Bargaining over Contingent Contracts Under Incomplete Information

Geoffroy de Clippel* Jack Fanning* Kareen Rozen*

Updated June 2020

Abstract

We study bargaining over contingent contracts in problems where private information becomes public or verifiable when the time comes to implement the agreement. We suggest a simple, two-stage game that incorporates important aspects of bargaining. We characterize equilibria in which parties always reach agreement, and study their limits as bargaining frictions vanish. Under mild regularity conditions, we show all interim-efficient limits belong to Myerson (1984)'s axiomatic solution. Furthermore, all limits must be interim-efficient if equilibria are required to be sequential. Results extend to other bargaining protocols.

*Department of Economics, Brown University. We are grateful to seminar participants at Northwestern and participants at the BEET, Cowles, SAET, and Stony Brook Game Theory conferences for valuable comments.

1 Introduction

Parties often come to the bargaining table holding private information. If that information becomes public upon the agreement's implementation, then the terms of the contract can be made contingent on that future information. Contingent contracts play a central role in many economic models. Arrow-Debreu securities, options, futures and other derivatives are all contingent contracts. The first appearance of contingent contracts to study cooperation under incomplete information dates back to Wilson (1978). Bazerman and Gillespie (1999) emphasize to practitioners the importance of considering contingent contracts in different bargaining scenarios, including those with incomplete information. But which specific terms should one expect as a result of negotiations?

For instance, suppose a laptop manufacturer and a microchip supplier bargain over future monetary proceeds. The supplier knows whether it will be able to provide an older chip (Old) or a new-generation chip (New) by the date production starts. The laptop manufacturer knows the type of the other components it will use in the laptop (e.g. screen, memory modules, fan, etc.). The components may be relatively old (Old) or of the newest generation (New). This gives rise to four ex-post verifiable states of the world, (Old, Old), (Old, New), (New, Old) and (New, New). The sales profit when the manufacturer uses the older components is \$12M independently of the chip, as old components cannot exploit the benefits of the new chip. Fitting an older chip in a machine with new components lowers profit to \$9M due to compatibility issues, while machines with the newest-generation components and chip generate the highest profit, \$15M. The laptop manufacturer and the chip supplier each believe the other has probability 1/2 of having new-generation hardware available when production starts. The laptop manufacturer is risk neutral ($u_1(x) = x$) while the chip supplier, a privately held firm, is risk averse ($u_2(x) = \sqrt{x}$). The ex-post utility set in state t for a given profit $M(t) \in \{9, 12, 15\}$ is then $U(t) = \{v \in \mathbb{R}^2_+ : v_1 + v_2^2 \leq M(t)\}$.

The Nash bargaining solution is focal in complete information settings. When information is incomplete, as in the above example, writing a contract that picks the Nash bargaining solution for each ex-post informational state may sound reasonable at first. Given a profit m, the Nash solution is obtained by maximizing $(m - v_2^2)v_2$, the product of utilities over the feasible utility set, and results in giving one-third of the profit to the chip supplier. Thus, the ex-post Nash contingent contract distributes profits as follows:

Ex-post Nash	*Old Chip*	*New Chip*
Old Components	$\$8M, \$4M$	$\$8M, \$4M$
New Components	$\$6M, \$3M$	$\$10M, \$5M$

Notice that whatever his type, the chip supplier faces a substantial risk of $\pm\$0.5M$ with equal probability. This is inefficient at the *interim* stage, that is, given the bargainers' information when they negotiate. For instance, it is possible to rearrange the laptop manufacturer's payoff, while keeping his expected utility constant, to construct a contingent contract that fully insures the chip supplier. This inefficiency is quite pervasive: settling on ex-post Nash solutions is generically interim-inefficient in smooth bargaining problems (see the Online Appendix).[1] Of course, there are many possible interim efficient contracts. What would be reasonable, interim efficient terms of trade?

Providing a first answer, Harsanyi and Selten (1972) and Myerson (1984) axiomatically characterize two distinct extensions of the Nash bargaining solution to related incomplete information settings (that impose interim efficiency as an axiom). Loosely speaking, *Harsanyi-Selten's solution* selects contingent contracts that maximize the probability-weighted product of interim utilities. Applied to our example, it awards the chip-supplier with $\$4M$ in all states of the world:

Harsanyi-Selten	*Old Chip*	*New Chip*
Old Components	$\$8M, \$4M$	$\$8M, \$4M$
New Components	$\$5M, \$4M$	$\$11M, \$4M$

Myerson's solution, on the other hand, selects contingent contracts that are both equitable and efficient for a rescaling of the interim utilities. In general, it incorporates agents' incentive constraints to truthfully reveal their type, but these constraints are not needed in our setting with verifiable types. Applied to our example, it rewards the chip supplier with \$4.5M for a newer chip (associated with weakly larger profits) but gives him only \$3.5M for an old chip, irrespective of the laptop maker's type:

Myerson	*Old Chip*	*New Chip*
Old Components	$\$8.5M, \$3.5M$	$\$7.5M, \$4.5M$
New Components	$\$5.5M, \$3.5M$	$\$10.5M, \$4.5M$

[1] If ex-post Nash happens to be interim efficient in a bargaining problem where both agents have at least two types, then it becomes inefficient when one agent's utility is rescaled in some state.

As an alternative to axioms, we can also adopt a non-cooperative approach, studying equilibrium outcomes of bargaining games, and their properties. Remarkably, in a wide variety of complete information bargaining games, those outcomes have been found to be close to Nash's axiomatic solution. What happens under incomplete information? The vast existing literature on noncooperative bargaining under incomplete information (see surveys by Osborne and Rubinstein (1990), Binmore et al. (1992), Kennan and Wilson (1993) and Ausubel et al. (2002)) provides no answer due its almost exclusive focus on bargaining over direct terms of trade (e.g. what quantity to trade at what price), instead of considering contingent contracts or incentive compatible mechanisms (as envisaged by the earlier axiomatic papers).[2]

In this paper, we follow the non-cooperative approach and find a surprising link between equilibrium outcomes and Myerson's extension of the Nash solution, in a simple two-stage bargaining game with contingent contracts.[3] The first stage of the game is a *demand/offer stage*, where each bargainer independently suggests a state-contingent contract. The second stage is a *bargaining posture stage*, where each bargainer independently decides whether to insist on her own contract, or to be conciliatory and accept her opponent's contract. When both parties are conciliatory, bargaining is equally likely to end with an agreement on either contract. If an agent insists on her own contract, her aggressiveness leads to a (small) probability of disagreement even when her opponent is conciliatory. If both insist, there is disagreement.

[2]Bargaining over contingent contracts was first studied in de Clippel and Minelli (2004). Aside from providing axiomatic results, they suggest a stringent refinement to obtain Myerson (1983)'s principal-agent solution as equilibrium outcomes of the simple take-it-or-leave-it offer protocol. Clearly, this protocol entirely favors the agent making the offer, and outcomes are generally unrelated to Myerson (1984)'s bargaining solution. By contrast, in our bargaining game below, both agents enjoy equal opportunities, which completely changes the analysis and results. Kim (2017) studies a narrow class of bargaining problems where Myerson (1983)'s principal-agent and Myerson (1984)'s bargaining solutions do coincide. Again, standard refinements fail to select Myerson's solution in the take-it-or-leave-it offer protocol, but Myerson (1989)'s 'coherent' equilibria do.

[3]Given its prominence under complete information, an alternating-offers protocol is another natural choice. But the repeated offers' potential for signaling and information leakage makes it very challenging to analyze. Even in our simpler game, the analysis is far from trivial. Okada (2016) considers that environment, but imposes three stringent refinements that effectively require demands to match those under complete information in every state, which therefore imply convergence to the ex-post Nash solution as agents become patient. In particular, he requires that demands don't change even if agents' learn about the state from their opponent's behavior, and each type interprets deviations as coming from opponent types that would strictly benefit ex-post from such a proposal if accepted. Without the first requirement, it is easy to construct examples where many interim-efficient contracts are supported as weak PBE. The second requirement necessarily violates the standard no-signalling-what-you-don't-know principle.

Though simple, this protocol encompasses some basic features of bargaining, with offers, demands, and posturing. The posturing stage with its risk of disagreement when one bargainer insists is a natural way to introduce frictions into the Nash 'demand game', thereby circumventing the well-known, anything-goes situation that arises otherwise. We are particularly interested in what happens when frictions are vanishingly small. Our payoff structure is rather general and can accommodate different scenarios. For instance, it corresponds to an environment where offers go astray with positive probability (a model studied by Evans (2003) under complete information), as well as an environment where each party's acceptance of the other's proposal is stochastically delayed. Our equilibrium characterization and results also extend to stationary equilibria when posturing is modeled as a war of attrition.

Our main results go as follows. We first characterize a focal set of equilibria: those in which bargainers formulate deterministic (pure-strategy) demands and are conciliatory on path, so they always reach an agreement after equilibrium demands. We call these 'conciliatory equilibria'. Under complete information, there is a unique pure strategy equilibrium, which is also the only conciliatory equilibrium; and it converges to the Nash bargaining solution as bargaining frictions vanish. What happens as frictions vanish under incomplete information? Under some mild conditions on the bargaining problem, we show that interim-efficient limits of conciliatory equilibria not only exist, but must all be Myerson solutions. Moreover, all limits must be interim efficient when conciliatory equilibria satisfy Fudenberg and Tirole (1991)'s 'no-signalling-what-you-don't-know' principle. Imposing this principle on weak perfect-Bayesian equilibrium corresponds to a natural extension of sequential equilibrium in our infinite game. Combining these results, we get a strong prediction: *under mild conditions, all limits of sequential, conciliatory equilibria are Myerson solutions.*

Beyond their stand-alone value, our results combined with those of Myerson (1984) can be seen as an extension of the 'Nash program' (of justifying a solution both axiomatically and non-cooperatively) to problems of incomplete information. Our results suggest that cross-agent and cross-type tradeoffs in the solution that Myerson derived axiomatically are, at some level, well justified. Even without requiring equilibria to be sequential, we can rule out non-Myerson interim-efficient solutions, including the Harsanyi-Selten solution. This is quite unusual for two-sided, incomplete-information bargaining problems, where the opportunity to interpret deviations as coming from an opponent's 'worst' possible type can often make the equilibrium set so large that it is hard to say anything meaningful about expected outcomes (e.g. see the discussion in Ausubel et al. (2002)). The ability to offer contingent contracts in our setting helps limit the power of 'punishing with beliefs,' because an agent can offer a contract that would be acceptable to his opponent in every state of the world, and so secure payoffs associated with his true type (see the notion of best-safe payoff in Section 3.2).

We believe the setting of verifiable information is interesting in its own right. However, we also hope it provides a building block for the analysis of more general bargaining settings in which parties have unverifiable private information, in which case they would bargain over communication protocols (or mechanisms) to determine how the agreement might vary with the information they share. This is a topic we hope to address in future work. Our ability to rule out any non-Myerson outcome in the case of verifiable types, certainly provides grounds for skepticism of the reasonableness of non-Myerson solutions with non-verifiable types.

2 Framework

We consider an incomplete-information setting with two agents at the bargaining table. Agent $i = 1, 2$ has a finite set of possible *types* T_i. For now, we assume agents share a *common prior* p with full support over the type profiles (also called *states*) in $T = T_1 \times T_2$. Results extend to non-common priors, as discussed in Section 5.2. The set of states consistent with type t_i is defined as $T(t_i) = \{\hat{t} \in T : \hat{t}_i = t_i\}$. These states become public or verifiable when the time comes to implement an agreement.

Each type profile $t \in T$ is associated with an *ex-post utility possibility set* (or feasible utility set) $U(t) \subset \mathbb{R}^2_+$. The collection of ex-post utility sets is $U = \times_{t \in T} U(t)$. We assume $U(t)$ is convex, compact, and contains its disagreement point $(0, 0)$ for all $t \in T$ (e.g., absence of trade or production). The assumption that all ex-post utilities are larger or equal to the disagreement payoff ($U(t) \subset \mathbb{R}^2_+$) is substantive, but is met in many applications and must hold if agents retain the possibility of taking their disagreement payoff at the ex-post stage. A *bargaining problem* is summarized by the tuple $\mathcal{B} = (T, U, p)$.

Let $\overline{u}_i(t)$ be the highest utility that Agent i can get in state t; and let $\underline{u}_i(t)$ be the highest utility i can get *conditional on j getting $\overline{u}_j(t)$*:

$$\overline{u}_i(t) = \max_{u \in U(t)} u_i$$

$$\underline{u}_i(t) = \max_{u \in U(t):\, u_j = \overline{u}_j(t)} u_i,$$

for $i = 1, 2$. If $U(t)$ has no flat part, then agent i would pick $(\underline{u}_j(t), \overline{u}_i(t))$ if he were a dictator. If multiple options achieve his best utility $\overline{u}_i(t)$, then i would be indifferent between all of them, and $(\underline{u}_j(t), \overline{u}_i(t))$ is the one that is most favorable to j (guaranteeing ex-post efficiency). We assume throughout that $\overline{u}_i(t) > \underline{u}_i(t)$ for all t, so that there is always something to bargain over.

We study the bargaining problem $\mathcal{B}$ at the *interim stage*: each agent knows his own type, but not the type of his opponent. Formally, a bargaining agreement is a *contingent contract* $u \in U$, which associates a utility profile $u(t) \in U(t)$ for each $t \in T$. Different bargaining problems may involve different underlying variables (e.g., the split of profits, the quantity or price of a good to be sold, the time a service is rendered). Describing a bargaining agreement by the resulting utility profiles is a notationally convenient and unifying device to encapsulate bargainers' considerations.

Agents evaluate a contingent contract by its expected utility, with beliefs regarding the other's type derived from the prior p by Bayes' rule. The expected utility from the contingent contract u to bargainer i of type $t_i \in T_i$ is:

$$E[u_i|t_i] = \sum_{t \in T(t_i)} p(t|t_i) u_i(t).$$

The contingent contract x is *interim efficient* if it is not interim-Pareto dominated by another contingent contract; that is, there is no $u \in U$ such that $E[u_i|t_i] \geq E[x_i|t_i]$ for all i and t_i, with strict inequality for some i and t_i.[4] The contingent contract x is *ex-post efficient* if, for every $t \in T$, $x(t)$ is Pareto efficient within $U(t)$; that is, for all t, there is no $a \in U(t)$ with $a_i \geq x_i(t)$ for all i and strict inequality for some i.

