability of irrationality reaches 1 at time $T_{1i}'$. Agent $i$ on reaching time $T_{1i}'$ knows she is facing an irrational player but waits until time $T$, for fear of an irrational agent $j$ making a more aggressive demand in bargaining game 2.

If that inequality does not hold then it is possible (through sufficient concession at time $T$) to make agent $j$ indifferent between conceding at time $T_{1i}'$ or waiting until time $T$ to receive mass concession. This mass concession $(1 - c_i^2)$ must then be defined by the following equation:

$$
(1 - G(T_{1i}'))(D_1^i + (1 - \alpha')) = \delta^i B[(\alpha' + \phi(\alpha') - 1) - c_i^2(\alpha' + \alpha'- 1)]
$$

$$c_i^2 = \frac{\alpha' + \phi(\alpha') - 1}{\alpha' + \alpha'- 1} - \frac{\delta^i(\phi(\alpha') - \alpha') (D_1^i + (1 - \alpha'))}{\delta^i(\alpha' + \alpha'- 1) (D_1^i + (1 - \alpha'))} < 1 \quad (14)
$$

For such mass concession to be the equilibrium outcome of the subgame starting at $T$, we must adjust initial concession rates so that posterior probabilities of irrationality at $T^*$, $z_i^2$, give such concession as the equilibrium outcome. For this to be so then we need:

$$
c_i^2 = \frac{\bar{z}_i^2}{(1 - G_2(T^*_2))\bar{z}_i^2} = \frac{\bar{z}_i^2}{c_i^2(R)K_i^2} = \frac{\bar{z}_i^2}{c_i^2(R)K_i^2} (\frac{\bar{z}_i^2}{c_i^2(R)K_i^2})^{-\frac{1}{K_i^2}} \quad (15)
$$

Where:
Given then that we have a bound in agent 1’s o so agent 2 can demand at most \( \alpha - B \phi(\alpha') - \alpha \).

This can be uniquely solved to yield the \((1 - c'), (1 - c)\) which solve equation 14.

This completes the description of the equilibria. No other possible equilibrium structures fit the requirements of our “standard arguments” outlined above and our technical assumption that \((1 - \alpha') + D_1^i \neq \delta B(\phi(\alpha') - \alpha')\). We now prove that the equilibrium is unique. Clearly, if there is a type 1 equilibrium, there is no equilibrium of any other type. To see this, notice that in the type 1 equilibrium we have \((1 - \alpha') + D_1^i < \delta B(\phi(\alpha') - \alpha')\) for one player, say i, and \((1 - \alpha') + D_1^i + \delta B(1 - \phi(\alpha')) < \delta V_2(z)\) for player j, then we clearly cannot have concession on \([0,T)\) for player i regardless of the level of concession by j, and associated increased probability of irrationality. However, given this, agent j clearly, can have no incentive to concede before time T either. But all other equilibrium types involve concession before time T.

Suppose then that there is a type 2 equilibrium. Again then we have \((1 - \alpha') + D_1^i < \delta B(\phi(\alpha') - \alpha')\), agent i, who hence cannot be induced to concede before time T even if certain to be facing an irrational opponent. Given this, clearly agent j would never concede with positive probability on \((0,T)\). Given that we also have that \((1 - \alpha') + D_1^i + \delta B(1 - \phi(\alpha')) \leq \delta V_2(z)\), the only variable of interest is the level of concession at time 0 for agent i.

But given that \(V_2(z)\) is continuously increasing in the concession probability at time 0, the concession probability at time 0 is uniquely determined by equation 11.

