Online Appendix: Bargaining over Contingent Contracts Under Incomplete Information

Geoffroy de Clippel, Jack Fanning and Kareen Rozen

OA.1 Interim superior contracts when $\mathbf{x}$ is not interim efficient

This section states supporting lemmas for the proof of Proposition 6. Consider some sequence of smooth bargaining games satisfying (SBC) with $\delta^n \to 1$. Let $(\mathbf{x}^n, \mathbf{y}^n) \to (\mathbf{x}, \mathbf{y})$ be an associated conciliatory equilibrium demands, where limit outcome $\mathbf{c}^* = (\mathbf{x} + \mathbf{y})/2$ is not interim efficient. By Lemma 2, $E[x_i|t_i] = E[c^*_i|t_i]$ and so $\mathbf{x}$ is not interim efficient, but is weakly ex-post efficient with $x_2(\mathbf{t}) \geq \underline{u}_2(\mathbf{t})$. In the proof of Proposition 6 we claimed there must exist an ex-post efficient contract $\mathbf{e}$ which interim strictly dominates $\bar{\mathbf{x}}$ when restricted to $\mathcal{T}'_1 \times \mathcal{T}'_2$ with $|\mathcal{T}'_i| = 2$ and $\mathcal{T}'_i \subset \mathcal{T}_i$, where $\bar{\mathbf{x}}(\mathbf{t}) = (f_1(\mathbf{t}, x_2(\mathbf{t})), x_2(\mathbf{t})) \geq \mathbf{x}(\mathbf{t})$ is the horizontal projection of $\mathbf{x}$ onto the utility frontier (notice $\bar{\mathbf{x}}$ is ex-post efficient).

This claim is an implication of Lemma OA.1 and OA.2 if $x_i(\mathbf{t}) > \underline{u}_i(\mathbf{t})$ for all i and $\mathbf{t}$, where such an $\mathbf{x}$ must be ex-post efficient, and so $\mathbf{x} = \bar{\mathbf{x}}$. First Lemma OA.1 shows that there exists a contract that interim dominates $\mathbf{x} = \bar{\mathbf{x}}$ when restricted to some $\mathcal{T}'_1 \times \mathcal{T}'_2$ with $|\mathcal{T}'_i| = 2$.

Lemma OA. 1. *Consider a smooth bargaining problem where each agent has at least two types. Suppose $\mathbf{x}$ is an ex-post efficient contract with $x_i(\mathbf{t}) > \underline{u}_i(\mathbf{t})$ for $i = 1, 2$ and $\mathbf{t} \in \mathcal{T}$. If $\mathbf{x}$ is not interim efficient, then there are $\mathcal{T}'_i \subset \mathcal{T}_i$ for $i = 1, 2$ with $|\mathcal{T}'_i| = 2$ and a contract e^* that interim dominates $\mathbf{x}$ when restricted to $\mathcal{T}'_1 \times \mathcal{T}'_2$ with $|\mathcal{T}'_i| = 2$.*

Lemma OA.2 then shows that if $\mathbf{x} = \bar{\mathbf{x}}$ is ex-post efficient but interim dominated when restricted to $\mathcal{T}'_1 \times \mathcal{T}'_2$ then it is also interim strictly dominated. Hence, combined with Lemma OA.1 there must exist a contract that interim strictly dominates $\mathbf{x} = \bar{\mathbf{x}}$ when restricted to $\mathcal{T}'_1 \times \mathcal{T}'_2$.

Lemma OA. 2. *Suppose $|\mathcal{T}_i| = 2$ for $i = 1, 2$. If $\mathbf{x}$ is both ex-post efficient and weakly interim efficient, then it is also interim efficient.*

Although we know $x_2(\mathbf{t}) > \underline{u}_2(\mathbf{t})$ for all $\mathbf{t}$, by Lemma 3 ((SBC) implies (BC)), we cannot be sure $x_1(\mathbf{t}) > \underline{u}_1(\mathbf{t})$, for all $\mathbf{t}$. Suppose then that $x_1(\mathbf{t}) \leq \underline{u}_1(\mathbf{t})$ in some state. In this case, Lemma OA.3 directly shows there exists $\mathcal{T}'_i \subset \mathcal{T}_i$ for $i = 1, 2$ with $|\mathcal{T}'_i| = 2$ such that $\bar{\mathbf{x}}$ is not weakly interim efficient restricted to $\mathcal{T}'_1 \times \mathcal{T}'_2$ (given that the bargaining problem satisfies (SBC)), and so establishes our claim.

Lemma OA. 3. *If $x_1(\mathbf{t}) \leq \underline{u}_1(\mathbf{t})$ for some $\mathbf{t}$ and the bargaining problem satisfies (SBC), then there exists $\mathcal{T}'_i \subset \mathcal{T}_i$ for $i = 1, 2$ with $|\mathcal{T}'_i| = 2$ such that $\bar{\mathbf{x}}$ is not weakly interim efficient restricted to $\mathcal{T}'_1 \times \mathcal{T}'_2$.*

We can extend these claims further (and thus the claims of Proposition 6) when $|\mathcal{T}_i| = 2$. If the bargaining problem is smooth but may not satisfy (SBC), Lemma OA.4 directly shows $\bar{\mathbf{x}}$ is not weakly interim efficient. If the bargaining problem is not smooth but satisfies $|\arg\max_{\mathbf{u}(\mathbf{t})\in\mathcal{U}(\mathbf{t})} u_i(\mathbf{t})| = 1$ for all i and $\mathbf{t}$, Lemma 2 ensures $\mathbf{x}$ is ex-post efficient and so Lemma OA.2 implies it is not weakly interim efficient.

Lemma OA. 4. *For a smooth bargaining problem with two types for each agent, $|\mathcal{T}_i| = 2$ for $i = 1, 2$, if Agent 1's limit conciliatory equilibrium contract $\mathbf{x}$ is not interim efficient, then $\bar{\mathbf{x}}$ defined by $\bar{\mathbf{x}}(\mathbf{t}) = (f_1(\mathbf{t}, x_2(\mathbf{t})), x_2(\mathbf{t})) \geq \mathbf{x}(\mathbf{t})$ is not weakly interim efficient.*

We prove these and additional technical lemmas in section OA.6.

OA.2 Proof of Proposition 1 (Myerson uniqueness)

This section proves the claim of Proposition 1 that interim utilities are uniquely defined in the Myerson solution. For $\alpha > 0$ and $i = 1, 2$, let $S_i(\alpha, \mathbf{t})$ represent i's payoff under the equal-split solution when $\mathcal{V}(\mathbf{t})$ is expanded by allowing transfers at a rate of α utils for bargainer 2 against one util for bargainer 1:

$$S_1(\alpha, \mathbf{t}) = \frac{1}{2\alpha} \max_{\mathbf{v}\in\mathcal{V}(\mathbf{t})} (\alpha v_1 + v_2) \text{ and } S_2(\alpha, \mathbf{t}) = \alpha S_1(\alpha, \mathbf{t}).$$

Lemma OA. 5. *Take $\alpha'' > \alpha' > 0$, $\mathbf{x}' \in \arg\max_{\mathbf{v}\in\mathcal{V}(\mathbf{t})}(\alpha' v_1 + v_2)$ and $\mathbf{x}'' \in \arg\max_{\mathbf{v}\in\mathcal{V}(\mathbf{t})}(\alpha'' v_1 + v_2)$. The following comparative statics hold.*

(i) $x_1'' \geq x_1'$ *and* $x_2'' \leq x_2'$.

(ii) $S_1(\alpha', \mathbf{t}) \geq S_1(\alpha'', \mathbf{t})$, *the inequality being strict if* $x_2'' > 0$.

(iii) $S_2(\alpha', \mathbf{t}) \leq S_2(\alpha'', \mathbf{t})$, *the inequality being strict if* $x_1' > 0$.

Proof.

(i) Suppose $x_1' > x_1''$. Then $\alpha' x_1' + x_2' \geq \alpha' x_1'' + x_2''$ and $\alpha'' x_1'' + x_2'' \geq \alpha'' x_1' + x_2'$ lead to the contradiction $\alpha'(x_1' - x_1'') \geq x_2'' - x_2' \geq \alpha''(x_1' - x_1'')$.

(ii) We have $S_1(\alpha', \mathbf{t}) = (\alpha' x_1' + x_2')/(2\alpha') \geq (\alpha' x_1'' + x_2'')/(2\alpha') \geq (\alpha'' x_1'' + x_2'')/(2\alpha'') = S_1(\alpha'', \mathbf{t})$, with the second inequality being strict when $x_2'' > 0$.

(iii) We have $S_2(\alpha', \mathbf{t}) = (\alpha' x_1' + x_2')/2 \leq (\alpha'' x_1' + x_2')/2 \leq (\alpha'' x_1'' + x_2'')/2 = S_2(\alpha'', \mathbf{t})$, with the first inequality being strict when $x_1' > 0$.