Our verifiable-state product structure ($T = T_1 \times T_2$) fits many settings of private information, but not all. For instance, firms considering a joint venture may have private information about the variety of outputs they can produce and the prices those outputs will fetch, the availability of different inputs and their prices, the physical locations of and existing contracts with suppliers and customers, all of which might

[4]Aside from introducing interim efficiency, Holmström and Myerson (1983) also discusses the related, strategic notion of *durability*. They show the two concepts are distinct in general. It is easy to check they coincide in our context (contingent contracts without IC constraints).

become verifiable ex-post. It does not, however, fit situations where there is a verifiable state of the world (e.g. the weather, future stock prices) which agents have *unverifiable* beliefs about.

Despite calling $U(t)$ an ex-post utility set, residual uncertainty may be present even after the state is known. For instance, suppose an oil Firm 1 already conducted a survey of oil reserves in different locations, while an engineering Firm 2 conducted a survey on the difficulty of oil extraction and transportation at those locations. These surveys represent agents' verifiable types. When pooled, they allow firms to extract oil from an optimal location, although the quantity of oil and the extraction costs remain random variables.

2.1 Efficiency and Weighted Utilitarianism

The *interim utility-possibility set* $\mathcal{U}(\mathcal{B})$ is the set of interim utilities $(E[x_i|t_i])_{i,t_i}$ achievable through contracts x for the bargaining problem $\mathcal{B}$. This set inherits compactness and convexity from each $U(t)$. By the supporting-hyperplane theorem, if a contract x is interim efficient, then there is a nonzero vector of weights $\hat{\lambda} = (\hat{\lambda}_i(t_i))_{i,t_i} \in \mathbb{R}_+^{T_1} \times \mathbb{R}_+^{T_2}$ such that $\sum_{i=1,2}\sum_{t_i \in T_i} \hat{\lambda}_i(t_i)E[u_i|t_i]$ is maximized within $\mathcal{U}(\mathcal{B})$ by the contract $u = x$. In this case, we say $\hat{\lambda}$ is *interim orthogonal* to $\mathcal{U}(\mathcal{B})$ at x.[5] Similarly, for each $t \in T$, if $x(t)$ is Pareto efficient within $U(t)$ then there is a nonzero vector of weights $\lambda(t) \in \mathbb{R}_+^2$ such that $\sum_{i=1,2} \lambda_i(t)a_i(t)$ is maximized within $U(t)$ by the allocation $x(t)$. In this case, we say $\lambda(t)$ is *ex-post orthogonal* to $U(t)$ at $x(t)$. The lemma below summarizes useful relationships, and is proved in the Online Appendix.

Lemma 1. *The following relationships hold:*

(i) If the allocation rule x is interim efficient, then it is ex-post efficient.

(ii) If x is interim efficient, then there is a non-zero $\hat{\lambda} \in \mathbb{R}_+^{T_1} \times \mathbb{R}_+^{T_2}$ which is interim orthogonal to $\mathcal{U}(\mathcal{B})$ at x. Conversely, if a vector $\hat{\lambda} \in \mathbb{R}_{++}^{T_1} \times \mathbb{R}_{++}^{T_2}$ is interim orthogonal to $\mathcal{U}(\mathcal{B})$ at x, then x is interim efficient.

(iii) If x is ex-post efficient, then for each $t \in T$ there is a non-zero $\lambda(t) \in \mathbb{R}_+^2$ which is ex-post orthogonal to $U(t)$ at $x(t)$. Conversely, if $\lambda(t) \in \mathbb{R}_{++}^2$ is ex-post orthogonal to $U(t)$ at $x(t)$ for each t, then x is ex-post efficient.

[5] More formally, $\hat{\lambda}$ is orthogonal to $\mathcal{U}(\mathcal{B})$ at the interim utility vector associated with x.

(iv) $\hat{\lambda} \in \mathbb{R}_+^{T_1} \times \mathbb{R}_+^{T_2}$ *is interim orthogonal to* $\mathcal{U}(\mathcal{B})$ *at* x *if, and only if,* $\lambda(t) = \left(\frac{\hat{\lambda}_1(t_1)}{p(t_1)}, \frac{\hat{\lambda}_2(t_2)}{p(t_2)}\right)$ *is ex-post orthogonal to* $U(t)$ *at* $x(t)$ *for all* $t \in T$.

We will sometimes assume the bargaining problem is *smooth*, meaning for each t and ex-post efficient $u \gg 0$ in $U(t)$, there is a unique orthogonal vector to $U(t)$ at u.

2.2 Myerson Solution

Under complete information, the Nash bargaining solution is obtained by maximizing the product of the two agents' utility gains over the utility possibility set. The *ex-post Nash solution* gives agents the Nash solution in every state of the world. While this solution is clearly ex-post efficient, it is generically interim inefficient for smooth bargaining problems (see Online Appendix).

In the hope of attaining interim efficiency, one way to extend Nash's solution to accommodate incomplete information would be to introduce some interim welfare function W and maximize it over the set of all feasible contingent contracts. This is in fact the path followed by Harsanyi and Selten (1972) whose bargaining solution adapted to the present framework maximizes $\prod_{i=1,2} \prod_{t_i \in T_i} (E[x_i|t_i])^{p(t_i)}$ over the set of feasible contingent contracts x.

By contrast, Myerson (1984)'s bargaining solution is not derived from the maximization of a social welfare function over interim utilities, but instead defined constructively. While originally defined more generally to accommodate incentive constraints, it boils down to the following in our setting: an allocation rule x is a *Myerson solution* for the bargaining problem $\mathcal{B}$ if there is $\hat{\lambda} \in \Delta_{++}(T_1) \times \Delta_{++}(T_2)$ such that

$$E[x_i|t_i] = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) \frac{p(t_i)}{2\hat{\lambda}_i(t_i)} \max_{v \in U(t)} \sum_{j=1,2} \frac{\hat{\lambda}_j(t_j)}{p(t_j)} v_j, \tag{1}$$

for all $t_i \in T_i$ and $i = 1, 2$. A Myerson solution is always interim efficient. The set of Myerson solutions for $\mathcal{B}$ is denoted $MY(\mathcal{B})$. As the reader may check, if the ex-post Nash solution of a smooth bargaining problem happens to be interim efficient, then it is also a Myerson solution.

That a Myerson solution exists in our framework is an additional implication of our convergence results; but typically, few contracts meet the requirements.[6] Take an

[6]For our laptop-maker and chip-supplier example there is a unique Myerson solution, but in general this need not be true.

interim-efficient contract x with a strictly positive interim-orthogonal vector $\hat{\lambda}$. From Lemma 1, for each $t \in T$, the ex-post orthogonal vector to $x(t)$ is $\lambda(t) = (\frac{\hat{\lambda}_1(t_1)}{p(t_1)}, \frac{\hat{\lambda}_2(t_2)}{p(t_2)})$. The following three-step process identifies whether x is a Myerson solution:

Step 1. For each $t \in T$, construct from $U(t)$ a 'linearized' ex-post utility possibility set $V_\lambda(t) := \{w \in \mathbb{R}^2_+ : \lambda(t) \cdot w \leq \lambda(t) \cdot x(t) = \max_{v \in U(t)} \lambda(t) \cdot v\}$, which permits transfers using the weights defined in $\lambda(t)$.

Step 2. Find the Nash solution for $V_\lambda(t)$ by picking the midpoint $m(t)$ of the efficient frontier, that is, $m_i(t) = \frac{p(t_i)}{2\lambda_i(t_i)} \max_{v \in U(t)} \lambda(t) \cdot v$, for $i = 1, 2$.

Step 3. Finally, x is a Myerson solution if it gives both bargainers the exact same interim utilities as the contingent contract m.

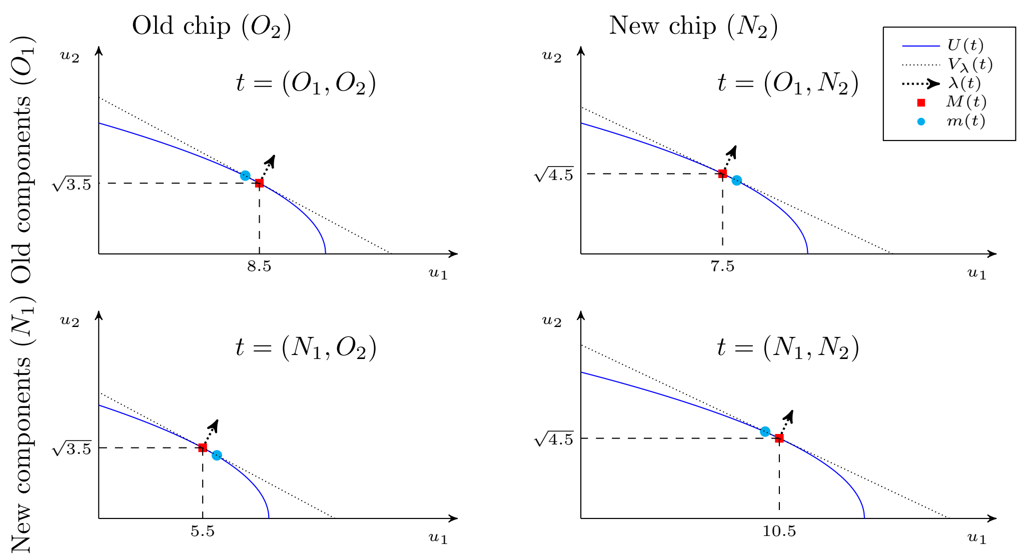

Figure 1: Procedure for finding Myerson solution in the introductory example.

Figure 1 illustrates this procedure for our example from the Introduction of the laptop manufacturer (Agent 1) and microchip supplier (Agent 2). Let the old and new types of Agent i be O_i and N_i, respectively. We can verify M is a Myerson solution because it delivers the same interim utilities as m. For instance, type O_1's gain of $M_1(O_1, O_1) - m_1(O_1, O_1) = \$0.75M$ relative to the midpoint in state (O_1, O_1) is exactly offset by his loss of $M_1(O_1, N_1) - m_1(O_1, N_1) = -\$0.75M$ relative to the midpoint in state (O_1, N_1).

Myerson derived this solution using three main axioms: probability invariance (a generalization of invariance to rescaling utilities), a suitably adapted version of

	A	C
A	0	δx
C	δy	$\frac{x+y}{2}$

Table 1: Prevailing Contingent Contract as a Function of Bargaining Stand (x is 1's offer; y is 2's offer; A=Aggressive; C=Conciliatory; 1 picks row)

independence over irrelevant alternatives, and a random dictatorship axiom. The last is an adaptation of Nash's symmetry axiom, but merits further explanation. Suppose there is a 'strong' solution, as in Myerson (1983), to the modified bargaining problem where the first agent has all the bargaining power and can make a take-it-or-leave-it offer to his opponent, and also a strong solution when the second agent has all the bargaining power. If taking a 50/50 mixture of these two solutions gives an interim-efficient contract, then the random dictatorship axiom states this mixture contract is a solution to the original problem with bargaining power on both sides.

2.3 Non-Cooperative Bargaining Protocol

We summarize our two-stage bargaining protocol, discussed in the Introduction, as follows. First, each agent $i = 1, 2$ simultaneously sends the other agent a proposed contract (an element of U). Agents choose a bargaining posture after observing the offers. If both take a conciliatory stand, then each contract is equally likely to be implemented. A risk of disagreement arises, however, if someone takes an aggressive stand. If one bargainer intransigently insists on his terms and the other is conciliatory, then the insistent agent's offer is implemented but payoffs are discounted by $\delta \in [0, 1)$.[7] The disagreement point prevails if both take an aggressive stand. Table 1 summarizes this information.

Our solution concept is (weak) Perfect Bayesian Equilibrium (PBE). An agent's strategy specifies which offer/demand to make, and which bargaining posture to adopt for each conceivable pair of offers. Throughout the paper we use the letter x (resp., y) to denote contingent contracts proposed by Agent 1 (resp., 2). An agent's belief system specifies a probability distribution over his opponent's types for each conceivable pair of offers. A PBE then consists of a strategy and belief system for each agent, such that each agent's strategy maximizes his expected payoff at each information set given his belief and opponent's strategy, with beliefs given by Bayes' rule whenever

[7]Equivalently, players agree on the insistent demand with probability δ and otherwise disagree.

possible. To clarify, we impose no restrictions on the beliefs of different types of an agent following an opponent's off-equilibrium path offer. We discuss the implications of requiring consistent beliefs in subsection 4.3.

As hinted in the Introduction, our bargaining protocol admits many interpretations beyond the above scenario. For increased generality, consider the variant where payoffs are discounted by a factor δ' with $\delta' \in (\delta, 1]$ should both bargainers take a conciliatory stand. Introducing δ' does not change the strategic features of our bargaining game. Indeed, the original payoff structure can be recovered by dividing all payoffs by δ', which simply amounts to rescaling the discount factor in case a single bargainer takes an aggressive stand.

Consider the bargaining protocol introduced and analyzed by Evans (2003) under complete information. Agents formulate demands/offers as before, but each offer goes astray with probability $\varepsilon > 0$. Instead of facing a positive risk of disagreement by insisting on one's demands, agents must decide whether to accept their counterpart's offer without knowing whether their own offer went through. Of course, it would be strategically equivalent for players to decide after the demand/offer stage which offers to accept, before knowing whether they'll receive one. Under this interpretation, participants get (i) $(1-\varepsilon^2)$ times the average of the two contracts, if both accept, (ii) $(1-\varepsilon)$ times the contract suggested by the rejecting party if the other accepts, and (iii) 0 if both reject. This matches the payoffs for $\delta = 1-\varepsilon$ and $\delta' = 1-\varepsilon^2$.

Alternatively, frictions may take the form of delays. Bargainers make acceptance decisions privately and independently at time 0, but these decisions are recorded with a random delay. The first contract accepted is implemented. With exponential discounting, the problem has the payoff structure above with $\delta' = \int\int e^{-r\min\{t_1,t_2\}}dF(t_1)dF(t_2)$ and $\delta = \int e^{-rt_i}dF(t_i)$, where F is an atomless CDF on $\mathbb{R}_+$ determining the time an agent's decision is recorded. Finally, concessions are modeled as an infinite-horizon war of attrition in Section 5.1.