If there is a type 3 equilibrium then we must have \((1 - \alpha') + D_1^i < \delta B(\phi(\alpha') - \alpha')\) for both agents. Furthermore, we have \(T_{1e}^i \leq T_{1m}^i\) so continuous concession up to time \(T_{1e}^i \equiv T_{1e}^j\) exhausts all rational types of agent i. But time \(T_{1e}^j\) is the first time that it could ever be possible to make both agents indifferent between conceding at \(T_{1m}^i\) and T, remember there can be mass concession at time T for at most 1 agent in equilibrium, and so it is infeasible for there to be an equilibrium where it was not true that both agents’ rational types were not exhausted by time T (i.e. no type 4 equilibrium). But in which case, given that agents must concede at given continuous rates, on \((0, T_{1m}^i)\) the only thing to determined is the concession probability at time 0 by agent j, to ensure a conditional probability of irrationality of 1 at time \(T_{1m}^i\), and clearly this is uniquely determined.

Finally when there is a type 4 equilibrium, we have \(T_{1m}^i = T_{1m}^j < T_{1e}^i\). Given that at most one agent receives mass concession at time T, we cannot have that agents concede continuously only up to \(T_{1m}^i\) as then at least one agent would have an incentive to offer mass concession at time t, which we know is impossible in equilibrium. Time \(T_{1m}^i = T_{1m}^j\) allows agent i to be indifferent between conceding at \(T_{1m}^i\) or at T, when receiving no mass concession at time T. For any \(T > T_{1m}^i\) conditional on not having conceded already, we must then have that i strictly prefers to wait until T, even when certainly facing an irrational agent. But equally as in the above analysis only h concession t > \(T_{1m}^i\). Given continuous concession on \((0, T_{1m}^i)\) then we have \(c_2^i, c_2^j\) and \(c', c'\) uniquely determined thanks to the analysis above.

**Proof of lemma 4**

Suppose there is inefficiency in the limit, in which case by the above analysis of the subgame following \(\alpha' + \alpha' > 1\), it must come from a type 1 equilibrium, in which neither agent concedes on \([0,T)\). Now suppose agent 1 mimics a type \(\alpha'\) with positive probability in the limit, and agent 2 counterdemands \(\alpha^2 > 1 - \alpha'\) with positive probability, such that the result is a type 1 equilibrium, then payoffs are \(\delta V_2(z) - D_1^i \geq (1 - \alpha') + \delta B(1 - \phi(\alpha'))\).

If this is the case, then it is certainly true that, agent 2 mimics a type such that \(K_2^1 > K_2^1\) or else agent 2’s payoff converges to \(\delta B(1 - \alpha') - D_2^1\), which is certainly lower than the payoff to immediate acceptance, \((1 + \delta B)(1 - \alpha')\). But in which case agent 2’s payoff to imitating all types following the demand \(\alpha'\) converges to \(\delta B\alpha^2 - D_2^1\), while agent 1’s payoff following this particular counterdemand converge to \(\delta B(1 - \alpha^2) - D_1^1\).

Moreover, it cannot be that \(\alpha'\) is arbitrarily close (within \(\epsilon\) of) to \(\alpha_2^1 = \max\{\min\{\frac{B_1 - D_1^i - D_1^j}{2B}, 0\}, 1\}\), as if this were so agent 1 can demand at most \(\alpha_2^2 + \epsilon\), while ensuring \(K_2^2 > K_2^2\). But then for sufficiently small \(\epsilon\) we have \(\delta B(\alpha_2^2 + \epsilon) - D_2^1 < (1 + \delta B)(\alpha_2^2 + \epsilon))\) and 2 would have reason to accept 1’s offer immediately.

Given then that we have a bound in 1’s offer \(\alpha' = \alpha_2^1 + \bar{\epsilon}\), (for some fixed \(\bar{\epsilon}\)) it must be that for a sufficiently rich type space, agent 2, must mimic a type arbitrarily close (from below) to \(\alpha_2^2 + \bar{\epsilon}\). To see this notice that given a fixed \(\alpha'\), increasing \(\alpha^2\), while still ensuring \(K_2^2 > K_2^2\), can only increase agent 2’s limit payoff from waiting until
time \( T \), which is \( \delta^2 \alpha^2 - D_1^+ \).