□

Lemma OA. 6. *If $\mathbf{x}'$ and $\mathbf{x}''$ are two Myerson solutions with associated interim utility weights that are collinear, then each type of each bargainer is indifferent between $\mathbf{x}'$ and $\mathbf{x}''$ at the interim stage.*

Proof. This is immediate, since bargainer i's expected utility when of type t_i is equal to her expected utility from picking the midpoint in the ex-post linearized problems, which must coincide when linearizing at $\mathbf{x}'$ or $\mathbf{x}''$ when the associated interim utility weights are collinear. $\square$

Rescaling the problem if necessary, we assume the prior is uniform (the necessary transformation is highlighted in Section OA.1). Let $\hat{\boldsymbol{\lambda}}'$ ($\hat{\boldsymbol{\lambda}}''$) be the interim utility weight supporting $\mathbf{x}'$ ($\mathbf{x}''$) as a Myerson solution. Define

$$\alpha'(\mathbf{t}) = \frac{\hat{\lambda}_1'(t_1)}{\hat{\lambda}_2'(t_2)} \text{ and } \alpha''(\mathbf{t}) = \frac{\hat{\lambda}_1''(t_1)}{\hat{\lambda}_2''(t_2)}.$$

Bargainers' excess utility in state $\mathbf{t}$, when comparing their payoff in the solution to the midpoint of the linearized bargaining set, is:

$$\Delta_i'(\mathbf{t}) = x_i'(\mathbf{t}) - S_i(\alpha'(\mathbf{t}), \mathbf{t}) \text{ and } \Delta_i''(\mathbf{t}) = x_i''(\mathbf{t}) - S_i(\alpha''(\mathbf{t}), \mathbf{t}),$$

for $i = 1, 2$. By virtue of being Myerson solutions, we have:

$$\sum_{\mathbf{t} \in \mathcal{T}(t_i)} \Delta_i'(\mathbf{t}) = \sum_{\mathbf{t} \in \mathcal{T}(t_i)} \Delta_i''(\mathbf{t}) = 0, \tag{OA.1}$$

for $i = 1, 2$ and all $t_i \in \mathcal{T}_i$. It can also be immediately verified (i.e. given $\alpha'(\mathbf{t})x_1'(\mathbf{t}) + x_2'(\mathbf{t}) = 2S_2(\alpha', \mathbf{t})$) that:

$$\alpha'(\mathbf{t})\Delta_1'(\mathbf{t}) + \Delta_2'(\mathbf{t}) = \alpha''(\mathbf{t})\Delta_1''(\mathbf{t}) + \Delta_2''(\mathbf{t}) = 0, \tag{OA.2}$$

for all $\mathbf{t} \in \mathcal{T}$.

For each $\mathbf{t} = (t_1, t_2)$ such that $\Delta_1'(\mathbf{t}) \geq \Delta_1''(\mathbf{t})$, define $\mathbf{g}(\mathbf{t}) \in \mathcal{T}(t_1)$ as a minimizer of $\alpha'(\boldsymbol{\tau})/\alpha''(\boldsymbol{\tau})$ over $\hat{\mathcal{T}}(t_1) = \{\boldsymbol{\tau} \in \mathcal{T}(t_1) | \Delta_1'(\boldsymbol{\tau}) \leq \Delta_1''(\boldsymbol{\tau})\}$ (which is nonempty, by (OA.1)). If $\alpha'(\mathbf{g}(\mathbf{t})) > \alpha''(\mathbf{g}(\mathbf{t}))$, then $\alpha'(\boldsymbol{\tau}) > \alpha''(\boldsymbol{\tau})$, for all $\boldsymbol{\tau} \in \hat{\mathcal{T}}(t_1)$, in which case $\Delta_1'(\boldsymbol{\tau}) \geq \Delta_1''(\boldsymbol{\tau})$, for all $\boldsymbol{\tau} \in \mathcal{T}(t_1)$ (by Lemma OA.5 and the definition of $\hat{\mathcal{T}}$). This is compatible with (OA.1) only if this weak inequality is binding for all $\boldsymbol{\tau} \in \mathcal{T}(t_1)$. By Lemma OA.5 (*ii*), such an equality holds only if $\Delta_1'(\boldsymbol{\tau}) = \Delta_1''(\boldsymbol{\tau}) = \bar{u}_1(\mathbf{t})/2 > 0$, which would contradict (OA.1). Thus it must be that

$$\alpha'(\mathbf{g}(\mathbf{t})) \leq \alpha''(\mathbf{g}(\mathbf{t})). \tag{OA.3}$$

Since $\Delta_1'(\mathbf{g}(\mathbf{t})) \leq \Delta_1''(\mathbf{g}(\mathbf{t}))$, we must also have by (OA.2) that

$$\Delta_2'(\mathbf{g}(\mathbf{t})) \geq \Delta_2''(\mathbf{g}(\mathbf{t})). \tag{OA.4}$$

Similarly, for each $\mathbf{t} = (t_1, t_2)$ such that $\Delta_2'(\mathbf{t}) \geq \Delta_2''(\mathbf{t})$, define $\mathbf{h}(\mathbf{t}) \in \mathcal{T}(t_2)$ as a maximizer of $\alpha'(\boldsymbol{\tau})/\alpha''(\boldsymbol{\tau})$ over $\{\boldsymbol{\tau} \in \mathcal{T}(t_1) | \Delta_2'(\boldsymbol{\tau}) \leq \Delta_2''(\boldsymbol{\tau})\}$. Following a similar reasoning as in the previous paragraph, we conclude that

$$\alpha'(\mathbf{h}(\mathbf{t})) \geq \alpha''(\mathbf{h}(\mathbf{t})) \text{ and } \Delta_1'(\mathbf{h}(\mathbf{t})) \geq \Delta_1''(\mathbf{h}(\mathbf{t})). \tag{OA.5}$$

Suppose, by way of contradiction, that some type of some bargainer is not indifferent between $\mathbf{x}'$ and $\mathbf{x}''$ at the interim stage. By Lemma OA.6, $\hat{\boldsymbol{\lambda}}'$ and $\hat{\boldsymbol{\lambda}}''$ are not collinear, say $\alpha'(\mathbf{t}^*) > \alpha''(\mathbf{t}^*)$. Hence $\Delta_1'(\mathbf{t}^*) \geq \Delta_1''(\mathbf{t}^*)$, by Lemma OA.5. Construct the infinite sequence of states $(\mathbf{t}^k)_{k\geq 0}$ by $\mathbf{t}^0 = \mathbf{t}^*$, $\mathbf{t}^{k+1} = \mathbf{g}(\mathbf{t}^k)$ for all even $k \geq 0$ and $\mathbf{t}^{k+1} = \mathbf{h}(\mathbf{t}^k)$ for all odd $k \geq 1$. By (OA.3) and (OA.5), $\alpha'(\mathbf{t}^k) \geq \alpha''(\mathbf{t}^k)$ for even k's, and $\alpha'(\mathbf{t}^k) \leq \alpha''(\mathbf{t}^k)$ for odd k's. With finitely many states, there exist $0 \leq i < j$ with $\mathbf{t}^i = t^j$ and $\mathbf{t}^k \neq t^\ell$ for all $i \leq k < \ell \leq j$. We assume i is even (a similar argument applies if i is odd instead).

We now prove that $\alpha'(\mathbf{t}^k) = \alpha''(\mathbf{t}^k)$ for all $i \leq k \leq j$. This is obvious if $j = i+1$ (or $\mathbf{g}(\mathbf{t}^i) = t^i$), since $\alpha'(\mathbf{t}^i) \geq \alpha''(\mathbf{t}^i)$ (i is even) and $\alpha'(\mathbf{t}^j) \leq \alpha''(\mathbf{t}^j)$ (j is odd). Suppose then that $j > i+1$, in which case it must be even too in order to have $\mathbf{t}^i = t^j$. Let $\mathbf{t}^k = (t_1^k, t_2^k)$ and notice that $t_1^k = t_1^{k+1}$ for k even and $t_2^k = t_2^{k+1}$ odd. We have:

$$\begin{aligned}1 = \frac{\hat{\lambda}_2'(t_2^j)}{\hat{\lambda}_2'(t_2^i)} &= \frac{\hat{\lambda}_1'(t_1^i)}{\hat{\lambda}_2'(t_2^i)} \cdot \frac{\hat{\lambda}_2'(t_2^{i+1})}{\hat{\lambda}_1'(t_1^{i+1})} \cdot \frac{\hat{\lambda}_1'(t_1^{i+2})}{\hat{\lambda}_2'(t_2^{i+2})} \cdots \cdot \frac{\hat{\lambda}_2'(t_2^j)}{\hat{\lambda}_1'(t_1^j)} \\ &\geq \frac{\hat{\lambda}_1''(t_1^i)}{\hat{\lambda}_2''(t_2^i)} \cdot \frac{\hat{\lambda}_2''(t_2^{i+1})}{\hat{\lambda}_1''(t_1^{i+1})} \cdot \frac{\hat{\lambda}_1''(t_1^{i+2})}{\hat{\lambda}_2''(t_2^{i+2})} \cdots \cdot \frac{\hat{\lambda}_2''(t_2^j)}{\hat{\lambda}_1''(t_1^j)} = \frac{\hat{\lambda}_2''(t_2^j)}{\hat{\lambda}_2''(t_2^i)} = 1\end{aligned}$$

To avoid a contradiction $1 > 1$, $\alpha'(\mathbf{t}^k) = \alpha''(\mathbf{t}^k)$ for all $i \leq k \leq j$, as claimed.