3 Conciliatory Equilibria

Our first main result provides a full characterization of *conciliatory equilibria*, whereby agents formulate deterministic demands, and take conciliatory stands on path.[8] The

[8] When there is a single state of the world, the refinement to pure strategies alone leads to a unique equilibrium, in which postures are conciliatory. With private information, equilibria can exist where

characterization proceeds in two steps. First, we show that for any conciliatory equilibrium, there is a pooling conciliatory equilibrium that generates the same outcome. This is reminiscent of Myerson (1983)'s inscrutability principle for the informed principal problem. Second, we fully characterize pooling conciliatory equilibria.

3.1 Inscrutability Principle

Consider a conciliatory equilibrium. It may be *separating*, in that some types of Agent i propose distinct contracts. The other agent may then infer something about i's type from the offer, influencing his bargaining posture. This is indeed a central feature of bargaining under incomplete information: offers can signal types, thereby impacting which agreements crystallize.

The next result shows, however, that there is no loss of generality in restricting attention to pooling strategies when it comes to conciliatory equilibria. An agent follows a *pooling* strategy if he proposes the same contingent contract independently of his type. This does not mean the intuition in the paragraph above is incorrect, but rather that the signaling which shapes agreements under incomplete information can be incorporated in new contracts that are part of a pooling conciliatory equilibrium with the same outcome.

To illustrate, suppose Agent 1 can be of two types, t_1 or t_1', and there is a conciliatory equilibrium where he proposes the contingent contract x when his type is t_1 and x' when his type is t_1'. Consider the contingent contract that coincides with x when his type is t_1 and with x' when his type is t_1'. It turns out that proposing this alternative contract, independently of his type, is part of another conciliatory equilibrium that generates the exact same outcome. The next proposition proves this, and extends the idea to show that any conciliatory equilibrium can be replicated by a pooling conciliatory equilibrium.

Proposition 1. *For any conciliatory equilibrium, there is a pooling conciliatory equilibrium that yields the same outcome in all states.*

agents adopt pure strategies on the equilibrium path, but some types posture aggressively. Mixed strategy equilibria exist even when there is a single state.

3.2 Characterization

By Proposition 1, we restrict attention to a pooling conciliatory equilibrium (x, y). Since taking an aggressive stand has an intrinsic cost ($\delta < 1$), Agent 1 may prefer to be conciliatory when offered a contract $\hat{y} \neq y$ slightly less appealing than x. This decision typically depends on his beliefs regarding Agent 2's type (updated given $\hat{y}$), and the likelihood Agent 2 insists on $\hat{y}$ following x (an off-path information set, as it follows $\hat{y}$). Still, being conciliatory would be the best response, independently of 1's belief and 2's bargaining stand, if

$$\frac{x_1(t) + \hat{y}_1(t)}{2} > \delta x_1(t) \text{ and } \hat{y}_1(t) > 0.$$

The first (resp., second) inequality guarantees Agent 1's willingness to be conciliatory when Agent 2 is conciliatory (resp., aggressive). Of course, in that case, being conciliatory is the best course of action whatever the mixed-strategy Agent 2 uses at the concession stage. The two inequalities can be rewritten as $\hat{y}_1(t) > \max\{\gamma x_1(t), 0\}$, with

$$\gamma := 2\delta - 1 \in [-1, 1). \tag{2}$$

It is 'safe' for Agent 2 to propose such a contract $\hat{y}$, as Agent 1 will surely be conciliatory. Define Agent 2's *best-safe* payoff given x at the type profile t by:

$$y_2^{bs|x}(t) = \sup\{u_2 \mid u \in U,\ u_1 > \max\{\gamma x_1(t), 0\}\} = \max\{u_2 \mid u \in U,\ u_1 \geq \gamma x_1(t)\}.$$

Since Agent 2 can always deviate to a contract that gives him a payoff arbitrarily close to this best-safe payoff, we conclude that

$$E[y_2|t_1] \geq E[y_2^{bs|x}|t_1], \text{ for all } t_2 \in T_2$$

Similarly, it must be that $E[x_1|t_1] \geq E[x_1^{bs|y}|t_1]$, for all $t_1 \in T_1$, where

$$x_1^{bs|y}(t) = \arg\max\{u_1 \mid u \in U(t), u_2 \geq \gamma y_2(t)\},$$

for each $t \in T$, is Agent 1's best-safe payoff at t given y.[9]

[9] Myerson (1983) introduced the notion of best-safe mechanism for an informed principal. In a more restrictive framework, Maskin and Tirole (1992) uses the notion of best-safe mechanism – called a Rothschild-Stiglitz-Wilson allocation in their paper – to characterize the equilibrium set of

Additionally, an agent's offer cannot be too favorable to himself in a conciliatory equilibrium. Since he anticipates that Agent 1 will be conciliatory given y, Agent 2's expected payoff is $\delta E[y_2|t_2]$ if he takes an aggressive posture, and $E[(x_2 + y_2)/2|t_2]$ if he is conciliatory. For Agent 2 to be conciliatory whatever his type, we must have

$$E[x_2|t_2] \geq \gamma E[y_2|t_2], \text{ for all } t_2 \in T_2.$$

Similarly, for Agent 1 to be conciliatory whatever his type, we must have

$$E[y_1|t_1] \geq \gamma E[x_1|t_1], \text{ for all } t_1 \in T_1.$$

The next proposition shows that the above inequalities, which are necessary for (x, y) to be part of a pooling conciliatory equilibrium, are also sufficient. For notational simplicity, we define $x_2^{bs|y} = \gamma y_2(t)$ and $y_1^{bs|x} = \gamma x_1(t)$.[10]

Proposition 2. *Let x, y be contingent contracts in U. There is a pooling conciliatory equilibrium where all types of Agent* 1 *propose x, and all types of Agent* 2 *propose y, if and only if:*

$$E[x_i|t_i] \geq E[x_i^{bs|y}|t_i] \text{ and } E[y_i|t_i] \geq E[y_i^{bs|x}|t_i] \tag{3}$$

for all $t_i \in T_i$ and all $i = 1, 2$.

Existence of pooling conciliatory equilibria follows as a corollary. For each $y \in U$, let $\hat{x}^{bs|y} \in U$ be the unique ex-post efficient contract such that $\hat{x}_1^{bs|y} = x_1^{bs|y}$. We can define $\hat{y}^{bs|x}$ analogously, for each $x \in U$. It is easy to check that the map associating $(\hat{x}^{bs|y}, \hat{y}^{bs|x})$ to each (x, y) is continuous, and so there is a fixed point (x, y) that satisfies $x(t) = \hat{x}^{bs|y}(t) \geq x^{bs|y}(t)$ and $y(t) = \hat{y}^{bs|x}(t) \geq y^{bs|x}(t)$. We have thus found a pooling conciliatory equilibrium. Indeed, $(x(t), y(t))$ is an ex-post equilibrium for each t: an equilibrium of our bargaining protocol applied to $U(t)$, while assuming that t is common knowledge. This shows equilibrium existence is not an issue; rather, there is typically a large set of equilibrium outcomes.

the informed-principal noncooperative game. The principal's best-safe payoffs are directly expressed in terms of the problem's exogenous variables. By contrast, in the present paper, each agent's best-safe payoff varies with the other agent's equilibrium offer. This is a substantial difference both conceptually and computationally. In particular, the equilibrium set characterization will now involve a fixed-point condition, see below.

[10] Notice that $x^{bs|y}(t)$ and $y^{bs|x}(t)$ need not belong to $U(t)$. Indeed, γ can be negative and, more generally, $U(t)$ need not be comprehensive over $\mathbb{R}_+^2$ either.

4 Vanishing Bargaining Frictions

We are interested in understanding what happens to conciliatory-equilibrium outcomes as the bargaining friction vanishes: that is, when δ, and thus $\gamma = 2\delta - 1$, tend to one. Let $C(\mathcal{B})$ be the set of all such outcomes, that is, those contingent contracts c for which one can find a sequence $\delta^n \to 1$ and a sequence of contingent contracts $c^n \to c$ such that c^n is a conciliatory-equilibrium outcome of the non-cooperative bargaining game associated to δ^n.

For a start, observe that $C(\mathcal{B})$ is nonempty, because it contains the ex-post Nash contingent contract. As pointed out after Proposition 2, for each t and each n, there is a pooling conciliatory equilibrium with demands $x^n(t), y^n(t) \in U(t)$ such that $x^n(t) = \hat{x}^{bs|y^n}(t)$ and $y^n(t) = \hat{y}^{bs|x^n}(t)$. A standard argument, as in Binmore et al. (1986), implies that the associated limit outcome $c = \lim \frac{x^n + y^n}{2}$ corresponds to the ex-post Nash solution.

Efficiency is a property to be desired in bargaining, at least whenever it is achievable. Let $C^*(\mathcal{B})$ be the set of contingent contracts in $C(\mathcal{B})$ that are also interim efficient. In this section, we first establish that efficiency is indeed achievable: $C^*(\mathcal{B})$ is always nonempty. Next, quite remarkably, we will show that under mild assumptions on $\mathcal{B}$, all elements of $C^*(\mathcal{B})$ are Myerson solutions. Finally, we show that interim-efficiency must occur at the limit when conciliatory equilibria are sequential.

4.1 Interim Efficiency is Achievable

The first result here establishes that interim-efficient limits always exist.

Proposition 3. *$C^*(\mathcal{B})$ is nonempty.*

To show this, we first apply a fixed-point theorem and Proposition 2 to establish that for any $\delta < 1$, there exists a pooling conciliatory equilibrium with each player proposing an interim-efficient contract; see Proposition 7 in the Appendix. By compactness, there are convergent sequences $\delta^n \to 1$, $x^n \to x$ and $y^n \to y$ such that the offers (x^n, y^n) are both interim efficient and comprise a pooling conciliatory equilibrium. To conclude the proof of Proposition 3, we show that the outcome $(x + y)/2$ is interim efficient.

Though perhaps intuitive, the result is not straightforward. First, interim efficiency is usually not preserved when averaging. We prove, however, that the sequences $(x^n)_{n\geq 1}$ and $(y^n)_{n\geq 1}$ must converge to each other in the space of interim utilities. Second, proving that the limit of interim-efficient contingent contracts is itself interim efficient requires some effort. It is natural to proceed by contraposition. If some contingent contract gives *strictly* higher interim utility to *all* types of both agents, then clearly this contract will also be interim superior to some contracts in the sequence before the limit. The issue is that interim inefficiency of the limit only guarantees the existence of some contingent contract giving at least as much interim utility to all types of all agents, and strictly more to at least one type of one agent. Weak inequality need not be preserved before the limit, and a subtler argument is needed to derive that a contract along the sequence is itself interim inefficient.

4.2 Convergence to Myerson

We next show that for smooth bargaining problems, any ex-post strictly individually rational outcome in $C^*(\mathcal{B})$ is a Myerson solution. This is a remarkably strong result. It rules out equilibria which always converge to other interim-efficient bargaining solutions, such as Harsanyi-Selten's. We also provide a rather mild boundary condition on $\mathcal{B}$ which guarantees that all elements of $C^*(\mathcal{B})$ are ex-post strictly individually rational. A contingent contract c is *ex-post strictly individually rational* if $c_i(t) > 0$ for all t and each agent i.

In general, we may describe i's maximal utility in state t given j's payoff v_j through the function $f_i(t,.) : [0, \overline{u}_j(t)] \to \mathbb{R}$ defined by:

$$f_i(t, v_j) = \max\{u_i : (u_i, v_j) \in U(t)\}.$$

Since $U(t)$ is convex, $f_i(t, v_j)$ is strictly decreasing on the interval $[\underline{u}_j(t), \overline{u}_j(t)]$. The notion of a smooth bargaining problem, introduced at the end of Section 2.1, is equivalent to requiring that $f_i(t,.)$ is continuously differentiable on $(0, \overline{u}_j(t))$, for all t and $i = 1, 2$ with $j \neq i$. For such problems, let $f_i'(t,.)$ be the continuous extension of this derivative over $[0, \overline{u}_j(t)]$. Part of the result below shows that elements of $C^*(\mathcal{B})$ must be ex-post strictly individually rational when the right derivative at zero is not 'too negative', or more precisely,

$$f_i'(t, 0) > -\frac{p(t)\overline{u}_i(t)}{\sum_{t'\in T(t_j)\setminus\{t\}} p(t')\overline{u}_j(t')}, \tag{BC}$$

for all $i = 1, 2$, $j \neq i$, and $t \in T$. This boundary condition (hence 'BC') means that, in each state, starting from a utility pair where j gets nothing, j's payoff can be increased without decreasing i's utility by much. Observe that (BC) is automatically satisfied whenever $\underline{u}_i(t) > 0$ for all i and $t \in T$. The rationale for the specific bound on the RHS, which is thus relevant only when $\underline{u}_i(t) = 0$ (and thus $f_i'(t, 0) \leq 0$) for some i and t, will become clear shortly.

Proposition 4 (Convergence to Myerson)**.** *Let $\mathcal{B}$ be a smooth bargaining problem, and let c be a contingent contract in $C^*(\mathcal{B})$. We have:*

(a) If c is ex-post strictly individually rational, then it is a Myerson solution.

(b) If $\mathcal{B}$ satisfies (BC), then c is ex-post strictly individually rational.

Hence, $C^*(\mathcal{B}) \subseteq MY(\mathcal{B})$ for all smooth $\mathcal{B}$ satisfying (BC).