Given such a continuation payoff at \( T \), agent 2 can only be induced to concede with any probability before time \( T \) in the limiting equilibrium if the conditional probability of 1’s irrationality at time \( T, \frac{\alpha}{\alpha^2} \) in the continuation game approaches zero at a rate strictly faster than \( \frac{\alpha}{\alpha^2} \) (so that \( V_2(\frac{\alpha}{\alpha^2}) \) is bounded away from \( \delta^2 \alpha^2 \)). But for this to happen, 1 must concede to 2 at time zero with probability approaching 1 (in which case 2 certainly obtains a payoff greater than \( \delta^2 \alpha^2 \)). And so it must be that in the limiting type 1 equilibrium agent 2 is mimicking the maximal type such that \( K_2^1 > K_2^1 \), that is arbitrarily close to \( \alpha_2^1 + \bar{e} \).

Is it possible that agent 2 mimics types other than this maximal type, which we call \( \alpha^2 \approx \alpha_2^1 + \bar{e} \), with positive probability in the limit? We know that the payoff in the type 1 equilibrium payoff is increasing towards the limit. We also know that among conceivable type 1 equilibria, imitating \( \alpha^2 \) most profitable. In all non-generic limiting equilibrium apart from type 1 equilibria, at least one agent concedes with at time zero (with probability approaching 1), clearly this cannot be agent 2 or she would have an incentive to accept 1’s initial offer instead, and so the payoff from these other equilibria, in which agent 1 demands \( \alpha^1 \) in the limit must be: \( \alpha^2 + \delta^2 \beta_2 \alpha^2 \), Further ignoring the non-generic case of \( \alpha^2 + \delta^2 \beta_2 \alpha^2 \approx \delta^2 \alpha^2 \), in the limit we can be sure that only \( \alpha^2 \) is imitated. This ensures that agent 1’s payoff from demanding \( \alpha^1 \) is exactly \( \delta^2 B(1 - \alpha^2) - D_1^+ \approx \delta^2 \alpha^2 \) in the limit.

But now agent 1 can make a demand arbitrarily close (from below) to \( \alpha_2^1 + (1 - \alpha^2) \approx \alpha_2^1 - \bar{e} \), which ensures that \( K_2^1 > K_2^1 \) if there is a counterdemand. But then agent 1 has a strictly larger continuation payoff in the limit from waiting until time \( T \), than in the type 1 equilibrium above \( \delta^2 \alpha_2^1 B - D_1^+ > \delta^2 B_2 \alpha^2 - \bar{e} - D_1^+ \). Agent 1 obtains this continuation payoff unless we have, 2’s conditional probability of irrationality \( \alpha_2^1 \), approaching zero strictly faster than \( \bar{e} \), but if this is to be so, we must have concession by agent 2 to the demand \( \alpha_2^1 \) with probability approaching 1 at time zero, and in which case, agent 1’s payoff following demand \( \alpha_2^1 \) is again in the limit strictly larger than \( \delta^2 B(\alpha_2^1 - \bar{e}) - D_1^+ \). But then agent 1’s choice of \( \alpha^1 \) which led to a type 1 equilibrium being played in the limit is not optimal, and so we cannot have a type 1 equilibrium played with positive probability in the limit. And hence there can be no delay in agreement in the limit.

**Proof of Lemma 6**

We again prove this by having showing agent \( j \) can obtain a payoff strictly larger than \( (1 + \delta^2 B)\alpha_2^1 \) in the limit. Suppose \( i=1 \), then clearly if agent 1 demands \( \alpha^1 < \alpha_2^1 - \bar{e} L \), where \( L \in (0, 1/2) \) is a constant to be determined, she does strictly worse. Also consider a demand \( \alpha^1 > \alpha_2^1 + \bar{e} L \) with positive probability in the limit, agent 2 can counterdemand \( \alpha_2^1 + \bar{e} L > \alpha_2^1 \), ensuring \( K_2^1 > K_2^1 \), and so agent 1 must concede with probability approaching 1 in the limit (by analysis similar to that of the previous paragraph).