By the previous paragraph, $\lambda_1'(\mathbf{t}_1^k)/\lambda_1''(\mathbf{t}_1^k) = \lambda_2'(\mathbf{t}_2^k)/\lambda_2''(\mathbf{t}_2^k)$, for all $i \leq k \leq j$. Since a single bargainer's type changes in moving from $\mathbf{t}^k$ to $\mathbf{t}^{k+1}$, an inductive argument establishes $\lambda_1'(t_1)/\lambda_1''(t_1) = \lambda_2'(t_2)/\lambda_2''(t_2)$, or $\alpha'(\mathbf{t}) = \alpha''(\mathbf{t})$, for all $\mathbf{t} \in \mathcal{A} \times \mathcal{B}$, where $\mathcal{A} = \{t_1^k | i \leq k \leq j\}$ and $\mathcal{B} = \{t_2^k | i \leq k \leq j\}$. Let $\bar{\mathcal{A}} = \{t_1 \in \mathcal{T}_1 | \alpha'(\mathbf{t}) = \alpha''(\mathbf{t})$, for some $t_2 \in \mathcal{B}\}$ and $\bar{\mathcal{B}} = \{t_2 \in \mathcal{T}_2 | \alpha'(\mathbf{t}) = \alpha''(\mathbf{t})$, for some $t_1 \in \mathcal{A}\}$. Similar reasoning implies $\lambda_1'(t_1)/\lambda_1''(t_1) = \lambda_2'(t_2)/\lambda_2''(t_2)$, or $\alpha'(\mathbf{t}) = \alpha''(\mathbf{t})$, for all $\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}$. Take now $\mathbf{t} = (t_1, t_2) \in \bar{\mathcal{A}} \times (\mathcal{T}_2 \setminus \bar{\mathcal{B}})$ and $t_2' \in \mathcal{B}$. Given $t_1 \in \bar{\mathcal{A}}$ we have $\alpha'((t_1, t_2')) = \alpha''((t_1, t_2'))$, or $\lambda_1'(t_1)/\lambda_1''(t_1) = \lambda_2'(t_2')/\lambda_2''(t_2')$ (*). Consider now $t_1' \in \mathcal{A}$. If $\alpha'((t_1', t_2)) < \alpha''((t_1', t_2))$, then $\Delta_1'((t_1', t_2)) \leq \Delta_1''((t_1', t_2))$ and we contradict $\alpha'(\mathbf{g}(\mathbf{t}^k)) = \alpha''(\mathbf{g}(\mathbf{t}^k))$ for k such that $t_1^k = t_1'$. Given that $\alpha'((t_1', t_2)) \neq \alpha''((t_1', t_2))$, since $t_2 \notin \bar{\mathcal{B}}$, it must be that $\alpha'((t_1', t_2)) > \alpha''((t_1', t_2))$, or $\lambda_1'(t_1')/\lambda_1''(t_1') > \lambda_2'(t_2)/\lambda_2''(t_2)$ (**). Since $(t_1', t_2') \in \mathcal{A} \times \mathcal{B}$, $\lambda_1'(t_1')/\lambda_1''(t_1') = \lambda_2'(t_2')/\lambda_2''(t_2')$ (***). Combining (*), (**) and (***), we conclude that $\lambda_1'(t_1)/\lambda_1''(t_1) > \lambda_2'(t_2)/\lambda_2''(t_2)$. Hence for all $\mathbf{t} \in \bar{\mathcal{A}} \times (\mathcal{T}_2 \setminus \bar{\mathcal{B}})$ we have $\alpha'(\mathbf{t}) > \alpha''(\mathbf{t})$, and $\Delta_1'(\mathbf{t}) \geq \Delta_1''(\mathbf{t})$, by Lemma OA.5; and similar reasoning implies $\alpha'(\mathbf{t}) < \alpha''(\mathbf{t})$, and hence $\Delta_2'(\mathbf{t}) \geq \Delta_2''(\mathbf{t})$, for all $\mathbf{t} \in (\mathcal{T}_1 \setminus \bar{\mathcal{A}}) \times \bar{\mathcal{B}}$.

Given the last sentence, (OA.1) implies $\sum_{\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}} \hat{\lambda}_1'(t_1)[\Delta_1'(\mathbf{t}) - \Delta_1''(\mathbf{t})] \leq 0$ and $\sum_{\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}} \hat{\lambda}_2'(t_2)[\Delta_2'(\mathbf{t}) - \Delta_2''(\mathbf{t})] \leq 0$. Combined together these give:

$$\sum_{\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}} \hat{\lambda}_1'(t_1)[\Delta_1'(\mathbf{t}) - \Delta_1''(\mathbf{t})] + \hat{\lambda}_2'(t_2)[\Delta_2'(\mathbf{t}) - \Delta_2''(\mathbf{t})] \leq 0.$$

Using (OA.2) and $\alpha'(\mathbf{t}) = \alpha''(\mathbf{t})$ for $\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}$, we get

$$\sum_{\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}} \Big(\hat{\lambda}_1'(t_1)[\Delta_1'(\mathbf{t}) - \Delta_1''(\mathbf{t})] + \hat{\lambda}_2'(t_2)[\Delta_2'(\mathbf{t}) - \Delta_2''(\mathbf{t})]\Big) = 0.$$

Hence, our inequalities must be equalities. In particular, $\Delta_1'(\mathbf{t}) = \Delta_1''(\mathbf{t})$ for all $\mathbf{t} \in \bar{\mathcal{A}} \times (\mathcal{T}_2 \setminus \bar{\mathcal{B}})$, and $\Delta_2'(\mathbf{t}) = \Delta_2''(\mathbf{t})$, for all $\mathbf{t} \in (\mathcal{T}_1 \setminus \bar{\mathcal{A}}) \times \bar{\mathcal{B}}$. This requires $\Delta_1'(\mathbf{t}) = \Delta_1''(\mathbf{t}) = \bar{u}_1(\mathbf{t})/2 > 0$ for $\mathbf{t} \in \bar{\mathcal{A}} \times (\mathcal{T}_2 \setminus \bar{\mathcal{B}})$ since $\alpha'(\mathbf{t}) > \alpha''(\mathbf{t})$, and $\Delta_2'(\mathbf{t}) = \Delta_2''(\mathbf{t}) = \bar{u}_2(\mathbf{t})/2 > 0$ for $\mathbf{t} \in (\mathcal{T}_1 \setminus \bar{\mathcal{A}}) \times \bar{\mathcal{B}}$ since $\alpha'(\mathbf{t}) < \alpha''(\mathbf{t})$. Given (OA.1), we must then have $\sum_{\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}} \hat{\lambda}_1'(t_1)\Delta_1'(\mathbf{t}) < 0$ (note $\bar{\mathcal{A}} \times (\mathcal{T}_2 \setminus \bar{\mathcal{B}}) \neq \emptyset$ as $t_2^* \notin \bar{\mathcal{B}}$ given $\alpha'(\mathbf{t}^*) > \alpha''(\mathbf{t}^*)$) and also $\sum_{\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}} \hat{\lambda}_2'(t_2)\Delta_2'(\mathbf{t}) \leq 0$. Combining these inequalities, however, gives $\sum_{\mathbf{t} \in \bar{\mathcal{A}} \times \bar{\mathcal{B}}} \hat{\lambda}_1'(t_1)\Delta_1'(\mathbf{t}) + \hat{\lambda}_2'(t_2)\Delta_2'(\mathbf{t}) < 0$, which contradicts (OA.2). □

OA.3 Characterization of Conciliatory Equilibria

This section proves the Propositions 2 and 3 which characterize conciliatory equilibrium.

Proof of Proposition 2 (Inscrutability). Take a separating conciliatory equilibrium. It is associated with partitions of the type spaces: $\mathcal{T}_1 = \mathcal{T}_1^{(1)} \cup \cdots \cup \mathcal{T}_1^{(m)}$, and $\mathcal{T}_2 = \mathcal{T}_2^{(1)} \cup \cdots \cup \mathcal{T}_2^{(n)}$. All types t_1 belonging to cell $\mathcal{T}_1^{(j)}$ propose $\mathbf{x}^{(j)}$, and all types t_2 belonging to cell $\mathcal{T}_2^{(j)}$ propose $\mathbf{y}^{(j)}$. We may assume wlog that $\mathbf{x}^{(j)} \neq \mathbf{x}^{(k)}$ and $\mathbf{y}^{(j)} \neq \mathbf{y}^{(k)}$ when $j \neq k$. We can thus define functions $j : \mathcal{T}_1 \to \{1, \ldots, m\}$ and $k : \mathcal{T}_2 \to \{1, \ldots, n\}$, where $j(t_1)$ is the index of the cell in the partition of $\mathcal{T}_1$ to which t_1 belongs ($t_1 \in \mathcal{T}_1^{(j(t_1))}$), and $k(t_2)$ is the index of the cell in the partition of $\mathcal{T}_2$ to which t_2 belongs ($t_2 \in \mathcal{T}_1^{(k(t_2))}$). Define best-safe contracts, $\mathbf{X}^{bs|\mathbf{y}^{(j)}}$ for each $j \in \{1, \ldots, n\}$ and $\mathbf{Y}^{bs|\mathbf{x}^{(k)}}$ for each $k \in \{1, \ldots, m\}$.

Consider a pooling strategy for Agent 1 where he offers $\mathbf{x}^*$ independently of t_1, with $\mathbf{x}^*(\mathbf{t}) = \mathbf{x}^{(j(t_1))}(\mathbf{t})$ for $\mathbf{t} = (t_1, t_2)$, and a pooling strategy for Agent 2 where he offers $\mathbf{y}^*$ independently of t_2, with $\mathbf{y}^*(\mathbf{t}) = \mathbf{y}^{(k(t_2))}(\mathbf{t})$ for $\mathbf{t} = (t_1, t_2)$. Followed by a conciliatory posture from all types, the same outcome prevails in all states as the original equilibrium. To conclude, we verify the conditions of Proposition 3.