We prove the first part of the proposition by deriving an appropriate approximation of each agent's best-safe payoff. To simplify the sketch of proof, suppose each x^n is an ex-post strictly individually rational and ex-post efficient demand. Then smoothness ensures a unique and strictly positive unit vector $\lambda^n(t)$ orthogonal to $U(t)$ at $x^n(t)$. As depicted in Figure 2, we can thus approximate Agent 2's best-safe payoff through his best-safe payoff from the linearized utility-possibility frontier $V_{\lambda^n}(t)$. This approximation,

$$\tilde{y}_2^{bs|x^n}(t) = x_2^n(t) - \frac{\lambda_1^n(t)}{\lambda_2^n(t)}\gamma^n x_1^n(t),$$

is at most $O((1-\gamma^n)^2)$ from $y_2^{bs|x^n}(t)$ by a second-order Taylor expansion. Combining this with our equilibrium conditions from Proposition 2 allows us to show that in the limit, Agent 2's expected payoff must be at least half of the linearized surplus in V_{λ^x}: that is, $E[c_2|t_2] \geq \frac{1}{2}E[\frac{\lambda^x \cdot x}{\lambda_2^x}|t_2]$, where $\lambda^n \to \lambda^x$. Similarly Agent 1 must get at least half of the linearized surplus in V_{λ^y} in expectation. Since c is interim efficient, it must also be ex-post efficient, and so $\lambda^x = \lambda^y$ (and $\lambda^x \cdot x = \lambda^x \cdot c = \lambda^x \cdot y$). The assumption that c is strictly ex-post individually rational ensures that λ^x is the unique ex-post orthogonal vector to $U(t)$ at $c(t)$. By Lemma 1, therefore, there is some $\hat{\lambda} \in \mathbb{R}_+^{T_1} \times \mathbb{R}_+^{T_2}$ which is interim orthogonal to $\mathcal{U}(\mathcal{B})$ at x such that $\lambda_i^x(t) = \frac{\hat{\lambda}_i(t_i)}{p(t_i)}$ for $i = 1, 2$. This is enough to show that each agent gets exactly half of the linearized surplus, making it a Myerson solution.

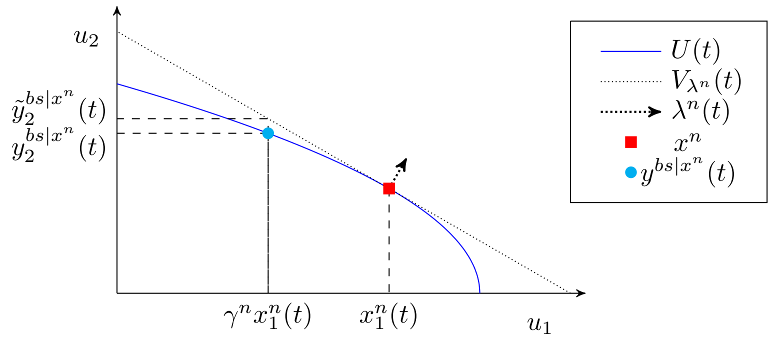

Figure 2: Approximation of Agent 2's best safe $y_2^b s|x^n(t)$ by $\tilde{y}_2^{bs|x^n}(t)$

We can now also explain how we prove part (b), which will also provide some intuition for (BC). Consider a feasible utility set with $\underline{u}_2(t) = 0$, which is the only case where condition (BC) is possibly binding, and suppose $c_2(t) = (\overline{u}_1(t), 0)$ so that $x_2^n(t) \to 0$. The reasoning from the previous paragraph remains valid, in that Agent 2's expected payoff given t_2 must be at least half of the expected linearized surplus $E[c_2|t_2] \geq \frac{1}{2}E[\frac{\lambda^x \cdot c}{\lambda_2^x}|t_2]$. This inequality is hard to satisfy if $\lambda_2^x(t)/\lambda_1^x(t)$ is very small, as then $\lambda^x(t) \cdot c(t)/\lambda_2^x(t)$ is very large (infinity if $\lambda_2^x(t)/\lambda_1^x(t) = 0$). We don't know much about $\lambda^x(t')$ for $t' \in T(t_2) \setminus \{t\}$ but it is certainly true that $c_2(t') \leq \lambda^x(t') \cdot c(t')/\lambda_2^x(t')$. Using this fact and $c_2(t') \leq \overline{u}_2(t')$, we see that it is infeasible for $c(t) = (\overline{u}_1(t), 0)$ and $E[c_2|t_2] \geq \frac{1}{2}E[\frac{\lambda^x \cdot c}{\lambda_2^x}|t_2]$ if ever $\frac{1}{2}\sum_{t' \in T(t_2)\setminus\{t\}} p(t'|t_2)\overline{u}_2(t') < \frac{1}{2}p(t|t_2)\frac{\lambda^x(t)\cdot c(t)}{\lambda_2^x(t)}$. But this condition exactly corresponds to (BC) for $i = 1$ after noticing that the ratio on the RHS is equal to $-\overline{u}_1(t)/2f_1'(0,t)$.

It should be emphasized that while smoothness and (BC) are sufficient to ensure the existence of a Myerson limit, they are far from necessary. However, there exist simple counter-examples to the result when relaxing any one of these assumptions.[11] Intuitively, when there is convergence to an interim-efficient limit outcome that is ex-post strictly individually rational and at a smooth point of the feasible utility set, the possibility of local deviations imposes considerable discipline on the relation between agents' demands in different states. This discipline (captured by the unique relationship between the interim and ex-post orthogonal vectors) is what ensures those limits must be a Myerson solution, and is lacking at a kink or boundary point.

[11] Available from the authors on request.

4.3 Sequential Equilibria

Under the notion of PBE studied thus far, beliefs following a demand/offer which is off-the-equilibrium path are left unrestricted. When more structure is desired, Kreps and Wilson (1982)'s sequential equilibrium comes to mind. However, the concept is defined for finite games, and generalizations to infinite games remain an active field of research (e.g., Myerson and Reny (2019)). In Section 4.3.1, we explain how the notion of sequential equilibrium naturally extends to our infinite bargaining game. We prove in Section 4.3.2 that, under rather mild regularity conditions, the limit of conciliatory sequential-equilibrium outcomes must be interim efficient, and hence Myerson solutions by Proposition 4. We cannot rely on general results to guarantee existence of sequential equilibria in our infinite game (identifying sequential equilibria is typically challenging even in finite games). Even so, we prove existence for a large class of problems where bargainers have two types each (see Section 4.3.3). Proving existence more generally remains an open question.[12]

4.3.1 Definition

Under the notion of sequential equilibrium, which is defined for finite games, beliefs should be justified at all information sets as limits of Bayesian-updated beliefs along some sequence of totally-mixed strategies which approximate the equilibrium strategies. Of course, it is impossible for a single strategy to mix between each of a continuum of offers with positive probability. To deal with this issue, consider any *finite* subset $\hat{U}$ of U, and define the $\hat{U}$-discretization of our game as its variant where demands/offers are restricted to $\hat{U}$. The discretization is *meaningful* given a conciliatory equilibrium if $\hat{U}$ contains equilibrium demands/offers. Fix now a belief system, which specifies i's belief about j's type after each demand/offer j may make in the original game. One can naturally restrict the equilibrium strategies and belief system to any meaningful discretization, simply by ignoring agents' bargaining stands and beliefs after infeasible demands/offers. With a slight abuse of terminology, we will not repeat this obvious step, and instead use the conciliatory equilibrium and the belief system of the original game as if they were defined in the discretizations.

[12]If sequential equilibria fail to exist in some cases (we haven't encountered an example yet), then the notion of PBE seems most appropriate as a fallback. Some may also simply generally prefer the notion of PBE over that of sequential equilibrium. In all these cases, our previous results apply and offer strong non-cooperative support for the Myerson solution.

Definition 1. *A conciliatory equilibrium, specifying strategies and a belief system, is a* sequential equilibrium *if it forms a sequential equilibrium in all meaningful discretizations.*

To provide further intuition regarding the restrictions imposed by sequential equilibrium, we suggest an equivalent definition based on Fudenberg and Tirole (1991)'s 'no-signaling-what-you-don't-know' principle (which was developed, once again, for finite games). Their idea is that while an opponent's unexpected demand/offer may reveal information to an agent about that opponent's type (no restriction is made in that regard), the agent's own demand/offer and own type provide no additional information."

If players' types are independent, the above idea is easy to formalize: i can hold any belief about j's type after j makes an unexpected demand/offer, but this belief cannot vary with i's type, or i's demand. In fact, however, the assumption of independent types is without loss of generality: any Bayesian game with state-dependent utility is strategically equivalent to a Bayesian game with independent types (Myerson, 1985). For instance, we could redefine bargaining problems to get a uniform prior by transforming each contract x into $\tilde{x}_i(t) = |T_{-i}|p(t_{-i}|t_i)x_i(t)$. Conciliatory PBEs and sequential equilibria of the resulting non-cooperative game are derived by applying the same transformation to equilibria of the original game.[13]

Definition 2. *If necessary, first reformulate the game so that types are independent. Assuming that types are independent, a belief system respects the* no-signaling-what-you-don't-know principle *if Agent i's belief about j's type after j made any off-equilibrium-path demand/offer does not vary with i's type, or i's demand/offer.*

Beliefs in sequential equilibria clearly satisfy this principle, since Bayesian updated beliefs associated to approximating strategies for the deviating agent can reveal information only about his type. Fudenberg and Tirole (1991) observe that, for finite two-stage games, imposing the no-signaling-what-you-don't-know principle is equivalent to restricting attention to sequential equilibria. Their result clearly extends to infinite games under the above definitions; the easy proof is left to the reader.

Observation 1. *A conciliatory PBE is a sequential equilibrium if, and only if, the belief system satisfies the no-signaling-what-you-don't-know principle.*

[13]The result for sequential equilibria was proved in Fudenberg and Tirole (1991, Proposition 5.1). Of course, they restricted attention to finite games, but their result carries over at once to infinite games under Definition 1.

4.3.2 Interim Efficiency at the Limit of Sequential Equilibria

We start by providing some intuition for our result (that limits must be efficient). For this, consider two-type bargaining problems ($T_1 = \{t_1', t_1''\}$, $T_2 = \{t_2', t_2''\}$) with a uniform prior p over types. We begin by focusing on a particular, seemingly robust type of conciliatory equilibrium: the ex-post PBE. With δ close to 1, ex-post equilibrium demands (x, y) are close to the limit ex-post Nash solution (epN for short), and satisfy $x = x^{bs|y}$, $y = y^{bs|x}$. We now explain why the ex-post PBE cannot be sequential if there is a contingent contract $e^* \in U$ that is strictly interim superior to epN.

We can assume, without loss of generality, that Agent 1 strictly prefers e^* over epN when types match, while Agent 2 strictly prefers e^* over epN when types mismatch: $E[t] > epN_1(t)$ for $t = (t_1', t_2')$ or (t_1'', t_2''), and $E[t] > epN_2(t)$ for $t = (t_1', t_2'')$ or (t_1'', t_2').[14] Now, if both of Agent 1's types were conciliatory at the information set (x, e^*), then Agent 2 could profitably deviate by proposing the interim superior e^* instead of y (getting $E[\frac{x_2+e_2^*}{2}|t_2] > E[\frac{x_2+y_2}{2}|t_2]$ by being conciliatory himself). Instead, suppose only one of Agent 1's types, say t_1', is conciliatory at (x, e^*), while type t_1'' insists on x. We know that Agent 2 strictly prefers e^* over epN in state (t_1', t_2''). By being conciliatory, type t_2'' profits by proposing e^* instead of y. Indeed, the deviation's expected payoff is

$$\frac{1}{2}\frac{x_2(t_1', t_2'') + e_2^*(t_1', t_2'')}{2} + \frac{\delta}{2}x_2(t_1'', t_2'') \geq \delta E[x_2|t_2] + \frac{e_2^*(t_1', t_2'') - x_2(t_1', t_2'')}{4},$$

which is strictly greater than the equilibrium payoff $E[\frac{x_2+y_2}{2}|t_2]$, as δ is close to 1 and both x and y are close to each other, and close to epN.

The arguments above imply that both of Agent 1's types must react aggressively at (x, e^*) to deter 2's deviation. Notice that, if type t_1' is conciliatory given some belief after e^*, then he will also be conciliatory for any larger probability of t_2' (e^* gives him more than $epN_1 \approx \gamma x_1$ in state (t_1', t_2')). Also, if type t_1' maintained his prior belief (that he faces type t_2' with probability half), then he would certainly be

[14] This follows at once after proving that each type-agent prefers e^* over epN for one type of the opponent, and vice versa for the opponent's other type. To see this, suppose 1 of type t_1' strictly prefers e^* over epN whatever is opponent's type. Since epN is ex-post efficient, Agent 2 strictly prefers epN over e^* in those states. Since e^* is interim strictly superior to epN, Agent 2 must strictly prefer e^* over epN in both (t_1'', t_2') and (t_1'', t_2'). But then e^* is interim strictly inferior to epN for Agent 1 of type t_1'', a contradiction.

conciliatory, because e^* delivers a higher expected payoff than epN. Thus, to trigger an aggressive stand, Agent 1 of type t_1' must believe that t_2'' is strictly more likely than t_2' at the information set (x, e^*). Similarly for Agent 1 of type t_1'' to be aggressive, he must believe that t_2' is strictly more likely than t_2'' at (x, e^*). Thus these two types *must* hold different beliefs to deter 2's deviation to e^*. This is not permitted in a sequential equilibrium.

We see that the ex-post PBE cannot be supported by a sequential equilibrium for δ close to 1 when the ex-post Nash solution is interim inefficient. The logic applies more generally, to all sequences of conciliatory equilibria (pooling and separating) with inefficient limits, independently of the prior when each agent has two types. With more than two types our argument extends assuming an additional regularity condition. The condition strengthens (BC) slightly, requiring that at the margin, j's utility can be increased in any state t without decreasing i's utility:

$$f_i'(t, 0) \geq 0. \tag{SBC}$$

This inequality holds (even strictly) whenever $\underline{u}_i(t) > 0$. If $\underline{u}_i(t) = 0$, then the inequality requires the orthogonal vector to $U(t)$ at $(0, \overline{u}_j(t))$ to place a zero weight on i. In other words, the Pareto frontier of $U(t)$ must be flat (if $i = 2$) or vertical (if $i = 1$) at the margin. This would be guaranteed, for instance, if utility possibility sets arise from utility functions satisfying Inada's conditions.

Remember that $C(\mathcal{B})$ denotes contingent contracts that can be approximated by a sequence of conciliatory-equilibrium outcomes as δ tends to 1, while $C^*(\mathcal{B})$ is the subset of contracts in $C(\mathcal{B})$ that are interim efficient. We let $C^s(\mathcal{B})$ be the subset of $C(\mathcal{B})$ for which equilibria along the sequence are sequential. We next establish $C^s(\mathcal{B}) \subseteq C^*(\mathcal{B})$ under mild conditions.