Finally, suppose \( \alpha^1 \in [\alpha_2^1 - \bar{e} L, \alpha_2^1 + \bar{e} L] \). We notice immediately that after any counterdemand we cannot have a type 2 equilibrium in which agent 1 receives mass concession, as we cannot have \((1 - \alpha^1)(1 + \delta^2 B) \leq -D_1^+ + (1 - \alpha^2)\delta^2 B \). We focus first on possible type 4 equilibrium. But now we again notice that \( T_{1m} = T > T_{1m}^* = T_{2m}^* \). Moreover, given that \( L < 1/2 \), we have a lower bound on the benefit of waiting for agent 2, and so \((1 - G(T_{1m})) \geq \bar{e} \). Then the cost to waiting until time \( T \) instead of conceding at \( T_{1m}^* = \bar{e} \delta(D_1^+ + (1 - \alpha^2)) \), while the maximum benefit to waiting even if agent 2 concedes with probability 1 at time \( T \) is \( \delta^2 B(\alpha^1 + \alpha^2 - 1) \). Agent 2 can thus ensure that cost to waiting is less than the benefit (and thus ensure concession at time 0, with probability approaching 1 in the limit) by setting \( \delta^2 \leq \frac{\epsilon(1 + D_1^+ + \delta^2 B_2(1 - \alpha^1)^{1+\alpha})}{\delta(B_2^1 + \delta B_2^1(1 - \alpha^1)^{1+\alpha})} \leq \frac{\epsilon(1 + D_1^+ + \delta^2 B_2(1 - \alpha^1)^{1+\alpha})}{\delta(B_2^1 + \delta B_2^1(1 - \alpha^1)^{1+\alpha})} \). However, for sufficiently small \( L \), we have that the right hand side is strictly smaller than \((1 - \alpha^1) \) and so agent 2 can again guarantee a payoff strictly larger than \((1 + \delta^2 B)\alpha_2^1 \) in the limit. Let the counterdemand that guarantees concession against \( \alpha^1 \in [\alpha_2^1 - \bar{e} L, \alpha_2^1 + \bar{e} L] \) for sufficiently small \( L \), be \( \alpha^2 > \alpha_2^1 \). Against such a counterdemand, agent 1 is not prepared to wait between \( T_{1m}^* \) and \( T \), even to receive probability 1 concession there, and so clearly cannot wait between 0 and \( T \) either, so against such a counterdemand there is no type 1 continuation equilibrium either.

A similar procedure shows that agent 1 can obtain a limiting payoff strictly larger than \((1 + \delta^2 B)\alpha_2^1 \) when for \( \alpha^2 \in [\alpha_2^1 - \bar{e}, \alpha_2^1 + \bar{e}] \), we have \( \delta^2 \alpha^2 > \alpha_2^1 + \bar{e} \), but \( \delta^1 \alpha^1 \approx \alpha^1 \). Again we cannot have a type 2 equilibrium for any counterdemand as \((1 - \alpha^1)(1 + \delta^2 B) \leq -D_1^+ + (1 - \alpha^1)\delta^2 B \).

Looking then to a counterdemand to \( \alpha^1 \) such that \( \alpha^2 \in [\alpha_2^1 - \bar{e} L, \alpha_2^1 + \bar{e} L] \), for \( L < 1/2 \), such that we have a type 4 equilibrium, we have a lower bound on the benefit of waiting for agent 1, and so \((1 - G(T_{1m})) \geq \bar{e} \). Then, let \( \alpha^{1'} = \min[\alpha_2^1 + \bar{e} L, \alpha_2^1 - \bar{e} L, \alpha_2^1 + \bar{e} L] \). For sufficiently small \( L \) and a rich type space.
Now notice that against \( a_{1}^{T} \) agent 2 is unwilling to wait from time \( T_{1m} \) until time \( T \), even if that did induce probability 1 concession at time \( T \). And so, agent 2 is certainly not willing to wait from time 0 either, so this will certainly not induce a type 1 continuation equilibrium, irrespective of whether \( K_{2}^{*} > K_{1}^{*} \). Clearly a similar analysis holds for any counterdemand greater than \( a_{1}^{T} + \delta L \), even if agent \( K_{2}^{*} > K_{1}^{*} \), agent 2 cannot be induced to wait, and so must concede with probability 1 in the limit. Finally, given that \( a_{1}^{T} \leq a_{1}^{T} + \delta L \), we do not need to worry about counterdemands less than \( a_{1}^{T} + \epsilon L \).