The condition $E[x_1^*|t_1] \geq E[X_1^{bs|\mathbf{y}^*}|t_1]$ follows by observing that in the original separating equilibrium, if Agent 1 (of any type) were to instead propose $\bar{X}^{bs|\mathbf{y}^*}$ then all types of Agent 2 will be conciliatory, for whatever beliefs 2 may have following this deviation. This follows by a similar computation as in the proof of Proposition 3. In the original separating equilibrium, Agent 1 of a type $t_1 \in \mathcal{T}_1^{(j)}$ instead proposes $\mathbf{x}^{(j)}$, to which all types of Agent 2 respond with a conciliatory posture. The rationality of Agent 1 sending $\mathbf{x}^{(j)}$ thus requires that $E[x_1^{(j)}|t_1] \geq E[X_1^{bs|\mathbf{y}^*}|t_1]$. By construction of $\mathbf{x}^*$, when $t_1 \in \mathcal{T}_1^{(j)}$ we have $E[x_1^{(j)}|t_1] = E[\mathbf{x}_1^*|t_1]$, yielding the desired inequality. A symmetric argument for Agent 2 implies the condition $E[y_2^*|t_2] \geq E[Y_2^{bs|\mathbf{x}^*}|t_2]$.

To conclude, we show $E[x_2^*|t_2] \geq E[X_2^{bs|\mathbf{y}^*}|t_2]$ holds for all t_2 ($E[y_1^*|t_1] \geq E[Y_1^{bs|\mathbf{x}^*}|t_1]$ for all t_1 is derived analogously). After receiving $\mathbf{x}^{(j)}$ in the separating equilibrium, Agent 2 of type $t_2 \in \mathcal{T}_2^{(k)}$ has Bayesian-updated beliefs $p(t_1|t_2, \mathcal{T}_1^{(j)})$. By a similar computation as for Proposition 3, being conciliatory after Agent 1's proposal requires

$$\text{(OA.6)} \qquad E[x_|^{(j)} t_2, \mathcal{T}_1^{(j)}] \geq \gamma E[y_2^{(k)}|t_2, \mathcal{T}_1^{(j)}] = E[X_2^{bs|\mathbf{y}^{(k)}}|t_2, \mathcal{T}_1^{(j)}]$$

for every type $t_2 \in \mathcal{T}_2^{(k)}$, every $k \in \{1, \ldots, n\}$ and every $j \in \{1, \ldots, m\}$. Multiply the inequality (OA.6) associated with each $j \in \{1, \ldots, m\}$ by the probability $p(\mathcal{T}_1^{(j)}|t_2)$ and sum up the corresponding inequalities over all j. The resulting inequality is equivalent to the desired one by the construction of $\mathbf{x}^*$ and $\mathbf{y}^*$. □

Proof of Proposition 3 (Characterization of Conciliatory Equilibria). It remains to show sufficiency. Suppose $\mathbf{x}$ and $\mathbf{y}$ satisfy equation (3) from the paper. We construct a conciliatory equilibrium in which all types of Agent 1 propose $\mathbf{x}$ and all types of Agent 2 propose $\mathbf{y}$. After the offer $\mathbf{x}$, Agent 2's updated belief over Agent 1's type coincides with his interim belief, and being conciliatory is optimal, as $E[x_2|t_2] \geq E[X_2^{bs|\mathbf{y}}|t_2]$, for all $t_2 \in \mathcal{T}_2$. Similar reasoning applies to Agent 1 after $\mathbf{y}$.

We now define beliefs and strategies, and check incentives after a unilateral deviation. Without loss, suppose Agent 1 proposes $\mathbf{x}'$ instead, while 2 proposes $\mathbf{y}$. For any type t_2, define Agent 2's beliefs and action as follows. Let $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) = \{t_1 \in \mathcal{T}_1 : x_2'(t_1, t_2) < \gamma y_2(t_1, t_2)\}$. If $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) \neq \emptyset$, let the probability type t_2 believes that he faces t_1 given $\mathbf{x}'$ be $\mu_2(t_1|t_2, \mathbf{x}', \mathbf{y}) = 1$ for some $t_1 \in \mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y})$, so Agent 2 takes an aggressive stand against $\mathbf{x}'$. If $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) = \emptyset$ then let $\mu_2(t_1|t_2, \mathbf{x}', \mathbf{y}) = 1$ for some arbitrary $t_1 \in \mathcal{T}_1$, with Agent 2 conciliatory following $\mathbf{x}'$. Agent 1's belief following $\mathbf{y}$ coincides with his interim belief, and he is conciliatory following 2's proposal.

We show the off-equilibrium behavior following a unilateral deviation is rational. If Agent 2 expects 1 to be conciliatory towards $\mathbf{y}$, then it is rational to posture aggressively against 1's deviation $\mathbf{x}'$ given his off-equilibrium belief when $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) \neq \emptyset$, and to be conciliatory otherwise. For Agent 1, posturing aggressively against $\mathbf{y}$ after proposing $\mathbf{x}'$, when he is of type t_1, gives him $\delta \sum_{t_2 \in \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) x_1'(t_1, t_2)$ in expectation, where $\mathcal{T}_2(\mathbf{x}', \mathbf{y}) = \{t_2 : \mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) = \emptyset\}$ is the set of Agent 2's types who will be conciliatory after $\mathbf{x}'$. By being conciliatory, Agent 1 of type t_1 gets:

$$\sum_{t_2 \in \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) \frac{x_1'(t_1, t_2) + y_1(t_1, t_2)}{2} + \delta \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) y_1(t_1, t_2).$$

Multiplying payoffs by $2/\gamma$ and rearranging, being conciliatory is preferable if and only if

$$\sum_{t_2 \in \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) x_1'(t_1, t_2) \leq \frac{1}{\gamma} E[y_1|t_1] + \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) y_1(t_1, t_2).$$

Since $x_2'(\mathbf{t}) \geq \gamma y_2(\mathbf{t})$ for $\mathbf{t} = (t_1, t_2)$ such that $t_2 \in \mathcal{T}_2(\mathbf{x}', \mathbf{y})$ we must have $x_1'(\mathbf{t}) \leq X_1^{bs|\mathbf{y}}(\mathbf{t})$. Imposing this inequality as an equality and rearranging, a conciliatory posture is certainly preferable if

$$E[X_1^{bs|\mathbf{y}}|t_1] \leq \frac{1}{\gamma} E[y_1|t_1] + \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1)(y_1(t_1, t_2) + X_1^{bs|\mathbf{y}}(t_1, t_2)).$$

By equation (3) from the paper, we have $E[y_1|t_1] \geq E[Y_1^{bs|\mathbf{x}}|t_1] = \gamma E[x_1|t_1] \geq \gamma E[X_1^{bs|\mathbf{y}}|t_1]$. Hence, a conciliatory posture is preferable, since $\mathbf{y}(\mathbf{t}) \geq 0$ and $\mathbf{X}^{bs|\mathbf{y}}(\mathbf{t}) \geq 0$.

We now show that deviating from $\mathbf{x}$ to $\mathbf{x}'$ is not profitable for Agent 1. Agent 1's expected payoff is equal to $\delta y_1(\mathbf{t})$ in any state $\mathbf{t} = (t_1, t_2)$ where Agent 2 refuses $\mathbf{x}'$ (i.e. if $t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})$), and is equal to $(x_1'(\mathbf{t}) + y_1(\mathbf{t}))/2$ for states where 2 is conciliatory. So 1 has no strict incentive to deviate from $\mathbf{x}$ to $\mathbf{x}'$ if and only if

(OA.7)

$$\sum_{t_2 \in \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) \frac{x_1'(t_1, t_2) + y_1(t_1, t_2)}{2} + \delta \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) y_1(t_1, t_2) \leq \frac{E[x_1|t_1] + E[y_1|t_1]}{2}$$

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

$$\text{(OA.8)} \qquad \sum_{t_2 \in \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) x_1'(t_1, t_2) + \gamma \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})} p(t_2|t_1) y_1(t_1, t_2) \leq E[x_1|t_1].$$

Notice that $x_1'(t_1, t_2) \leq X_1^{bs|\mathbf{y}}(t_1, t_2)$ when $t_2 \in \mathcal{T}_2(\mathbf{x}', \mathbf{y})$, by definition of $\mathcal{T}_2(\mathbf{x}', \mathbf{y})$, and that $\gamma y_1(\mathbf{t}) \leq X_1^{bs|\mathbf{y}}(\mathbf{t})$, by definition of $X_1^{bs|\mathbf{y}}$. Thus the LHS of equation (OA.8) is less or equal to $E[X_1^{bs|\mathbf{y}}|t_1]$, which is less than the RHS of equation (OA.8), thanks to (3) from the paper. Thus Agent 1 does not find it profitable to unilaterally deviate to $\mathbf{x}'$, as claimed.