Proposition 5. *Let $\mathcal{B}$ be a bargaining problem that either has $|T_i| = 2$ for $i = 1, 2$, or is smooth and satisfies (SBC).*[15] *Then $C^s(\mathcal{B}) \subseteq C^*(\mathcal{B})$.*

[15] The result also holds if one agent has no private information ($|T_i| = 1$ for some i). Ex-post efficiency and interim efficiency are equivalent in that case, and Lemma 6 shows $C(\mathcal{B})$ contains only ex-post efficient contracts. So $C(\mathcal{B}) \subseteq C^*(\mathcal{B})$, and a fortiori $C^s(\mathcal{B}) \subseteq C^*(\mathcal{B})$.

4.3.3 Existence

Proving general existence results for adaptations of sequential equilibrium to infinite games is notoriously hard. We now establish that assumptions underlying our convergence result in Proposition 4 are also essentially sufficient for the nonemptiness of $C^s(\mathcal{B})$ when agents have two types each. The only additional assumption is that $f_i(t,\cdot)$ is twice-differentiable at $\overline{u}_{-i}(t)$, and $f_i''(t,\overline{u}_{-i}(t)) < 0$, for all i and $t \in T$. It is thus required that the utility possibility set is *strictly* convex at utility pairs where one agent gets his best possible payoff. Finding sufficient conditions for the non-emptiness of $C^s(\mathcal{B})$ with larger type spaces remains an open question (we have not found a counter-example).

Proposition 6. *Suppose that $\mathcal{B}$ is smooth, (BC) holds, $f_i''(t,\overline{u}_{-i}(t)) < 0$, and $|T_i| = 2$, for all $i = 1,2$ and $t \in T$. Then $C^s(\mathcal{B}) \neq \emptyset$.*[16]

We now briefly sketch the main ideas in the proof, which appears in the Online Appendix. Suppose first that utility is transferable and types are independent: risk neutral bargainers divide $M(t)$ dollars in each state t. The set of conciliatory equilibria is easy to describe in this case. Namely, each bargainer demands a fraction $\frac{1}{1+\gamma}$ of the expected money available (e.g., $E[x_1|t_1] = \frac{E[M|t_1]}{1+\gamma}$), while offering his opponent an expected share of $\frac{\gamma}{1+\gamma}$ (e.g., $E[x_2|t_2] = \frac{\gamma E[M|t_2]}{1+\gamma}$). We show that each conciliatory equilibrium outcome can be supported by a sequential equilibrium in this case. Suppose Agent 1 unilaterally deviates by demanding $\hat{x}$ instead of the equilibrium x. One must define a belief for Agent 2 that is independent of t_2 and an equilibrium in the bargaining-posture stage that makes both types of Agent 1 no better-off compared to his equilibrium payoff. Getting both types of Agent 2 to take an aggressive stand against $\hat{x}$ and both types of Agent 1 to take a conciliatory stand would do it, but there won't be beliefs supporting this when $\hat{x}$ is generous to Agent 2 (compared to δy). When no such beliefs exist, Agent 2 will take a conciliatory stand for some type, but there is a sense in which Agent 1 is too generous towards Agent 2 in such deviations, and one can find some equilibrium of the continuation game that leaves Agent 1 no better off. The argument here relies on Farkas' lemma.

For general bargaining problems, we introduce the idea of a joint principal-agent equilibrium. Essentially, it is a pair (x,y) of contingent contracts such x (resp., y) is

[16] In fact, we prove a slightly stronger result: there exists a threshold $\underline{\delta} < 1$ such that conciliatory sequential equilibria exist for all $\delta \in [\underline{\delta}, 1]$.

the analogue of Myerson (1983)'s principal-agent solution when Agent 2's (resp., 1's) outside option is γy (resp., γx). The reasoning from the paragraph above extends to any joint principal-agent equilibrium. The last step is to show the existence of a joint principal-agent equilibrium, for which we use the facts that $\mathcal{B}$ is smooth, satisfies (BC) and has $f_i''(t, \overline{u}_{-i}(t)) < 0$. As should be clear from above, these conditions are not necessary for existence.

5 Extensions

This section considers various extensions to our original model. Primary among these is the extension of our results to an infinite horizon war of attrition game. We also extend to non-common priors and asymmetry in discounting. Details are provided in the Online Appendix.

5.1 Concession as a War of Attrition

Our results extend to a dynamic bargaining game, where the concession stage is a war of attrition. At period 0, agents independently propose contingent contracts, as in the demand/offer stage of our static game. Subsequently, and as long as no agent has conceded, an intermediary reaches out to one agent in each period $s \in \{1, 2, ...\}$, to inquire whether he'd like to concede. Future payoffs are discounted using a common discount factor δ. Which of the two agents the intermediary contacts first is determined by uniform randomization. The intermediary alternates thereafter (contacting agent i in all odd periods and agent j in all even periods). In any given period, there is an exogenous probability $\varepsilon \in (0, 1)$ that the intermediary and the designated agent do not get in touch (e.g., the agent is unavailable, the intermediary gets sidetracked, or a technical issue arises).[17]

A *stationary equilibrium* is a perfect-Bayesian equilibrium where each type of each agent decides whether to concede in period s solely based on the demands/offers and his beliefs about his opponent in period s (but not explicitly on the time-period s).[18] Strategies induce *initial concession* if both agent 1 and 2 concede in period 1

[17]Hence, following any initial demands/offers, all future periods are on the equilibrium path, which means that agent's beliefs can be determined by Bayes' rule from his beliefs after observing an opponent's initial demand and that opponent's strategy.

[18]Strategies can still be time dependent, because beliefs may vary over time. Notice by the way

conditional on being called by the intermediary.

In the Online Appendix, we show that the set of payoffs in stationary equilibria with initial concession is equivalent to that of conciliatory equilibria in our simple two-stage game, when $\gamma = \delta(1-\varepsilon)/(1-\varepsilon\delta^2)$. And so, *any interim efficient limits of these equilibria* (as $\delta \to 1$) *are Myerson solutions*, under our previous assumptions about the bargaining problem.

5.2 Non-common priors

So far we have assumed that agents share a common prior. We now relax this assumption, letting $p_i \in \Delta(T_1 \times T_2)$ denote Agent i's prior. Disagreement about priors give rise to the possibility of mutually-beneficial bets at the interim stage. Consider, for instance, problems such as the introductory example where there is an amount $\$m(t)$ to split in state t. To isolate the effect of non-common priors, suppose both bargainers have the same utility function $u : \mathbb{R}_+ \to \mathbb{R}$ that is smooth, strictly concave and with an infinite marginal utility of money at zero. With these assumptions, the ex-post bargaining problems satisfy all our assumptions, including the strong boundary condition. Symmetry of preferences means the ex-post Nash solution–which will split $m(t)$ equally between both bargainers in all states–will be interim efficient, *for any common prior.*[19] By contrast, the ex-post Nash solution is *interim inefficient for all non-common prior environments.*[20]

The point we want to emphasize is this: *all our main results extend to situations where bargainers derive their beliefs from different priors.* Indeed, Agent i's expected utility from x under p_i is identical to his expected utility, *under a uniform prior over the state space*, from receiving $\tilde{x}_i(t) = |T_{-i}|p_i(t_{-i}|t_i)x_i(t)$ in each state t. Thus, the bargaining problem U under the priors (p_1, p_2) is *strategically equivalent* to the bargaining problem $\tilde{U}$ under the uniform common prior, where $\tilde{U}(t)$ is the set of $\tilde{x}(t)$

that higher-order beliefs may vary and matter as well (e.g., beyond i's belief about t_i, i's assessment about what j believes regarding t_i may also matter, etc.).

[19]The vector $(1,1)$ is orthogonal to $U(t)$ at the Nash solution. For the prior p, take $\lambda_i(t_i)$ as the marginal $p(t_i)$ and apply Lemma 1.

[20]Following Morris (1994), for all non-common prior (p_1, p_2), there exists $\phi : T \to \mathbb{R}^2$ such that (a) $\phi_1(t) + \phi_2(t) = 0$ for all t, and (b) $E_i[\phi_i|t_i] > 0$ for all i, t_i. Consider a small (infinitesimal) monetary transfer $\$\frac{\phi_i(t)dm}{u'(0.5m(t))}$ in each state t between the two agents (budget balanced, by (a)). The marginal impact on i's ex-post utility in state t is $\phi(t)dm$. By (b), the new contingent contract gives strictly larger interim utility to all types of both agents.

for $x \in U$.[21]

5.3 Asymmetric Bargaining Power

Under complete information, asymmetric bargaining power can easily be accommodated in the Nash solution, by maximizing a weighted Nash product $u_1^{\alpha}u_2^{1-\alpha}$. The parameter $\alpha \in [0,1]$ captures 1's relative bargaining power.

A natural way to introduce asymmetry in our non-cooperative bargaining game is to change the outcome that prevails when both agents take a conciliatory stand, say $(1-\alpha)x + \alpha y$ instead of the plain average of the offers x and y. Under complete information, a standard argument shows that, for $\alpha \in (0,1)$, the Nash equilibrium outcome converges to the weighted Nash solution discussed above, as δ tends to 1. Notice that players do *less* well when their own proposal is agreed to with greater frequency. What are the limit equilibrium outcomes arising under incomplete information?

Assume that the bargaining problem is smooth, and that the stronger boundary condition (SBC) holds. Following the same reasoning as in the proof of Proposition 4, any limit equilibrium outcome c^* must be an *α-weighted Myerson solution*:[22] there exists $\hat{\lambda} \in \Delta_{++}(T_1) \times \Delta_{++}(T_2)$ such that

$$E[c_i^*|t_i] = \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) \frac{\alpha_i \max_{v \in V(t)} \sum_{j=1,2} \lambda_j(t_j) v_j}{\lambda_i(t_i)},$$

where $\alpha_1 = \alpha$, $\alpha_2 = 1-\alpha$, and $\lambda_j(t_j) = \hat{\lambda}_j(t_j)/p(t_j)$, for all t_j and both $j = 1, 2$. In terms of the three-step process proposed in Section 2.2 to describe the Myerson solution, only the second step is modified: a share α_i of the surplus in $V_\lambda(t)$ is now allocated to agent i.

Finally, remember that we considered some alternative bargaining protocols in Section 2.3. For the one introduced and analyzed by Evans (2003) under complete information, suppose now that there is a probability ε_i that i's demand/offer goes

[21] In other words, equilibrium outcomes must satisfy the analogue in our framework of Myerson (1984)'s probabilistic invariance axiom.

[22] The proof is available from the authors on request. The stronger boundary condition can be relaxed. What matters is that the limit equilibrium outcome is strictly individually rational. In Proposition 4, we showed that the weaker boundary condition (BC) guarantees this for $\alpha = 1/2$. That weaker condition can easily be adapted for any $\alpha \in (0,1)$.

astray. It is not difficult to check that, if both ε_1 and ε_2 vanish, then (under the usual assumptions) the limit equilibrium outcome will be the α-weighted Myerson solution where $\alpha = \frac{1}{1+\lim \frac{\varepsilon_1}{\varepsilon_2}}$. Having one's demand/offer go astray less often thus corresponds to a higher weight in the limit. In the bargaining game where acceptance is stochastically delayed, a natural asymmetry is differential discounting, so that agent i's discount rate is r_i. It is again easy to check that as agents become patient, the limit equilibrium outcome will be the α-weighted Myerson solution where $\alpha = \frac{1}{1+\lim \frac{r_1}{r_2}}$. Greater patience corresponds to a higher weight.

Proposition 5 (limits of sequential equilibria are efficient) goes through unchanged.[23]

References

Ausubel, L., P. Cramton, and R. Deneckere (2002). Bargaining with incomplete information. *Handbook of Game Theory 3*, 1897–1945.

Bazerman, M. H. and J. J. Gillespie (1999). Betting on the future: The virtues of contingent contracts. *Harvard Business Review 77*, 155–162.

Binmore, K., M. J. Osborne, and A. Rubinstein (1992). Chapter 7 noncooperative models of bargaining. In *Handbook of Game Theory with Economic Applications*, Volume 1 of *Handbook of Game Theory with Economic Applications*, pp. 179 – 225. Elsevier.

Binmore, K., A. Rubinstein, and A. Wolinsky (1986). The nash bargaining solution in economic modelling. *The RAND Journal of Economics 17*, 176–188.

de Clippel, G. and E. Minelli (2004). Two-person bargaining with verifiable information. *Journal of Mathematical Economics 40*, 799–813.

Evans, R. (2003). A one-period version of rubinstein's bargaining game. *Contributions to Theoretical Economics 3*.

Fudenberg, D. and J. Tirole (1991). Perfect bayesian equilibrium and sequential equilibrium. *Journal of Economic Theory 53*, 236–260.

Harsanyi, J. C. and R. Selten (1972). A generalized nash solution for two-person bargaining games with incomplete information. *Management Science 18*(5-part-2), 80–106.

Holmström, B. and R. B. Myerson (1983). Efficient and durable decision rules with incomplete information. *Econometrica*, 1799–1819.

Kennan, J. and R. Wilson (1993). Bargaining with private information. *Journal of Economic Literature 31*(1), 45–104.

[23] We have not verified whether Proposition 6 extends.

Kim, J. Y. (2017). A noncooperative foundation of the neutral bargaining solution. *Unpublished manuscript*.

Kreps, D. M. and R. Wilson (1982). Sequential equilibria. *Econometrica: Journal of the Econometric Society*, 863–894.

Maskin, E. and J. Tirole (1992). The principal-agent relationship with an informed principal, ii: Common values. *Econometrica 60*, 1–42.

Morris, S. (1994). Trade with heterogeneous prior beliefs and asymmetric information. *Econometrica 62*, 1327–1347.

Myerson, R. (1983). Mechanism design by an informed principal. *Econometrica 51*(6), 1767–1797.

Myerson, R. B. (1984). Two-person bargaining problems with incomplete information. *Econometrica 52*(2), 461–487.

Myerson, R. B. (1989). Credible negotiation statements and coherent plans. *Journal of Economic Theory 48*(1), 264–303.