**Proof of Lemma 7**

The existence of a unique equilibrium of type 1 or 2 is trivial, given appropriate parameters described above, is obvious. The only case we must be concerned with is for a type 3 equilibrium. In this case, we must have \( T_{1e}^{*} < T_{1e} \) and \( T_{2e}^{*} < T_{1m}^{*} \), and certainly cannot have mass concession by agent 2 at time zero. Clearly \( T_{1e}^{*} \), \( T_{1m}^{*} \), and \( T_{2e}^{*} \) can be defined for any initial probabilities of agents irrationality, taking \( T_{1m}^{*} = 0 \) when \( \delta \bar{V}_{2}(z_{1}^{*}) - D_{1}^{*} \geq (1 - a_{2}^{T}) + B(1 - maxC) \), moreover, these times are clearly continuous in agents initial levels of probability of irrationality. Given this we can define \( T_{1e}^{3}(c_{1}) \), \( T_{2e}^{3}(c_{1}) \), \( T_{1m}^{3}(c_{1}) \), representing such times defined when agent 1’s prior probability of irrationality is given by \( \frac{\alpha}{\alpha} \).

As discussed above, reducing \( c_{1} \) increases \( z_{1}^{*}(t) \) and hence \( V_{2}(z_{1}^{*}(t)) \), this slows down agent 2’s rate of concession, and thus increases \( T_{2e}^{3}(c_{1}) \). Similarly, it must decrease \( T_{1e}^{3}(c_{1}) \), because agent 2’s initial probability of irrationality is higher, and also decrease \( T_{1m}^{3}(c_{1}) \), because \( V_{2}(z_{1}^{*}(t)) \) is higher.

Given that \( T_{1e}^{3}(1) > T_{1e}(1) \), while \( T_{2e}^{3}(c_{1}) > T_{1m}^{3}(c_{1}) = 0 \), by monotonicity (and continuity) we must have a unique level of initial concession by agent 1 such that \( T_{1e}^{3}(c_{1}) = T_{1m}^{3}(c_{1}) \), let such a time be \( T_{1e} \). If there is also a level of initial concession such that \( T_{1e}^{3}(c_{1}) = T_{1m}^{3}(c_{1}) \) (which must also be unique by monotonicity) let this be \( T_{1m} \). Clearly our only candidates for equilibrium involve both agents conceding continuously up to \( T_{1m} \) or \( T_{1e} \), at which point, by construction agent 2 reaches probability 1 of irrationality.

If \( T_{1m}^{*} < T_{1e}^{*} \) then agent 1 reaches \( T_{1m}^{*} \) with a probability of irrationality less than 1, and with the residual probability waits until time \( T \). Suppose then there is another equilibrium in which both agents concede continuously up to \( T_{1e}^{*} \). At point \( T_{1e} \) agent 1 must have a probability of irrationality equal to 1, but \( V_{2}(1) > V_{2}(c_{1}(T_{1m}^{*})) \) but at time \( T_{1e} \) there is also a lower cost to waiting until time \( T \), \( 1 - G_{1}(T_{1e})(1 - a_{2}^{T} + D_{1}) \), and so with greater benefits and lower costs agent 1 must strictly prefer waiting until \( T \), rather than conceding immediately; so this clearly cannot be an equilibrium.