It remains to ensure there exist mutually optimal continuation strategies given beliefs after mutual deviations to $\mathbf{x}'$ and $\mathbf{y}'$. We define beliefs to be consistent with those after unilateral deviations, so $\mu_2(t_1|t_2, \mathbf{x}', \mathbf{y}') = \mu_2(t_1|t_2, \mathbf{x}', \mathbf{y})$, and $\mu_1(t_2|t_1, \mathbf{x}', \mathbf{y}') = \mu_1(t_1|t_2, \mathbf{x}, \mathbf{y}')$. These beliefs and posturing strategies determine expected continuation payoffs. Let those continuation payoffs correspond to payoff functions in an auxiliary posturing game with $\mathcal{T}_1 \cup \mathcal{T}_2$ players. That finite game has a Nash equilibrium, and we let postures following deviations $\mathbf{x}', \mathbf{y}'$ be given by one of those equilibria. □

OA.4 Non-Emptiness of $\mathcal{C}^*(\mathbf{B})$

This section establishes Proposition 4's claim that there is always at least one efficient limit of conciliatory equilibria, $\mathcal{C}^*(\mathbf{B}) \neq \emptyset$. We begin by defining the correspondence $\mathcal{F} : \mathcal{U} \rightrightarrows \mathcal{U}$ that associates to any contingent contract $\mathbf{v} \in \mathcal{U}$ the set of contingent contracts $\mathbf{u} \in \mathcal{U}$ that are weakly interim superior to $\mathbf{v}$:

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

The technical Lemma OA.7 (proved in section A4) establishes that this correspondence is continuous with non-empty, compact, and convex values

Lemma OA. 7. *$\mathcal{F}$ is continuous with non-empty, compact, and convex values.*

We use this result to establish the existence of an equilibrium with interim efficient demands.

Proposition OA. 1. *For any δ, there exists some pooling conciliatory equilibrium with interim-efficient demands.*

Proof. For $\gamma \leq 0$, the demands $(\mathbf{x}, \mathbf{y})$ with $\mathbf{x}(\mathbf{t}) = (\overline{u}_1(\mathbf{t}), \underline{u}_2(\mathbf{t}))$ and $\mathbf{y}(\mathbf{t}) = (\underline{u}_1(\mathbf{t}), \overline{u}_2(\mathbf{t}))$ form such an equilibrium. Henceforth, assume $\gamma > 0$ (equivalently $\delta > 1/2$).

Let $\overline{\mathcal{U}}(\mathbf{t}) = \{\mathbf{u} \in \mathbb{R}^2_+ | (\exists \mathbf{v} \in \mathcal{U}(\mathbf{t})) : \mathbf{u} \leq \mathbf{v}\}$ be the comprehensive closure of $U(\mathbf{t})$, and let $\mathcal{G} : \overline{\mathcal{U}} \rightrightarrows \mathcal{U}$ be the correspondence that is defined by $\mathcal{G}(\mathbf{v}) = \arg\max_{\mathbf{u}\in\mathcal{F}(\mathbf{v})} \prod_{t_i,\, i} (E[u_i|t_i] - E[v_i|t_i] + 1)$, where $\mathcal{F}$ was defined right before Lemma OA.7. The set $\mathcal{G}(\mathbf{v})$ is compact and convex, since it is obtained by maximizing a concave function over a compact and convex set. Clearly, it selects contingent contracts that are interim efficient in $\mathcal{U}$. The Theorem of the Maximum implies $\mathcal{G}$ is upper hemi-continuous ($\mathcal{F}$ is continuous, thanks to Lemma OA.7).

Let then $\mathcal{H} : \mathcal{U}^2 \rightrightarrows \mathcal{U}^2$ be the correspondence defined as follows: $\mathcal{H}(\mathbf{x}, \mathbf{y}) = (\mathcal{G}(\mathbf{X}^{bs|\mathbf{y}}), \mathcal{G}(\mathbf{Y}^{bs|\mathbf{x}}))$. This is well-defined since $\mathbf{X}^{bs|\mathbf{y}}$ and $\mathbf{Y}^{bs|\mathbf{x}}$ belong to $\bar{\mathcal{U}}$ (but not necessarily $\mathcal{U}$). Notice $\mathbf{X}^{bs|\mathbf{y}}$ is continuous in $\mathbf{y}$ and that $\mathbf{Y}^{bs|\mathbf{x}}$ is continuous in $\mathbf{x}$. Let $(\mathbf{x}, \mathbf{y})$ be a fixed-point of $\mathcal{H}$, by Kakutani's fixed point theorem. The construction of $\mathcal{H}$ ensures interim efficiency of $\mathbf{x}$ and $\mathbf{y}$, and that $E[x_i|t_i] \geq E[X_i^{bs|\mathbf{y}}|t_i]$ and $E[y_i|t_i] \geq E[Y_i^{bs|\mathbf{x}}|t_i]$ for all t_i and i. Hence, by Proposition 3, demands $(\mathbf{x}, \mathbf{y})$ are sustained by a pooling conciliatory equilibrium. □

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

Lemma OA. 8. *Consider a sequence of feasible contingent contracts $\mathbf{x}^n \to \mathbf{x} \in \mathcal{U}$. If each $\mathbf{x}^n$ is interim efficient, then $\mathbf{x}$ is interim efficient.*

We are now ready to prove the non-emptiness of $\mathcal{C}^*(\mathbf{B})$.

Proof of Proposition 4 ($\mathcal{C}^*(\mathbf{B})$ is non-empty). Fix a sequence $\delta^n \to 1$, and an associated sequence of pooling conciliatory equilibria with interim efficient demands $(\mathbf{x}^n, \mathbf{y}^n)$ (see Proposition OA.1). Since $\mathcal{U}(\mathbf{t})$ is compact, we may assume (considering a subsequence if needed) $(\mathbf{x}^n, \mathbf{y}^n)$ converges to a limit $(\mathbf{x}, \mathbf{y})$ as n tends to infinity. By Lemma OA.8, $\mathbf{x}$ and $\mathbf{y}$ are interim efficient. By Lemma 2, $E[x_i|t_i] = E[y_i|t_i]$ for all i, t_i. So the limit equilibrium outcome $(\mathbf{x} + \mathbf{y})/2$ is interim efficient and belongs to $\mathcal{C}^*(\mathbf{B})$. □

OA.5 Proof of Proposition 7 ($\mathcal{C}^{ns}(\mathbf{B}) \neq \emptyset$)

This section establishes Proposition 7's claim that conciliatory equilibria satisfying NSWYDK exist for large δ. This has two parts, we first show in Proposition OA.2 joint principal equilibrium exists in smooth problems satisfying (BC) and (NLB), and then use such an equilibrium to establish Proposition 7.

Proposition OA. 2. *For any smooth bargaining problem which satisfies (BC) and (NLB), there exists $\overline{\delta} < 1$ such that if $\delta > \overline{\delta}$ there is a joint principal equilibrium.*

Proof. Let $\mathcal{EPE}$ be the set of ex-post efficient contingent contracts in $\mathcal{U}$.

We now construct a correspondence from $\mathcal{EPE} \times \mathcal{EPE}$ into itself, and prove that it admits a fixed-point. We will use this result in the second-half of the proof to establish the existence of a joint principal equilibrium for sufficiently large δ.

Fix $(\mathbf{x}, \mathbf{y})$, a pair of ex-post efficient contingent contracts in $\mathcal{EPE}$. For $i = 1, 2$, we take a few steps to define a subset $\mathcal{F}_i(\mathbf{x}, \mathbf{y})$ of $\mathcal{EPE}$. We start by detailing the construction for $i = 1$. The case $i = 2$ proceeds analogously, as explained below.

First, let $\boldsymbol{\lambda}^{\mathbf{x}}(\mathbf{t}) \in \mathbb{R}^2_+$ denote the unique normalized unit vector that is orthogonal to $\mathcal{U}(\mathbf{t})$ at $\mathbf{x}(\mathbf{t})$ and continuous in $\mathbf{x}$. In each state $\mathbf{t}$, we expand $\mathcal{U}(\mathbf{t})$ using the supporting hyperplane defined by $\boldsymbol{\lambda}^{\mathbf{x}}(\mathbf{t})$. Then we select the payoff pair on that hyperplane that pays $\gamma y_2(\mathbf{t})$ to the second bargainer. This is not well-defined though if $\lambda_1^{\mathbf{x}}(\mathbf{t}) = 0$. For that purpose, we introduce a large number K,[1] and define the continuous function $\mathbf{g}$ by:

$$[\mathbf{g}(\mathbf{x}, \mathbf{y})](\mathbf{t}) = \left(\min\left\{K, (x_1(\mathbf{t}) + \frac{\lambda_2^{\mathbf{x}}(\mathbf{t})}{\lambda_1^{\mathbf{x}}(\mathbf{t})}(x_2(\mathbf{t}) - \gamma y_2(\mathbf{t}))\right\}, \gamma y_2(\mathbf{t}))\right).$$