Myerson, R. B. and P. Reny (2019). Perfect conditional ε-equilibria of multi-stage games with infinite sets of signals and actions. *Working Paper*.

Okada, A. (2016). A non-cooperative bargaining theory with incomplete information: Verifiable types. *Journal of Economic Theory 163*, 318–341.

Osborne, M. J. and A. Rubinstein (1990). *Bargaining and Markets*. Academic Press.

Wilson, R. (1978). Information, efficiency, and the core of an economy. *Econometrica 46*(4), 807–816.

Appendix

A1 Preliminaries

We now state, and prove in the Online Appendix, some useful technical results. We begin with the correspondence $F : U \rightrightarrows U$ that associates to any contingent contract $v \in U$ the set of contingent contracts $u \in U$ that are weakly interim superior to v:

$$F(v) = \{u \in U : E[u_i|t_i] \geq E[v_i|t_i] \text{ for all } t_i \in T_i \text{ and } i = 1, 2\}.$$

Lemma 2. *F is continuous with non-empty, compact, and convex values.*

The next lemma establishes that the set of interim efficient contingent contracts is closed. This is true under complete information when there are two agents, but not for three or more agents. With two agents under incomplete informations, there are more than two type-agents and it is not clear a priori that interim efficiency is preserved through limits.

Lemma 3. *Consider a sequence of feasible contingent contracts* $x^n \to x \in U$*. If each* x^n *is interim efficient, then* x *is interim efficient.*

Next, say that contract z *interim strictly dominates* x if $E[z_i|t_i] > E[x_i|t_i]$ for all i, t_i, and x is *weakly interim efficient* if there is no such contract z.

Lemma 4. *Suppose* $|T_i| = 2$ *for* $i = 1, 2$*. If* x *is both ex-post efficient and weakly interim efficient, then it is also interim efficient.*

Contract e^* *interim dominates* x when restricted to $T_1' \times T_2'$ if $E[e_i^*|t_i, T_{-i}'] \geq E[x_i|t_i, T_{-i}']$ for all $i, t_i \in T_i'$, with strict inequality for some i, t_i.

Lemma 5. *Consider a smooth bargaining problem where each agent has at least two types. Suppose* x *is an ex-post efficient contract with* $x_i(t) > \underline{u}_i(t)$ *for* $i = 1, 2$ *and* $t \in T$*. If* x *is not interim efficient, then there are* $T_i' \subset T_i$ *for* $i = 1, 2$ *with* $|T_i'| = 2$ *and a contract* e^* *that interim dominates* x *when restricted to* $T_1' \times T_2'$*.*

A2 Characterization of Conciliatory Equilibria

Proof of Proposition 1 (Inscrutability) Take a separating conciliatory equilibrium. It is associated with partitions of the type spaces T_1 and T_2:

$$T_1 = T_1^{(1)} \cup \cdots \cup T_1^{(m)}, \text{ and } T_2 = T_2^{(1)} \cup \cdots \cup T_2^{(n)}.$$

All types t_1 belonging to cell $T_1^{(j)}$ propose $x^{(j)}$, and all types t_2 belonging to cell $T_2^{(j)}$ propose $y^{(j)}$. We may assume wlog that $x^{(j)} \neq x^{(k)}$ and $y^{(j)} \neq y^{(k)}$ when $j \neq k$. We can thus define functions $j : T_1 \to \{1, \ldots, m\}$ and $k : T_2 \to \{1, \ldots, n\}$, where $j(t_1)$ is the index of the cell in the partition of T_1 to which t_1 belongs ($t_1 \in T_1^{(j(t_1))}$), and $k(t_2)$ is the index of the cell in the partition of T_2 to which t_2 belongs ($t_2 \in T_1^{(k(t_2))}$). Define best-safe contracts, $x^{bs|y^{(j)}}$ for each $j \in \{1, \ldots, n\}$ and $y^{bs|x^{(k)}}$ for each $k \in \{1, \ldots, m\}$.

Consider a pooling strategy for Agent 1 where he offers x^* independently of t_1, with $x^*(t) = x^{(j(t_1))}(t)$ for $t = (t_1, t_2)$, and a pooling strategy for Agent 2 where he offers y^* independently of t_2, with $y^*(t) = y^{(k(t_2))}(t)$ for $t = (t_1, t_2)$. Followed by a conciliatory posture from all types, these strategies yield the same outcome in all states as the original separating conciliatory equilibrium. To conclude the proof we show these new strategies are part of a conciliatory equilibrium, by verifying the conditions of Proposition 2.

The desired condition $E[x_1^*|t_1] \geq E[x_1^{bs|y^*}|t_1]$ follows by observing that in the original separating equilibrium, if Agent 1 (of any type) were to deviate and propose $x^{bs|y^*}$ then all types of Agent 2 will take a conciliatory posture, for whatever beliefs 2 may have following this deviation. This follows by a similar computation as in the proof of Proposition 2. In the original separating equilibrium, Agent 1 of a type $t_1 \in T_1^{(j)}$ instead proposes $x^{(j)}$, to which all types of Agent 2 respond with a conciliatory posture. The rationality of Agent 1 sending $x^{(j)}$ thus requires that $E[x_1^{(j)}|t_1] \geq E[x_1^{bs|y^*}|t_1]$. By construction of x^*, when $t_1 \in T_1^{(j)}$ we have $E[x_1^{(j)}|t_1] = E[x_1^*|t_1]$, yielding the desired inequality. A symmetric argument for Agent 2 implies the condition $E[y_2^*|t_2] \geq E[y_2^{bs|x^*}|t_2]$.

To conclude the proof, we show the condition $E[x_2^*|t_2] \geq E[x_2^{bs|y^*}|t_2]$ holds for all t_2; the condition that $E[y_1^*|t_1] \geq E[y_1^{bs|x^*}|t_1]$ for all t_1 is derived analogously. Observe that after receiving the proposal $x^{(j)}$ in the separating equilibrium, an Agent 2 of type $t_2 \in T_2^{(k)}$ has Bayesian-updated beliefs given by $p(t_1|t_2, T_1^{(j)})$. Agent 2 is conciliatory following Agent 1's proposal when he also has the option of posturing aggressively. By a similar computation as for Proposition 2, we conclude that being conciliatory requires

$$E[x_2^{(j)}|t_2, T_1^{(j)}] \geq \gamma E[y^{(k)_2}|t_2, T_1^{(j)}] = E[x_2^{bs|y^{(k)}}|t_2, T_1^{(j)}] \quad (4)$$

for every type $t_2 \in T_2^{(k)}$, every $k \in \{1, \ldots, n\}$ and every $j \in \{1, \ldots, m\}$. Multiply the inequality (4) associated with each $j \in \{1, \ldots, m\}$ by the probability $p(T_1^{(j)}|t_2)$ and sum up the corresponding inequalities over all j. The resulting inequality is equivalent to the desired one by the construction of x^* and y^*. □

Proof of Proposition 2 (Characterization of Conciliatory Equilibria) Necessity was established in the text. For sufficiency, suppose the contingent contracts x and y satisfy (3). We construct a conciliatory equilibrium in which all types of Agent 1 propose x and all types of Agent 2 propose y. Following the offer x, Agent 2's updated belief over Agent 1's type coincides with his interim belief, and being conciliatory is a best response since $E[x_2|t_2] \geq E[x_2^{bs|y}|t_2]$, for all $t_2 \in T_2$. Similar reasoning applies to Agent 1 following y.

We now define beliefs and strategies, and check incentives after a unilateral deviation. Without loss, suppose Agent 1 proposes x' instead, while 2 proposes y. For any type t_2, define Agent 2's beliefs and action as follows. Let $T_1(t_2, x', y) = \{t_1 \in T_1 : x'_2(t_1, t_2) < \gamma y_2(t_1, t_2)\}$. If $T_1(t_2, x', y) \neq \emptyset$, let the probability type t_2 believes that he faces t_1 given x' be $\mu_2(t_1|t_2, x', y) = 1$ for some $t_1 \in T_1(t_2, x', y)$, so Agent 2 takes an aggressive stand against x'. If $T_1(t_2, x', y) = \emptyset$ then let $\mu_2(t_1|t_2, x', y) = 1$ for some arbitrary $t_1 \in T_1$, with Agent 2 conciliatory following x'. Agent 1's belief following y coincides with his interim belief, and he is conciliatory following 2's proposal.

We now show that the off-equilibrium behavior following a unilateral deviation is rational. If Agent 2 expects y to result in a conciliatory posture, then it is rational for him to posture aggressively against 1's deviation x' given his off-equilibrium belief when $T_1(t_2, x, y) \neq \emptyset$, and to be conciliatory otherwise. Moving on to Agent 1's strategy, posturing aggressively against y after proposing x', when he is of type t_1, gives him an expected payoff of

$$\delta \sum_{t_2 \in T_2(x', y)} p(t_2|t_1) x'_1(t_1, t_2),$$

where $T_2(x', y) = \{t_2 : T_1(t_2, x', y) = \emptyset\}$ is the set of Agent 2's types who will be conciliatory after x'. By being conciliatory, Agent 1 of type t_1 gets:

$$\sum_{t_2 \in T_2(x', y)} p(t_2|t_1) \frac{x'_1(t_1, t_2) + y_1(t_1, t_2)}{2} + \delta \sum_{t_2 \in T_2 \setminus T_2(x', y)} p(t_2|t_1) y_1(t_1, t_2).$$

Multiplying the payoffs by $\frac{2}{\gamma}$ and rearranging, we see that being conciliatory is preferable to being aggressive if and only if

$$\sum_{t_2 \in T_2(x', y)} p(t_2|t_1) x'_1(t_1, t_2) \leq \frac{1}{\gamma} E[y_1|t_1] + \sum_{t_2 \in T_2 \setminus T_2(x', y)} p(t_2|t_1) y_1(t_1, t_2).$$

Since $x'_2(t) \geq \gamma y_2(t)$ for $t = (t_1, t_2)$ such that $t_2 \in T_2(x', y)$ then we must have $x'_1(t) \leq x_1^{bs|y}(t)$. Imposing this inequality as an equality and rearranging, we get that a conciliatory posture is certainly preferable if

$$E[x_1^{bs|y}|t_1] \leq \frac{1}{\gamma} E[y_1|t_1] + \sum_{t_2 \in T_2 \setminus T_2(x', y)} p(t_2|t_1)(y_1(t_1, t_2) + x_1^{bs|y}(t_1, t_2)).$$

By equation (3), we have $E[y_1|t_1] \geq E[y_1^{bs|x}|t_1] = \gamma E[x_1|t_1] \geq \gamma E[x_1^{bs|y}|t_1]$. Hence, a conciliatory posture is preferable, since $y(t) \geq 0$ and $x^{bs|y}(t) \geq 0$.

We now show that deviating from x to x' is not profitable for Agent 1. Agent 1's expected payoff is equal to $\delta y_1(t)$ in any state $t = (t_1, t_2)$ where Agent 2 refuses x' (i.e. if $t_2 \in T_2 \setminus T_2(x', y)$), and is equal to $\frac{x_1'(t)+y_1(t)}{2}$ for states where 2 is conciliatory. Thus 1 has no strict incentive to deviate by proposing x' instead of x if and only if

$$\sum_{t_2 \in T_2(x',y)} p(t_2|t_1) \frac{x_1'(t_1,t_2) + y_1(t_1,t_2)}{2} + \delta \sum_{t_2 \in T_2 \setminus T_2(x',y)} p(t_2|t_1) y_1(t_1,t_2) \leq \frac{E[x_1|t_1] + E[y_1|t_1]}{2} \tag{5}$$

Multiplying both sides of the inequality by 2 and rearranging, we get:

$$\sum_{t_2 \in T_2(x',y)} p(t_2|t_1) x_1'(t_1,t_2) + \gamma \sum_{t_2 \in T_2 \setminus T_2(x',y)} p(t_2|t_1) y_1(t_1,t_2) \leq E[x_1|t_1]. \tag{6}$$

Notice that $x_1'(t_1, t_2) \leq x_1^{bs|y}(t_1, t_2)$ when $t_2 \in T_2(x', y)$, by definition of $T_2(x', y)$, and that $\gamma y_1(t) \leq x_1^{bs|y}(t)$, by definition of $x_1^{bs|y}$. Thus the LHS of equation (6) is less or equal to $E[x_1^{bs|y}|t_1]$, which itself is less than the RHS of equation (6), thanks to our equilibrium conditions from equation (3). Thus Agent 1 does not find it profitable to unilaterally deviate to x', as claimed.

It remains to ensure there exist mutually optimal continuation strategies given beliefs after mutual deviations to x' and y'. We define beliefs to be consistent with those after unilateral deviations, so $\mu_2(t_1|t_2, x', y') = \mu_2(t_1|t_2, x', y)$, and $\mu_1(t_2|t_1, x', y') = \mu(t_1|t_2, x, y')$. These beliefs and agents' posturing strategies determine expected continuation payoffs. Let those continuation payoffs correspond to payoff functions in an auxiliary posturing game with $T_1 \cup T_2$ players. That finite game must have a Nash equilibrium and so we let postures following deviations x', y' be defined by one of those equilibria. □

A3 Non-Emptiness of $C^*(\mathcal{B})$

We start by establishing, for any δ (before the limit), the existence of pooling conciliatory equillibria with interim efficient demands.

Proposition 7. *There exists some pooling conciliatory equilibrium with interim efficient demands.*

Proof. Let $\overline{U}(t) = \{u \in \mathbb{R}^2_+ | (\exists v \in U(t)) : u \leq v\}$, and let $\hat{\phi} : \overline{U} \rightrightarrows U$ be the correspondence defined by $\hat{\phi}(v) = \arg\max_{u \in F(v)} \prod_{t_i,\, i} (E[u_i|t_i] - E[v_i|t_i] + 1)$, where F was defined right before Lemma 2. The set $\hat{\phi}(v)$ is compact and convex, since it is obtained by maximizing a concave function over a set that is itself compact and convex. Clearly, it selects contingent contracts that are interim efficient in U. The Theorem of the Maximum then implies that $\hat{\phi}$ is upper hemi-continuous (F is continuous, thanks to Lemma 2).