If there is no \( T_{1m}^{*} \) (i.e. \( T_{1e}^{3}(c_{1}) < T_{1m}^{3}(c_{1}) \) for \( c_{1} \in [\alpha, 1] \)) then clearly the only possibility for equilibrium is for agent 1’s initial mass concession to determined so that agent 1 (and 2) reach \( T_{1e}^{3} \) with a conditional probability of irrationality equal 1. Given that \( T_{1e}^{*} = T_{1e}(c_{1}) < T_{1m}^{*}(c_{1}) \), clearly even at \( T_{1e}^{*} \) where \( V_{2}(T_{1e}^{*}) = V_{2}(1) \) agent 1 has no incentive to wait until time \( T \) instead of conceding in the way described.

If there is some \( T_{1m}^{*} \) such that \( T_{1m}^{*} > T_{1e}^{*} \), then clearly there is an equilibrium as described in the previous paragraph. It is also clear that there is no other equilibrium as for the level of initial concession \((1 - c_{1})\) such that both agents hit probability 1 of irrationality at \( T_{1e}^{*} \), we must have \( T_{2e}^{3}(c_{1}) > T_{1e}^{*} \). And thus initial concession by agent 1 must increase further for us to have \( T_{2e}^{3}(c_{1}) = T_{1e}^{*} \). But clearly then rational agents of agent 1 must be exhausted more quickly, meaning \( T_{1m}^{3}(c_{1}) \) decreases. But then clearly we cannot have an equilibrium with continuous concession (by both agents) up to \( T_{1m}^{*} \) as type 1 agents must already be exhausted.

**Proof of Lemma 8**

To see that this is so, first notice that we must have that \( T_{1e}^{3} \) converges to \( T \) in the limit (this is so even when setting \( B(t) = 0 \) for all \( t \)). So, whether \( T_{1e}^{3} < T_{1m}^{3} \) depends crucially on whether \( V_{2}(T_{1e}^{3}) \) is bounded away from \( B(1 - maxC) \) as we approach the limit. If this holds then we eventually have \( T_{1m}^{3} < T_{1e}^{3} \), as agent 1 would have a strict incentive to wait until time \( T \) instead of conceding at time \( T_{1e}^{3} \) (We must have \( 1 - G(T_{1e}^{3})(1 - a_{2}^{T} + D_{1}) < \delta^{3}(V_{2}(T_{1e}^{3}) - B(1 - maxC)) \) as \( T_{1e}^{3} \to T \)). If however \( V_{2}(T_{1m}^{3}) = B(1 - maxC) \) as we approach the limit then, it is equally clear we have \( 1 - G(T_{1e}^{3})(1 - a_{2}^{T} + D_{1}) > \delta^{3}(V_{2}(T_{1e}^{3}) - B(1 - maxC)) = 0 \) and so \( T_{1e}^{3} < T_{1m}^{3} \). In cases where \( T_{1e}^{3} < T_{1m}^{3} \) as we approach the limit we can also show that we also that \( T_{2e}^{3} < T_{1e}^{3} \), and so whether \( V_{2}(T_{1e}^{3}) \) is bounded away from \( B(1 - maxC) \) entirely determines whether agent 1 or agent 2 concedes in the war of attrition.

Suppose then that agent 1 demands \( a_{1}^{T} \leq \delta \bar{a} \), and suppose agent 2 counterdemands \( a_{2}^{T} > (1 - a_{1}^{T}) \) with positive probability in the limit. Then in this case, suppose that we have \( V_{2}(T_{1e}^{3}) \to B(1 - maxC) \) as we approach the limit, and so \( B(t) \to 0 \) for all \( t \leq T_{1e}^{3} \). Then conditional on no concession by agent 1 at time 0, we have.