Second, given that $\mathbf{g}(\mathbf{x}, \mathbf{y})$ typically falls outside of $\mathcal{U}$, we wish to project it back to feasible contingent contracts, in fact ones that are ex-post efficient. For the fixed-point to be useful, though, we have to proceed carefully. Let

$$\begin{aligned}\mathcal{H}(\mathbf{x}, \mathbf{y}) = \{\mathbf{u} \in \mathbb{R}^{\mathcal{T}_1}_+ | (\exists \mathbf{z} \in \mathcal{U})(\forall t_i) : & u(t_1) = E[z_1|t_1] \geq E[X_1^{bs|\mathbf{y}}|t_1], \\ & \text{and } E[z_2|t_2] = \gamma E[y_2|t_2]\}\end{aligned}$$

and define $\mathbf{h}(\mathbf{x}, \mathbf{y}) \in \mathcal{H}(\mathbf{x}, \mathbf{y})$ to be the vector of interim utilities for the first bargainer which is closest (minimum Euclidean distance) to the vector $(E[g(\mathbf{x}, \mathbf{y})_1|t_1])_{t_1 \in \mathcal{T}_1}$. It is not difficult to check that $\mathbf{h}$ is a continuous function.[2] We can then construct another continuous function $\mathbf{IE}$ such that $\mathbf{IE}(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{\mathcal{T}_1}_+ \times \mathbb{R}^{\mathcal{T}_2}_+$ is a interim efficient payoff profile satisfying the following inequalities for all t_1 and t_2:[3]

$$[IE(\mathbf{x}, \mathbf{y})]_1(t_1) \geq [\mathbf{h}(\mathbf{x}, \mathbf{y})]_1(t_1) \text{ and } [IE(\mathbf{x}, \mathbf{y})]_2(t_2) \geq \gamma E[y_2|t_2].$$

Finally, let $\mathcal{F}_1(\mathbf{x}, \mathbf{y})$ be the set of feasible contingent contracts that generate the interim utility profile $\mathbf{IE}(\mathbf{x}, \mathbf{y})$:

$$\mathcal{F}_1(\mathbf{x}, \mathbf{y}) = \{\mathbf{z} \in \mathcal{U} | (\forall i = 1, 2)(\forall t_i) : E[z_i|t_i] = [IE(\mathbf{x}, \mathbf{y})]_i(t_i)\}.$$

[1] K is taken large enough that $E[z_i|t_i] < Kp(t_{-i}|t_i)$ for all i, t_i and $\mathbf{z} \in \mathcal{U}$, that is, the expected utility of getting K in some state and zero elsewhere is infeasible.

[2] Clearly, H has compact, convex values and always contains $(E[X_1^{bs|\mathbf{y}}|t_1])_{t_1 \in \mathcal{T}_1}$. Lemma OA.7 implies that H is continuous. The Euclidean distance is continuous and so the theorem of the maximum implies that $\mathbf{h}$ is continuous.

[3] For instance, pick $\mathbf{IE}(\mathbf{x}, \mathbf{y})$ by maximizing the function $\prod_{t_i \in \mathcal{T}_i, i=1,2}(w_i(t_i) + 1)$ over the set of feasible interim utilities $w \in \mathbb{R}^{\mathcal{T}_1}_+ \times \mathbb{R}^{\mathcal{T}_2}_+$ that satisfy $w_1(t_1) \geq [h(\mathbf{x}, \mathbf{y})]_1(t_1)$ and $w_2(t_2) \geq \gamma E[y_2|t_2]$ for all t_i. Again the constraint set has compact, convex values and non-empty and continuous by Lemma OA.7 ensuring the continuity of $\mathbf{IE}$ by the theorem of the maximum.

By construction, any $\mathbf{z} \in F_1(\mathbf{x}, \mathbf{y})$ is interim efficient, and so ex-post efficient.

A symmetric construction applies to the second bargainer, which defines a correspondence $\mathcal{F}_2$ that associates a set of interim efficient contingent contract to any pair $(\mathbf{x}, \mathbf{y})$ of ex-post efficient contracts in $\mathcal{U}$.[4] The correspondence $\mathcal{F} = \mathcal{F}_1 \times \mathcal{F}_2$ is defined from $\mathcal{EPE} \times \mathcal{EPE}$ into itself, is upper-hemi continuous, and has compact, convex values. Since $\mathcal{EPE}$ is compact and homeomorphic to a convex set, $\mathcal{F}$ admits a fixed-point by Kakutani.

We now examine the properties of such fixed points $(\mathbf{x}, \mathbf{y}) \in \mathcal{F}(\mathbf{x}, \mathbf{y})$. First, notice that the interim efficient contracts $(\mathbf{x}, \mathbf{y})$ form an equilibrium because our construction ensured $[IE(\mathbf{x}, \mathbf{y})]_1(t_1) \geq [h(\mathbf{x}, \mathbf{y})]_1(t_1) \geq E[X_1^{bs|\mathbf{y}}|t_1]$ and $[IE(\mathbf{x}, \mathbf{y})]_2(t_2) \geq \gamma E[y_2|t_2]$ for all t_i.

We next claim that there exists $\overline{\delta} < 1$, such that for $\delta \geq \overline{\delta}$ we must have $x_i(\mathbf{t}) > \underline{u}_i(\mathbf{t})$. If this was not true, then there must exist some sequence of $\delta^n \to 1$ and associated sequence of equilibria arising from our fixed points such that for all n, $x_i^n(\mathbf{t}) \leq \underline{u}_i(\mathbf{t})$ for some player i and state $\mathbf{t}$. Considering a subsequence if necessary let $(\mathbf{x}^n, \mathbf{y}^n) \to (\mathbf{x}, \mathbf{y})$. By Lemma 3 and (BC), we must have $x_2(\mathbf{t}) > \underline{u}_2(\mathbf{t})$ and $y_1(\mathbf{t}) > \underline{u}_1(\mathbf{t})$ for all $\mathbf{t}$, and so it must be that $x_1(\mathbf{t}) \leq \underline{u}_1(\mathbf{t})$ for some $\mathbf{t}$. This combined with (NLB) implies that $(\mathbf{x} + \mathbf{y})/2$ is not ex-post efficient. By Lemma 2, however, we must have $E[x_i|t_i] = E[y_i|t_i] = E[x_i + y_i|t_i]/2$ for all t_i and by Lemma OA.8, $\mathbf{x}$ and $\mathbf{y}$ must be interim efficient. This contraction ensures $x_i(\mathbf{t}) > \underline{u}_i(\mathbf{t})$ for all sufficiently large δ. In this case, we clearly have a uniquely defined positive unit vector $\hat{\boldsymbol{\lambda}}^{\mathbf{x}} \in \Delta_{++}(\mathcal{T}_1) \times \Delta_{++}(\mathcal{T}_2)$ which is interim orthogonal to $\mathcal{U}(\mathbf{B})$ at $\mathbf{x}$.

We next claim that $E[x_1|t_1] = [h(\mathbf{x}, \mathbf{y})]_1(t_1) \leq E[g(\mathbf{x}, \mathbf{y})_1|t_1]$ and $E[x_2|t_2] = \gamma E[y_2|t_2]$ for all t_1, t_2. To establish this, first notice that $[h(\mathbf{x}, \mathbf{y})]_1(t_1) \leq E[g(\mathbf{x}, \mathbf{y})_1|t_1]$ for all t_1, or we could find points in $\mathcal{H}(\mathbf{x}, \mathbf{y})$ that are strictly closer to $(E[g(\mathbf{x}, \mathbf{y})_1|t_1])_{t_1 \in \mathcal{T}_1}$ than $\mathbf{h}(\mathbf{x}, \mathbf{y})$.[5] Also notice that if $[h(\mathbf{x}, \mathbf{y})]_1(t_1) = E[g(\mathbf{x}, \mathbf{y})_1|t_1]$ for all t_1 then we would have $E[x_1|t_1] = [h(\mathbf{x}, \mathbf{y})]_1(t_1) = E[g(\mathbf{x}, \mathbf{y})_1|t_1]$ and $E[x_2|t_2] = \gamma E[y_2|t_2] = E[g_2(\mathbf{x}, \mathbf{y})|t_2]$ for all t_2 because $\mathbf{g}(\mathbf{x}, \mathbf{y})$ is interim efficient in the bargaining problem where each $\mathcal{U}(\mathbf{t})$ is expanded by the supporting hyperplane defined by $\boldsymbol{\lambda}^{\mathbf{x}}(\mathbf{t}) = (\hat{\lambda}_1^{\mathbf{x}}(t_1)/p(t_1), \hat{\lambda}_2^{\mathbf{x}}(t_2)/p(t_2))$.[6] For our claim not to hold, therefore, there must be some type t_1' such that $[h(\mathbf{x}, \mathbf{y})]_1(t_1') < E[g(\mathbf{x}, \mathbf{y})_1|t_1']$.

By construction we have $E[x_i|t_i] = [IE(\mathbf{x}, \mathbf{y})]_i(t_i)$ and so $E[x_1|t_1] \geq [h(\mathbf{x}, \mathbf{y})]_1(t_1)$ and $E[x_2|t_2] \geq \gamma E[y_2|t_2]$. If $E[x_2|t_2'] > \gamma E[y_2|t_2']$ for some t_2', then we could increase $[h(\mathbf{x}, \mathbf{y})]_1(t_1')$ slightly to find a point in $\mathcal{H}(\mathbf{x}, \mathbf{y})$ that is closer to $(E[g(\mathbf{x}, \mathbf{y})_1|t_1])_{t_1 \in \mathcal{T}_1}$ than $\mathbf{h}(\mathbf{x}, \mathbf{y})$.[7] Similarly, if $E[x_1|t_1''] > [h(\mathbf{x}, \mathbf{y})]_1(t_1'')$ for some t_1'', then we could slightly

[4] Now $\gamma\mathbf{x}$ is used as an outside option for the first bargainer, and the second bargainer gets the remaining surplus in the linearized problem using in each $\mathbf{t}$ the vector that is orthogonal to $\mathcal{U}(\mathbf{t})$ at $\mathbf{y}(\mathbf{t})$. Details, which are simple once the construction of $\mathcal{F}_1$ is understood, are left to the reader.