Let then $\phi : U^2 \rightrightarrows U^2$ be the correspondence defined as follows: $\phi(x, y) = (\hat{\phi}(x^{bs|y}), \hat{\phi}(y^{bs|x}))$. This is well-defined since $x^{bs|y}$ and $y^{bs|x}$ belong to $\bar{U}$ (but not necessarily U). Notice $x^{bs|y}$ is continuous in y and that $y^{bs|x}$ is continuous in x. Let (x, y) be a fixed-point of ϕ, by Kakutani's fixed point theorem. The construction of ϕ ensures interim efficiency of x and y, and that $E[x_i|t_i] \geq E[x_i^{bs|y}|t_i]$ and $E[y_i|t_i] \geq E[y^{bs|x}|t_i]$ for all t_i and i. Hence, by Proposition 2, demands (x, y) are sustained by a pooling conciliatory equilibrium. □

The next lemma establishes that if players demands $(x^n, y^n) \to (x, y)$ as $\delta^n \to 1$, then x and y must give the same interim utilities

Lemma 6. *Consider a sequence of bargaining games with $\delta^n \to 1$ and an associated sequence of pooling conciliatory equilibria whose demands converge, $(x^n, y^n) \to (x, y)$. Then (i) $E[x_i|t_i] = E[y_i|t_i]$ for all t_i and i, and (ii) x and y are ex-post efficient.*

Proof. For (i), observe that in a conciliatory equilibrium, we must have $E[x_2^n|t_2] \geq E[x_2^{bs|y^n}|t_2] = \gamma^n E[y_2^n|t_2]$. In the limit as $\gamma^n \to 1$ we must have $E[x_2|t_2] \geq E[y_2|t_2]$. We must also have $E[y_2^n|t_2] \geq E[y_2^{bs|x^n}|t_2] \geq E[x_2^n|t_2]$. Hence $E[y_2|t_2] \geq E[x_2|t_2]$ and so $E[y_2|t_2] = E[x_2|t_2]$, and by identical logic $E[y_1|t_1] = E[x_1|t_1]$.

We now prove (ii). If $f_2(t, x_1(t')) > x_2(t')$ for some t', $E[y_2^n|t_2'] \geq E[y_2^{bs|x^n}|t_2']$ and $y_2^{bs|x^n}(t) \geq f_2(t, \gamma^n x_1^n(t)) \geq \gamma^n x_2^n(t)$, for all t, imply that $E[y_2|t_2'] \geq \lim E[y_2^{bs|x^n}|t_2'] \geq E[f_2(\cdot, x_1(\cdot))|t_2'] > E[x_2|t_2']$, a contradiction to (i). Suppose now x is not ex-post efficient. Given $f_2(t, x_1(t)) = x_2(t)$ for all t we must have $x_1(t') < \underline{u}_1(t')$ for some t'. Then $x_1^n(t') < \underline{u}_1(t')$ for large n, and so $y_2^{bs|x^n}(t') = \overline{u}_2(t') > f_2(t, x_1(t')) = x_2(t')$. This implies $E[y_2|t_2'] \geq \lim E[y_2^{bs|x^n}|t_2'] > E[f_2(\cdot, x_1(\cdot))|t_2'] \geq E[x_2|t_2']$, contradicting (i). □

Proof of Proposition 3 ($C^*(\mathcal{B})$ is non-empty) Fix a sequence $\delta^n \to 1$, and an associated sequence of pooling conciliatory equilibria with interim efficient demands (x^n, y^n) (see Proposition 7). Since $U(t)$ is compact, we may assume (considering a subsequence if needed) (x^n, y^n) converges to some limit (x, y) as n tends to infinity. By Lemma 3, x and y are interim efficient. By Lemma 6, $E[x_i|t_i] = E[y_i|t_i]$ for all i, t_i. So the limit equilibrium outcome $\frac{x+y}{2}$ is interim efficient and belongs to $C^*(\mathcal{B})$. □

A4 Convergence to Myerson

Proof of Proposition 4

Fix a smooth bargaining problem $\mathcal{B}$, and assume U is comprehensive, that is, $v \in \mathbb{R}^2_+$ belongs to $U(t)$ as soon as it contains some $u \geq v$. To see why this is without loss of generality, assume U is not comprehensive. Consider its comprehensive closure $\overline{U}$ defined at the beginning of the proof for Proposition 7. Notice $\underline{u}(t)$, $\overline{u}(t)$, and the set of interim efficient contingent contracts remain unchanged when considering $\overline{U}(t)$ instead of $U(t)$. Similarly, for all x, y in U, $x^{bs|y}(t)$ and $y^{bs|x}(t)$ remain unchanged. Hence the set of strictly individually rational contingent contracts that belong to $C^*(\mathcal{B})$ is unchanged when taking the comprehensive closure. The set of Myerson solutions also remains unchanged. Hence (a) in Proposition 4 holds if we can show it holds for $\overline{U}$, which is comprehensive. As for (b), taking comprehensive closures has no impact on whether (BC) holds. Based on the above observations, any conciliatory equilibrium outcome for U is also a conciliatory equilibrium outcome for $\overline{U}$. Hence (b) in Proposition 4 holds if we can show it holds for $\overline{U}$, which is comprehensive.

Next, we associate to any weakly efficient $u \in U(t)$ a unique positive unit vector $\lambda^u(t)$ that is orthogonal to $U(t)$ at u. This is unequivocally defined if u is strictly individually rational ($U(t)$ is smooth). What if $u_i = 0$ for some i? Then there could be multiple orthogonal unit vectors. Define $\lambda^u(t)$ by the continuous extension over strictly individually rational payoff pairs: $\lambda^u(t) = \lim_m \lambda^{u^m}(t)$ for any sequence of strictly individually rational and efficient payoff pair u^m that converges to u. This is well-defined since $U(t)$ is smooth.

Take an element c^* of $C^*(\mathcal{B})$. Let $\delta^n \to 1$ and an associated sequence of pooling conciliatory equilibria with equilibrium demands (x^n, y^n), such that $(x^n, y^n) \to (x, y)$ and $c^* = \frac{x+y}{2}$. By Lemma 6, x and y are both ex-post efficient. Our proof of Proposition 4 proceeds in two steps. First, Lemma 7 establishes that Agent 2 must get at least half of the linearize surpluses $V_{\lambda^x}(t)$ in expectation, while furthermore showing that this implies $x_2(t) > \underline{u}_2(t)$ and hence $c^*_2(t) > \underline{u}_2(t)$ if $\mathcal{B}$ satisfies (BC). Similarly, Agent 1 must get at least half of the linearized surpluses $V_{\lambda^y}(t)$ in expectation and

$y_1(t) > \underline{u}_1(t)$ and $c_1^*(t) > \underline{u}_1(t)$ given (BC). The second step is Lemma 8, which shows that, if c^* is strictly individually rational and interim efficient, then $\lambda^y = \lambda^x$ and each agent expects exactly half of the linearized surplus, so c^* must be a Myerson solution.

Lemma 7. *Let $M^x(t) = \lambda^x(t) \cdot x(t)$ and $M^y(t) = \lambda^y(t) \cdot y(t)$. Then $\lambda_2^x(t) > 0$ and $\lambda_1^y(t) > 0$ and:*

$$E[c_2^*|t_2] \geq \frac{1}{2}E[\frac{M^x}{\lambda_2^x}|t_2], \quad \textit{and} \quad E[c_1^*|t_1] \geq \frac{1}{2}E[\frac{M^y}{\lambda_1^y}|t_1]. \tag{7}$$

Furthermore, if in state t the bargaining problem satisfies (BC) for $i = 1$, then $x_2(t) > \underline{u}_2(t)$, and if it satisfies (BC) for $i = 2$, then $y_1(t) > \underline{u}_1(t)$, so that if it satisfies (BC) for $i = 1, 2$ then $c_i^*(t) \in (\underline{u}_i(t), \overline{u}_i(t))$.

Proof. We prove the claims regarding λ^x, M^x and $c_2^*(t)$ with the claims regarding λ^y, M^y and $c_1^*(t)$ proved analogously. While x is ex-post efficient by Lemma 6, this need not be true of x^n. Define, therefore, $\bar{x}^n(t)$ to be the vertical projection to the utility possibility frontier: $\bar{x}^n(t) = (x_1^n(t), f_2(t, x_1^n(t)))$, where this clearly also converges to x. By the fact that the bargaining problem is smooth, we have $f_2(t, .)$ is continuously differentiable on the set $(\underline{u}_1(t), \overline{u}_1(t))$, where this derivative is $f_2'(t, .)$. This function f_2' is continuously extended to the closed interval. Clearly, $f_2'(t, \bar{x}_1^n(t)) = -\frac{\lambda_1^n(t)}{\lambda_2^n(t)}$, where $\lambda^n(t)$ stands for $\lambda^{\bar{x}^n}(t)$. Then for some small $\varepsilon > 0$ with $\lambda_2^x(t) \neq \varepsilon$ define $\bar{\lambda}_2^n(t) = \max\{\varepsilon, \lambda_2^n(t)\}$, $\bar{\lambda}_1^n(t) = 1 - \bar{\lambda}_2^n(t)$, $\bar{\lambda}_2^x(t) = \max\{\varepsilon, \lambda_2^x(t)\}$ and $\bar{\lambda}_1^x(t) = 1 - \bar{\lambda}_2^x(t)$. Finally, define $\bar{M}^n(t) = \bar{\lambda}^n(t) \cdot \bar{x}^n(t)$ and $\bar{M}^x(t) = \bar{\lambda}(t) \cdot x(t)$. For our fixed ε, we claim:

$$y_2^{BS|x^n}(t) \geq \frac{\bar{M}^n(t)}{\bar{\lambda}_2^n(t)} - \gamma^n \frac{\bar{\lambda}_1^n(t)}{\bar{\lambda}_2^n(t)} x_1^n(t) - O((1 - \gamma^n)^2). \tag{8}$$

If $\lambda_2^x(t) > \varepsilon$ then $\lambda_2^n(t) > \varepsilon$ for sufficiently large n and equation (8) holds thanks to a Taylor's expansion of Agent 2's best safe payoff against x^n around $\bar{x}^n(t)$. The remainder $O((1-\gamma^n)^2)$ is a constant times a quadratic factor of the distance between $x_1^n(t)$ and $\gamma^n x_1^n(t)$; hence dividing it by $(1 - \gamma^n)^2$ gives an expression that converges to a constant as $\gamma^n \to 1$ (the smoothness assumption is important here). This is illustrated in Figure 2, where the the boundary of the linearized utility set $V_{\lambda^n}(t)$ is the line $z_2 = \frac{\bar{M}^n(t)}{\bar{\lambda}_2^n(t)} - \frac{\bar{\lambda}_1^n(t)}{\bar{\lambda}_2^n(t)} z_1$ which tangent to $\underline{U}(t)$ at $x^n(t)$ (where in this example $x^n(t) = \bar{x}^n(t)$ and $\lambda^n(t) = \bar{\lambda}^n(t)$). If on the other hand $\lambda_2^x(t) < \varepsilon$ so that $\lambda_2^n(t) < \varepsilon$ for sufficiently large n, then we directly have $y_2^{BS|x^n}(t) > \frac{\bar{M}^n(t)}{\bar{\lambda}_2^n(t)} - \gamma^n \frac{\bar{\lambda}_1^n(t)}{\bar{\lambda}_2^n(t)} x_1^n(t)$ for large

n because the slope of the linearized set is less steep than the slope of the utility frontier, i.e. $-\frac{1-\varepsilon}{\varepsilon} = -\frac{\bar{\lambda}_1^n(t)}{\bar{\lambda}_2^n(t)} > -\frac{\lambda_1^n(t)}{\lambda_2^n(t)} = f_2'(t, x_1^n(t))$. Also, if for all $\varepsilon > 0$ we have $\lambda_2^n(t) < \varepsilon$ for large enough n, then we have $x(t) = (\overline{u}_1(t), \underline{u}_2(t))$. Taking expectations,

$$E[y_2^{BS|x^n}|t_2] \geq E[\frac{\bar{M}^n}{\bar{\lambda}_2^n}|t_2] - \gamma^n E[\frac{\bar{\lambda}_1^n}{\bar{\lambda}_2^n}x_1^n|t_2] - O((1-\gamma^n)^2) \tag{9}$$

Moreover, we also have:

$$E[\frac{\bar{M}^n}{\bar{\lambda}_2^n}|t_2] - E[\frac{\bar{\lambda}_1^n}{\bar{\lambda}_2^n}x_1^n|t_2] = E[\bar{x}_2^n|t_2] \geq E[x_2^n|t_2] \geq E[x_2^{BS|y^n}|t_2] = \gamma^n E[y_2^n|t_2]. \tag{10}$$

The first equality follows from $\bar{M}^n = \bar{\lambda}^n \cdot \bar{x}^n$, the first inequality follows from $\bar{x}$'s definition, the second inequality follows from equilibrium conditions, and the second equality follows from the best safe's definition. So (10) implies

$$-E[\frac{\bar{\lambda}_1^n}{\bar{\lambda}_2^n}x_1^n|t_2] \geq \gamma^n E[y_2^n|t_2] - E[\frac{\bar{M}^n}{\bar{\lambda}_2^n}|t_2].$$

Combining this with (9) and the equilibrium condition $E[y_2^n|t_2] \geq E[y_2^{BS|x^n}|t_2]$,

$$E[y_2^n|t_2] \geq E[y_2^{BS|x^n}|t_2] \geq E[\frac{\bar{M}^n}{\bar{\lambda}_2^n}|t_2] + \gamma^n \left(\gamma^n E[y_2^n|t_2] - E[\frac{\bar{M}^n}{\bar{\lambda}_2^n}|t_2]\right) - O((1-\gamma^n)^2). \tag{11}$$

The above inequality simplifies to

$$(1-(\gamma^n)^2)E[y_2^n|t_2] \geq (1-\gamma^n)E[\frac{\bar{M}^n}{\bar{\lambda}_2^n}|t_2] - O((1-\gamma^n)^2). \tag{12}$$

Dividing this by $(1-(\gamma^n)^2) = (1-\gamma^n)(1+\gamma^n)$ we get:

$$E[y_2^n|t_2] \geq \frac{1}{1+\gamma^n}E[\frac{\bar{M}^n}{\bar{\lambda}_2^n}|t_2] - O(1-\gamma^n) \tag{13}$$