48
\begin{align*}
\bar{z}^3(T_{1e}^2) &= \frac{\bar{z}^3}{|1 - G(T_{1e}^2)|^{\hat{\alpha}}_1} < 1, \text{ while } [1 - G(T_{1e}^2)]^{\hat{\alpha}}_1 \to \bar{z}^3.
\end{align*}

Now consider what happens if agent 3 counterdemands \( a^3 = \max C^3 \) with positive probability in the limit, when agent 1 arrives at time \( T \) with reputation \( \bar{z}^3(T_{1e}^2) \), then in this continuation game we eventually have:

\begin{align*}
1 - c^3 &= 1 - \bar{z}^3(\bar{z}^3(T_{1e}^2)^{\hat{\alpha}}_2) \bar{z}^3 \left( 1 - \bar{z}^3(\bar{z}^3(T_{1e}^2)^{\hat{\alpha}}_2) (\bar{z}^3(T_{1e}^2)^{\hat{\alpha}}_2)^2 \right) \to 1
\end{align*}

Where \( \varepsilon(z) \geq 0 \) must converge to zero as \( \varepsilon_1 \) does and merely takes account of the fact that we may have \( [1 - G(T_{1e}^2)] \) slightly smaller than \( \bar{z}^3 \) as we approach the limit. To see the conclusion that \( 1 - c^3 \to 1 \) notice, the conclusion must be true if \( 1 < \frac{\bar{z}^3_1}{\bar{z}^3} + \frac{\bar{z}^3_2}{\bar{z}^3} = \varphi(\alpha^3, \alpha^2) \). But, this is precisely what we know must be true is \( \alpha^3 \leq \hat{\alpha}^3 \). But in which case we have must have that \( V^1(T_{1e}^2) \) bounded away from \( B(1 - \max C^3) \) in the limit, which is a contradiction.

Given this we we must have \( T_{1m}^{le} < T_{1e}^2 \) in the limit. To see that this implies that agent 2 must concede to agent 1 with probability approaching 1 in the limit, suppose that there was concession with only probability \( \kappa < 1 \) in the limit, this would mean that agent 2’s conditional probability of irrationality is at time \( 0^+ \) would be only boundedly greater than 1’s and so all the above arguments would still go through ensuring that \( T_{1m}^{le} < T_{1e}^2 \) in the subgame at time 0 which we clearly cannot have.

Suppose then that \( \alpha^3 > \hat{\alpha}^3 \) then agent 2 can demand \( \alpha^2 > (1 - \alpha^3) \) such that it is arbitrarily close (from below) to that for which \( \varphi(\alpha^1, \alpha^2) = 1 \), so ensuring that \( \frac{\bar{z}^3_1}{\bar{z}^3} + \frac{\bar{z}^3_2}{\bar{z}^3} < 1 \). Given this we can verify in a manner similar to that above, that we must have \( V^1(T_{1e}^2) = B(1 - \max C^3) \), and \( B(t) = 0 \) for all \( t \leq T_{1e}^2 \) in the limit and so \( T_{1m}^{le} > T_{1e}^2 \).

We must also have \( T_{1e}^{le} > T_{1e}^2 \) given that in this case we clearly have \( \frac{\bar{z}^3_1}{\bar{z}^3} < 1 \) and so mass concession must be made by agent 1. A similar argument to that above shows concession must be with probability approaching 1 in the limit. Having revealed rationality in bargaining over issue 1, agent 2 must then concede immediately to the demand \( \max C^3 \) on bargaining issue 2. A more aggressive counterdemand by agent 2 would ensure immediate mass concession in the limit, by similar analysis.

\section*{References}


Dilip Abreu, David Pearce, and Ennio Stacchetti. One sided uncertainty and delay in reputational bargaining. \textit{Unpublished manuscript}, 2012. 3, 26

Peter B. Zemsky Christopher Avery. Option values and bargaining delays. \textit{Games and Economic Behavior}, 7(2):139153, September 1994. 4

Chaim Fershtman. The importance of the agenda in bargaining. \textit{Games and Economic Behavior}, 2(3):224238, September 1990. 4