[5] Consider $v_1(t_1) = \min\{[h(\mathbf{x}, \mathbf{y})]_1(t_1), E[g(\mathbf{x}, \mathbf{y})_1|t_1]\}$. We can always assume $\mathcal{U}$ is comprehensive and so ensure $\mathbf{v} \in \mathcal{H}(\mathbf{x}, \mathbf{y})$, because adding all less efficient utility pairs to $\mathcal{U}$ doesn't affect $\mathcal{EPE}$.

[6] The vector $\hat{\boldsymbol{\lambda}}^{\mathbf{x}}$ is interim orthogonal even given feasible utility sets $\mathcal{V}(\mathbf{t}) = \{\mathbf{v} \in \mathbb{R}^2_+ | \mathbf{v} \cdot \boldsymbol{\lambda}^{\mathbf{x}}(\mathbf{t}) \leq \mathbf{x}(\mathbf{t}) \cdot \boldsymbol{\lambda}^{\mathbf{x}}(\mathbf{t})\} \supseteq \mathcal{U}(\mathbf{t})$.

[7] Consider $\mathbf{v} \in \mathcal{H}(\mathbf{x}, \mathbf{y})$ defined by $v_1(t_1) = \min\{E[\hat{x}_1|t_1], E[g(\mathbf{x}, \mathbf{y})_1|t_1]\}$ where $\hat{\mathbf{x}}(\mathbf{t}) = \mathbf{x}(\mathbf{t})$ except

increase $[h(\mathbf{x},\mathbf{y})]_1(t_1')$ to find a point in $\mathcal{H}(\mathbf{x},\mathbf{y})$ that is closer to $(E[g(\mathbf{x},\mathbf{y})_1|t_1])_{t_1\in\mathcal{T}_1}$ than $\mathbf{h}(\mathbf{x},\mathbf{y})$.[8] This establishes the claim.

Identical logic applies to player 2's demand $\mathbf{y}$. Thus we are ready to show that $\mathbf{x},\mathbf{y}$ with interim orthogonal vectors $\hat{\boldsymbol{\lambda}}^{\mathbf{x}},\hat{\boldsymbol{\lambda}}^{\mathbf{y}}$ form a joint principal equilibrium. We have established $E[x_2|t_2]=\gamma E[y_2|t_2]$. Because $\mathbf{x}$ is interim efficient we have:

$$\begin{aligned}[g(\mathbf{x},\mathbf{y})]_1(\mathbf{t})\frac{\hat{\lambda}_1^{\mathbf{x}}(t_1)}{p(t_1)} =&x_1(\mathbf{t})\frac{\hat{\lambda}_1^{\mathbf{x}}(t_1)}{p(t_1)}+(x_2(\mathbf{t})-\gamma y_2(\mathbf{t}))\frac{\hat{\lambda}_2^{\mathbf{x}}(t_2)}{p(t_2)}\\ &=\max_{\mathbf{u}\in\mathcal{U}(\mathbf{t})} u_1\frac{\hat{\lambda}_1^{\mathbf{x}}(t_1)}{p(t_1)}+(u_2-\gamma y_2(\mathbf{t}))\frac{\hat{\lambda}_2^{\mathbf{x}}(t_2)}{p(t_2)}\end{aligned}$$

for all $\mathbf{t}=(t_1,t_2)$, where the first equality is by definition. We now multiply this by $p(\mathbf{t})$ and sum it up over $\mathbf{t}\in\mathcal{T}$ to get:

$$\begin{aligned}\sum_{t_1\in\mathcal{T}_1} E[g(\mathbf{x},\mathbf{y})_1|t_1]\hat{\lambda}_1^{\mathbf{x}}(t_1) &= \sum_{t_1\in\mathcal{T}_1,t_2\in\mathcal{T}_2} p(t_2|t_1)x_1(\mathbf{t})\hat{\lambda}_1^{\mathbf{x}}(t_1)+p(t_1|t_2)(x_2(\mathbf{t})-\gamma y_2(\mathbf{t}))\hat{\lambda}_2^{\mathbf{x}}(t_2)\\ &=\sum_{t_1\in\mathcal{T}_1} E[x_1|t_1]\hat{\lambda}_1^{\mathbf{x}}(t_1)\end{aligned}$$

where the second equality holds because $E[x_2|t_2]=\gamma E[y_2|t_2]$ for all t_2. But now using the established claim that $E[x_1|t_1]\le E[g(\mathbf{x},\mathbf{y})_1|t_1]$ for all t_1, it is clear that the above equality can only hold if $E[x_1|t_1]=E[g(\mathbf{x},\mathbf{y})_1|t_1]$ for all t_1. But in which case,

$$E[x_1|t_1]=E[g(\mathbf{x},\mathbf{y})_1|t_1]=\sum_{t_2\in\mathcal{T}_2} p(t_2|t_1)[\max_{\mathbf{u}\in\mathcal{U}(\mathbf{t})} u_1+(u_2-\gamma y_2(\mathbf{t}))\frac{p(t_1)\hat{\lambda}_2^{\mathbf{x}}(t_2)}{p(t_2)\hat{\lambda}_1^{\mathbf{x}}(t_1)}].$$

This establishes equation (7) for $(\mathbf{x},\hat{\boldsymbol{\lambda}}^{\mathbf{x}})$ in the joint principal equilibrium definition. Identical logic then applies to $(\mathbf{y},\hat{\boldsymbol{\lambda}}^{\mathbf{y}})$, establishing the result. □

We now show that beliefs in a joint principal equilibrium (which exists thanks to Proposition OA.2) can be made to satisfy NSWYDK. Rescaling if necessary, consider a joint principal equilibrium with demands $\mathbf{x},\mathbf{y}$ in $\mathbf{B}=(\mathcal{T},\mathcal{U},\mathbf{p})$, where $\mathbf{p}$ is a common uniform prior (the necessary transformation is highlighted in Section OA.7). Let $\hat{\boldsymbol{\lambda}}^{\mathbf{x}},\hat{\boldsymbol{\lambda}}^{\mathbf{y}}$ be interim orthogonal unit vectors associated with those joint principal equilibrium with demands. We will focus on showing that following an arbitrary deviation $\hat{\mathbf{x}}$ by Agent 1, we can find a belief $\boldsymbol{\pi}_2\in\Delta(\mathcal{T}_1)$ for Agent 2 and continuation acceptance strategies, which make the deviation unprofitable. Notice that we can transform $\mathbf{B}$

for in state (t_1',t_2') where $\hat{\mathbf{x}}$ gives player 2 slightly less than $\mathbf{x}$, and player 1 slightly more. This clearly implies $v_1(t_1')>[h(\mathbf{x},\mathbf{y})]_1(t_1')$.

[8] For some arbitrary t_2', let $\hat{\mathbf{x}}(\mathbf{t})=\mathbf{x}(\mathbf{t})$ except for in state (t_1',t_2') and (t_1'',t_2'). Let $\hat{\mathbf{x}}$ give player 2 slightly less than $\mathbf{x}$ in state (t_1',t_2') slightly more in state (t_1'',t_2'), with player 1 getting the residual, so that type t_2' obtains the same interim utility under $\hat{\mathbf{x}}$ and $\mathbf{x}$. The point $\mathbf{v}\in\mathcal{H}(\mathbf{x},\mathbf{y})$ defined by $\mathbf{v}(t_1)=\min\{E[\hat{x}_1|t_1],E[g(\mathbf{x},\mathbf{y})_1|t_1]\}$ is closer to $(E[g(\mathbf{x},\mathbf{y})_1|t_1])_{t_1\in\mathcal{T}_1}$ than $\mathbf{h}(\mathbf{x},\mathbf{y})$.

into a strategically equivalent problem $\mathbf{B}^{\mathbf{x}} = (\mathcal{T}, \mathcal{U}^{\mathbf{x}}, \mathbf{p})$ using the invertible mapping $\boldsymbol{\phi}^{\mathbf{x}} : \mathcal{U} \to \mathbb{R}_+^{\mathcal{T}}$, with $\phi^{\mathbf{x}}(\mathbf{u})_i(\mathbf{t}) = \hat{\lambda}_i^{\mathbf{x}}(t_i) u_i(\mathbf{t})$, where $\mathcal{U}^{\mathbf{x}} = \{\mathbf{v} \in \mathbb{R}_+^{\mathcal{T}} : \mathbf{v} = \boldsymbol{\phi}^{\mathbf{x}}(\mathbf{u}), \mathbf{u} \in \mathcal{U}\}$. This transformation, implies $E[u_i|t_i]\hat{\lambda}_i^{\mathbf{x}}(t_i) = E[\phi^{\mathbf{x}}(\mathbf{u})_i|t_i]$ for all $\mathbf{u} \in \mathcal{U}$. We have a joint principal equilibrium with demands $\mathbf{x}, \mathbf{y}$ in the original problem if and only if we have a joint principal equilibrium with demands $\boldsymbol{\phi}^{\mathbf{x}}(\mathbf{x}), \boldsymbol{\phi}^{\mathbf{x}}(\mathbf{y})$ in the transformed problem. Notice, that in this new bargaining problem the vector (1,1) is ex-post orthogonal to $\hat{\mathcal{U}}^{\mathbf{x}}(\mathbf{t})$ at $\boldsymbol{\phi}^{\mathbf{x}}(\mathbf{x})$ for all $\mathbf{t}$. To avoid unnecessary notation assume that the joint principal equilibrium demands and deviation in $\mathbf{B}^{\mathbf{x}}$ are in fact $\mathbf{x}, \mathbf{y}$ and $\hat{\mathbf{x}}$ (not $\boldsymbol{\phi}^{\mathbf{x}}(\mathbf{x}), \boldsymbol{\phi}^{\mathbf{x}}(\mathbf{y})$ and $\boldsymbol{\phi}^{\mathbf{x}}(\hat{\mathbf{x}})$).