Since $x^n \to x$ and $\bar{x}^n \to x$, we have $\bar{\lambda}^n \to \bar{\lambda}^x$ and $\bar{M}^n \to \bar{M}^x$. Taking the limit of (13) as $n \to \infty$, and noting $E[x_i|t_i] = E[y_i|t_i]$ by Lemma 6 we get:

$$E[y_2|t_2] = E[x_2|t_2] = E[c_2^*|t_2] \geq \frac{1}{2}E[\frac{\bar{M}^x}{\bar{\lambda}_2^x}|t_2]. \tag{14}$$

Taking $\varepsilon \to 0$ we have $\frac{\bar{M}^x(t)}{\bar{\lambda}_2^x(t)} \to \frac{M^x(t)}{\lambda_2^x(t)}$ so long as $\lambda_2^x(t) > 0$. If $\lambda_2^x(t) = 0$ then $\frac{\bar{M}^x(t)}{\bar{\lambda}_2^x(t)}$ explodes as $\varepsilon \to 0$, contradicting the feasibility of equation (14) for all ε sufficiently small. This establishes that $\lambda_2^x(t) > 0$ and equation (7). Given that $\lambda_2^x(t) > 0$ it is clear that when $\underline{u}_2(t) > 0$ we must have $x_2(t) > \underline{u}_2(t)$ and $x_1(t) < \overline{u}_1(t)$ because otherwise we would have $\lambda_2^x(t) = 0$ when $x_2(t) = \underline{u}_2(t) > 0$. Finally, suppose $x_2(t') = \underline{u}_2(t') = 0$ and so $x_1(t') = \overline{u}_1(t')$ for some $t' \in T(t_2)$. Then equation (14) implies:

$$\sum_{t \in T(t_2)\setminus\{t'\}} \frac{p(t)}{p(t_2)} x_2(t) = E[x_2|t_2] \geq \frac{1}{2} E[\frac{M^x}{\lambda_2^x}|t_2]$$
$$\geq \frac{1}{2p(t_2)} [\frac{\lambda_1^x(t')}{\lambda_2^x(t')} p(t')\overline{u}_1(t') + \sum_{t \in T(t_2)\setminus\{t'\}} p(t)x_2(t)]$$

This is impossible if $T(t_2) = \{t'\}$, so suppose otherwise. Rearrange the far left and right terms above and use $\frac{\lambda_1^x(t')}{\lambda_2^x(t')} = -\frac{1}{f_1'(t',0)}$ and $x_2(t) \leq \overline{u}_2(t)$ to get:

$$-f_1'(t',0) \geq \frac{p(t')\overline{u}_1(t')}{\sum_{t \in T(t_2)\setminus\{t'\}} p(t)x_2(t)} \geq \frac{p(t')\overline{u}_1(t')}{\sum_{t \in T(t_2)\setminus\{t'\}} p(t)\overline{u}_2(t)}$$

Clearly, this cannot hold if the bargaining problem satisfies (BC). □

Lemma 8. *If c^* is interim-efficient and strictly individually rational, then it is a Myerson solution.*

Proof. If c^* is interim efficient then it is ex-post efficient, and so $\lambda^x(t) = \lambda^y(t)$, call it $\lambda(t)$, and $M^x = M^y = M = \lambda(t) \cdot c^*(t)$ (M^x and M^y are defined in the proof of Lemma 7). This is the unique orthogonal vector at $c^*(t)$ given that c^* is strictly individually rational. By Lemma 7 we know $\lambda_i(t) > 0$ for all i. By Lemma 1, c^* being interim efficient implies there exists a vector $\hat{\lambda} \in \mathbb{R}_{++}^{T_1} \times \mathbb{R}_{++}^{T_2}$ such that $\lambda_i(t) = \frac{\hat{\lambda}_i(t_i)}{p(t_i)}$ for all i and t. Hence (7) implies

$$\hat{\lambda}_i(t_i) E[c_i^*|t_i] \geq p(t_i) E[\frac{M}{2}|t_i]. \tag{15}$$

Summing up the inequalities in equation (15) over t_1 and over t_2, we get:

$$\sum_{i=1,2} \sum_{t_i \in T_i} \hat{\lambda}_i(t_i) E[c_i^*|t_i] \geq E[M], \tag{16}$$

Using the definitions of $M(t)$ and $\hat{\lambda}_i(t_i)$, we have:

$$\begin{aligned}
\sum_i \sum_{t_i \in T_i} \hat{\lambda}_i(t_i) E[c_i^*|t_i] &= \sum_{i=1,2} \sum_{t_i \in T_i} \hat{\lambda}_i(t_i) \sum_{t_{-i} \in T_{-i}} p(t_{-i}|t_i) c_i^*(t) \\
&= \sum_{t \in T} \sum_{i=1,2} p(t_{-i}|t_i) \hat{\lambda}_i(t_i) c_i^*(t) \\
&= \sum_{t \in T} p(t) \sum_{i=1,2} \lambda_i(t) c_i^*(t) \\
&= E[M]
\end{aligned}$$

But then (16) must hold with equality, and hence (15) must also hold with equality for each $i = 1, 2$, which means c^* is a Myerson solution. □

A5 Sequential Equilibria

Proof of Proposition 5 (Interim Efficiency at the Limit)

For this proposition, $\mathcal{B}$ is assumed to be smooth and satisfy (SBC) if $|T_i| > 2$ for some i. As explained in the main text, we can also assume without loss of generality that p is uniform. Suppose $c^* \in C^*(\mathcal{B})$ is not interim efficient. Let $\delta^n \to 1$ and consider an associated sequence of conciliatory equilibria. These need not be pooling equilibria, so let $x^n(t)$ and $y^n(t)$ correspond to agents actual equilibrium demands in state t. In other words, $x^n(t_1, t_2)$ is the demand of type t_1 in state (t_1, t_2) (rather than of type t_1') and $y^n(t_1, t_2)$ is the demand of type t_2. Considering a subsequence if necessary, let $(x^n, y^n) \to (x, y)$ and $c^* = \frac{x+y}{2}$. By Lemma 6, x and y are both ex-post efficient and $E[x_i|t_i] = E[y_i|t_i] = E[c_i^*|t_i]$ so that x and y are not interim efficient either. To get closer to the proof sketch provided in the main text, we would like to focus on type subsets with two elements for each agent. To do this, we could apply Lemma 5 to x. Unfortunately, while we know that $x_1(t) < \overline{u}_1(t)$ for all t, by Lemma 7 ((SBC) implies (BC)), we cannot be sure that $x_1(t) > \underline{u}_1(t)$, for all t. We must consider a complementary lemma to cover this case, which is proved in the Online Appendix.

Lemma 9. *If $x_1(t) = \underline{u}_1(t)$ for some t, then there are $T_i' \subset T_i$ for $i = 1, 2$ with $|T_i'| = 2$ and a contract e^* that interim dominates x when restricted to $T_1' \times T_2'$.*

If $|T_i| > 2$ for some agent i, then we know by Lemmas 5 and 9 that there is $T_j' \subset T_j$ for $j = 1, 2$ with $|T_j| = 2$ such that x is not interim efficient restricted to $T_1' \times T_2'$. If

$|T_i| = 2$ then let $T_i' = T_i$. By Lemma 4, x is not weakly interim efficient restricted to $T_1' \times T_2'$, and so there is some alternative (ex-post efficient) contract e which strictly interim dominates x when restricted to $T_1' \times T_2'$. Let e^* be defined by $e^* = e(t)$ if $t \in T_1' \times T_2'$ and $e^*(t) = x(t)$ otherwise. Clearly we have $E[e_i^*|t_i] > E[x_i|t_i]$ for all $t_i \in T_i'$ and $i = 1, 2$. Furthermore, let e^n be defined by $e^n = e(t)$ if $t \in T_1' \times T_2'$ and $e^n(t) = \hat{y}^{bs|x^n}(t)$ otherwise.[24] Given that $\hat{y}^{bs|x^n}(t) \to x(t)$, we clearly have $e^n \to e^*$ and hence $E[e_i^n|t_i] > E[x_i^n|t_i]$ for all $t_i \in T_i'$ and $i = 1, 2$ for sufficiently large n.

Consider a unilateral deviation for Agent 2, who proposes e^n. We show this deviation is profitable for some type in T_2' when n is large. The first step is to show that, for all sufficiently large n, some type of Agent 1 in T_1' is conciliatory after this deviation. Since e^* strictly interim dominates x when restricted to $T_1' \times T_2'$, we have:

$$\sum_{t \in \{t_1\} \times T_2'} (e_1^*(t) - x_1(t))p(t) > 0 \quad \text{and} \quad \sum_{t \in \{t_2\} \times T_1'} (e_2^*(t) - x_2(t))p(t) > 0 \qquad (17)$$

for all $t_1 \in T_1'$ and $t_2 \in T_2'$. If Agent 1 believes that 2 is conciliatory, then his payoff difference from being conciliatory instead of aggressive following the deviation is $\frac{1}{2}(e_1^n(t) - \gamma^n x_1^n(t))$ in state t. Let $\mu_1(t|t_1, e^n)$ be the probability that Agent 1 attributes to state t after 2's deviation, when of type t_1. Type $t_1 \in T_1$ is certainly conciliatory following e^n if

$$\mu_1^n(T_1' \times T_2'|t_1, e^n)[\sum_{t \in T(t_1) \cap T_1' \times T_2'} (e_1^n(t) - \gamma^n x_1^n(t))\mu_1^n(t|t_1, e^n, T_2')]$$
$$+ (1 - \mu_1^n(T_1' \times T_2'|t_1, e^n))[\sum_{t \in T(t_1) \setminus T_1' \times T_2'} (e_1^n(t) - \gamma^n x_1^n(t))\mu_1^n(t|t_1, e^n, T_2 \setminus T_2')] > 0.$$

As argued in the paper, we can assume an agent is always conciliatory if offered his best safe payoff in every state he considers possible. Hence, if $\mu_1^n(T_1' \times T_2'|t_1, e^n) = 0$ then Agent 1 is certainly conciliatory, in particular if $t_1 \notin T_1'$. Suppose $\mu_1^n(T_1' \times T_2'|t_1, e^n) > 0$. Clearly $e_1^n(t) - \gamma^n x_1^n(t) \geq 0$ if $t \notin T_1' \times T_2'$, and so type $t_1 \in T_1'$ must be conciliatory if $\sum_{t \in \{t_1\} \times T_2'} (e_1^n(t) - \gamma^n x_1^n(t))\mu_1^n(t|t_1, e^n, T_2') > 0$. Considering a subsequence if needed, say $\mu_1^n(t|t_1, T_2', e^n)$ converges to $\mu_1^*(t|t_1, T_2')$. As $\gamma^n x^n \to x$,

[24] Remember the definition of $\hat{y}^{bs|x^n}$ introduced at the very end of Section 3. It's derived from $y^{bs|x^n}$ to guarantee feasibility and ex-post efficiency.

$t_1 \in T_1'$ is conciliatory for all large n if

$$\sum_{t \in \{t_1\} \times T_2'} (e_1^*(t) - x_1(t))\mu_1^*(t|t_1, T_2') > 0. \tag{18}$$

By (17), $e_2^*(t') > x_2(t')$ for some $t' = (t_1', t_2')$ and so $e_1^*(t') < x_1(t')$ where we let $T_i' = \{t_i', t_i''\}$. As in footnote 25, we must have $e_2^*(t_1', t_2'') < x_2(t_1', t_2'')$, $e_2^*(t_1'', t_2'') > x_2(t_1'', t_2'')$ and $e_2^*(t_1'', t_2') < x_2(t_1'', t_2')$. Ex-post efficiency of e^* and x also implies $e_2^*(t_1', t_2') > x_2(t_1', t_2')$, $e_1^*(t_1', t_2'') > x_1(t_1', t_2'')$, $e_1^*(t_1'', t_2'') < x_1(t_1'', t_2'')$, and $e_1^*(t_1'', t_2') > x_1(t_1'', t_2')$. Suppose now, contradictory to what we set out to prove, that neither t_1', nor t_1'', take a conciliatory stand. Hence equation (18) is violated for both types, and it must be that $\mu_1^*((t_1', t_2')|t_1', T_2') > \frac{p(t_1', t_2')}{p(t_1' \times T_2')} = \frac{p(t_1'', t_2')}{p(t_1'' \times T_2')} > \mu_1^*((t_1'', t_2')|t_1'', T_2')$ (where the equality follows from assuming p is uniform from the start, without loss). Since equilibria along the sequence are sequential, it must be $\mu_1^*((t_1', t_2')|t_1', T_2') = \mu_1^*((t_1'', t_2')|t_1'', T_2')$. It follows from this contradiction that, as claimed, at least one of t_1' and t_1'' is conciliatory after e^n, for all large n. Without loss, assume this is type t_1'' for all sufficiently large n. Say type t_1' is conciliatory following e^n with probability $\alpha^n \in [0, 1]$ and assume $\alpha^n \to \alpha$ (consider a subsequence if needed).

We established above that $e_2^*(t_1'', t_2'') > x_2(t_1'', t_2'')$ while $e_2^*(t_1', t_2'') < x_2(t_1', t_2'')$. Agent 2 offering e^n and being conciliatory following Agent 1's offer of x^n ensures a limiting utility for type t_2'' of:

$$p(t_1'|t_2'')[\alpha \frac{e_2^*(t_1', t_2'') + x_2(t_1', t_2'')}{2} + (1-\alpha)x_2(t_1', t_2'')] + p(t_1''|t_2'')\frac{e_2^*(t_1'', t_2'') + x_2(t_1'', t_2'')}{2}$$
$$+ \sum_{t_1 \in T_1 \setminus T_1'} p(t_1|t_2'')x_2(t_1, t_2'').$$

This is decreasing in α given that $e_2^*(t_1', t_2'') < x_2(t_1', t_2'')$, and so is minimized when $\alpha = 1$. However, weknow that e^* strictly interim dominates x when restricted to $T_1' \times T_2'$, implying $E[e_2^*|t_2'', T_1'] > E[x_2|t_2'', T_1']$ and so the above deviation payoff is strictly larger than the limit of type t_2'' equilibrium payoffs $E[x_2|t_2''] = E[c_2^*|t_2'']$. Hence, deviating to e^n must be profitable for large n. $\square$