If $\gamma \leq 0$ then it is an equilibrium for each agent to demand the entire surplus and to accept regardless of her demand or her opponent's demand. We can therefore assume that agents don't update their beliefs after a deviation in this case. Henceforth, therefore, we restrict attention to the case of $\gamma > 0$. We first specify that Agent 1 always accepts in the continuation game; it will be verified later that this behavior is optimal. Define $D(\mathbf{t}) = \gamma y_2(\mathbf{t}) - \hat{x}_2(\mathbf{t})$; this is twice the difference between Agent 2's payoff from rejecting and accepting. Given a belief $\boldsymbol{\pi}_2 \in \Delta(\mathcal{T}_1)$, type t_2 must accept if $\boldsymbol{\pi}_2 \cdot D(., t_2) < 0$, and reject if $\boldsymbol{\pi}_2 \cdot D(., t_2) > 0$.

Let $V_1(\mathbf{t})$ be the expected payoff for Agent 1 in state $\mathbf{t} = (t_1, t_2)$ when Agent 2 accepts with probability $\alpha(t_2)$ then:

$$\begin{aligned} 2V_1(\mathbf{t}) =& \hat{x}_1(\mathbf{t}) + y_1(\mathbf{t}) - (1 - \alpha(t_2))(\hat{x}_1(\mathbf{t}) - \gamma y_1(\mathbf{t})) \\ \leq& M(\mathbf{t}) + y_1(\mathbf{t}) - \hat{x}_2(\mathbf{t}) - (1 - \alpha(t_2))((1 - \gamma)M(\mathbf{t}) + D(\mathbf{t})), \end{aligned}$$

where $M(\mathbf{t}) = x_1(\mathbf{t}) + x_2(\mathbf{t})$ and the inequality uses $\hat{x}_1(\mathbf{t}) \leq M(\mathbf{t}) - \hat{x}_2(\mathbf{t})$ and $y_1(\mathbf{t}) \leq M(\mathbf{t}) - y_2(\mathbf{t})$. And so:

$$2E[V_1|t_1] \leq E[M + y_1 - \hat{x}_2 - (1-\alpha)((1-\gamma)M + D)|t_1] = E[x_1 + y_1 + \alpha D - (1-\alpha)(1-\gamma)M|t_1]$$

where the equality follows from $E[x_2 - \gamma y_2|t_1] = 0$ in a joint principal equilibrium. The equilibrium payoff of t_1 is $E[x_1 + y_1|t_1]/2$, implying that the deviation is unprofitable for type t_1 whenever $E[\alpha D - (1 - \alpha)(1 - \gamma)M|t_1] \leq 0$.

Below we outline a series of lemmas, which show that there always exists a belief and optimal acceptance strategies for player 2 that imply $E[\alpha D|t_1] \leq 0$ for all $t_1 \in \mathcal{T}_1$, ensuring the deviation is unprofitable.

Let Δ_n be the n-dimensional simplex, that is, $\Delta_n = \{\boldsymbol{\pi} \in \mathbb{R}_+^n | \sum_k \pi_k = 1\}$.

Lemma OA. 9. *Let $\mathbf{h} : \Delta_n \to \mathbb{R}^n$ be a continuous function such that $\boldsymbol{\pi} \cdot \mathbf{h}(\boldsymbol{\pi}) \leq 0$ for all $\boldsymbol{\pi} \in \Delta_n$. Then there exists $\boldsymbol{\pi} \in \Delta_n$ such that $\mathbf{h}(\boldsymbol{\pi}) \leq 0$.*

Proof. Suppose, on the contrary, that $\mathbf{h}(\boldsymbol{\pi}) \notin \mathbb{R}_-^n$ for all $\boldsymbol{\pi}$. For each $\varepsilon > 0$ and $\boldsymbol{\pi} \in \Delta_n$, let $\mathbf{h}^\varepsilon(\boldsymbol{\pi}) = \mathbf{h}(\boldsymbol{\pi}) - \varepsilon \mathbf{1}_n$. There exists $\eta > 0$ so that $\mathbf{h}^\eta(\boldsymbol{\pi}) \notin \mathbb{R}_-^n$ for all $\boldsymbol{\pi}$. Otherwise, one can construct a sequence $\boldsymbol{\pi}^m$ converging to some $\boldsymbol{\pi}^* \in \Delta_n$ such that $\mathbf{h}(\boldsymbol{\pi}^m) - \mathbf{1}_n/m \leq 0$, for each m. Since $\mathbf{h}$ is continuous, $\mathbf{h}(\boldsymbol{\pi}^*) \leq 0$, a contradiction.

Let $\mathcal{G} : \Delta_n \to \Delta_n$ be the correspondence associating to each $\boldsymbol{\pi}$ the set of vectors separating $\mathbf{h}^\eta(\boldsymbol{\pi})$ and $\mathbb{R}^n_-$:

$$\mathcal{G}(\boldsymbol{\pi}) = \{\hat{\boldsymbol{\pi}} \in \Delta_n | \hat{\boldsymbol{\pi}} \cdot \mathbf{h}^\eta(\boldsymbol{\pi}) \geq 0\}.$$

Clearly, $\mathcal{G}$ has nonempty convex values and a compact graph. But then it admits a fixed point $\boldsymbol{\pi}^* \in \Delta_n$, in which case $\boldsymbol{\pi}^* \cdot \mathbf{h}^\eta(\boldsymbol{\pi}^*) \geq 0$, a contradiction, since $\boldsymbol{\pi}^* \cdot \mathbf{h}^\eta(\boldsymbol{\pi}^*) = \boldsymbol{\pi}^* \cdot \mathbf{h}(\boldsymbol{\pi}^*) - \eta < 0$. □

Lemma OA. 10. *Let $\mathcal{F} : \Delta_n \to \mathbb{R}^n$ be a correspondence with nonempty convex values and a compact graph. Then, for each m there exists a continuous function $\mathbf{h}^m : \Delta_n \to \mathbb{R}^n$ such that* $\max_{\boldsymbol{\pi} \in \Delta_n} \min_{\mathbf{v} \in \mathcal{F}(\boldsymbol{\pi})} ||\mathbf{h}^m(\boldsymbol{\pi}) - \mathbf{v}|| \leq 1/m$.

Proof. This result is from von Neumann (1937); alternatively see Border (1985), p68.
□ □

Lemma OA. 11. *Let $\mathcal{F} : \Delta_n \to \mathbb{R}^n$ be a correspondence with nonempty convex values and a compact graph. If $\boldsymbol{\pi} \cdot \mathbf{v} \leq 0$ for all $\boldsymbol{\pi} \in \Delta_n$ and all $\mathbf{v} \in \mathcal{F}(\boldsymbol{\pi})$, then there exists $\boldsymbol{\pi} \in \Delta_n$ and $\mathbf{v} \in \mathcal{F}(\boldsymbol{\pi})$ such that $\mathbf{v} \leq 0$.*

Proof. Suppose, on the contrary, that $\mathcal{F}(\boldsymbol{\pi}) \cap \mathbb{R}^n_- = \emptyset$ for all $\boldsymbol{\pi}$. For $\varepsilon > 0$, define $\mathcal{F}^\varepsilon : \Delta_n \to \mathbb{R}^n$ as $\mathcal{F}^\varepsilon(\boldsymbol{\pi}) = \{\mathbf{x} - \varepsilon \mathbf{1}_n | \mathbf{x} \in \mathcal{F}(\boldsymbol{\pi})\}$. There exists $\eta > 0$ so that $\mathcal{F}^\eta(\boldsymbol{\pi}) \cap \mathbb{R}^n_- = \emptyset$ for all $\boldsymbol{\pi}$. Otherwise, one can construct a sequence $\boldsymbol{\pi}^k$ converging to some $\boldsymbol{\pi}^*$ and a sequence $\mathbf{v}^k$ converging to some $\mathbf{v}$ such that $\mathbf{v}^k \leq 0$ and $\mathbf{v}^k + \mathbf{1}_n/k \in \mathcal{F}(\boldsymbol{\pi}^k)$, for each k. Since $\mathcal{F}$ has a compact graph $\mathbf{v} \in \mathcal{F}(\boldsymbol{\pi}^*)$, which contradicts the assumption that $\mathcal{F}(\boldsymbol{\pi}^*) \cap \mathbb{R}^n_- = \emptyset$.

Using Lemma OA.10, let $(\mathbf{h}^m)_{m \geq 1}$ be a sequence of continuous functions from Δ_n into $\mathbb{R}^n$ such that $\max_{\boldsymbol{\pi} \in \Delta_n} \min_{\mathbf{v} \in \mathcal{F}^\eta(\boldsymbol{\pi})} ||\mathbf{h}^m(\boldsymbol{\pi}) - \mathbf{v}|| \leq 1/m$. Focusing on m's large enough, we'll have $\boldsymbol{\pi} \cdot \mathbf{h}^m(\boldsymbol{\pi}) \leq 0$, for all $\boldsymbol{\pi} \in \Delta^n$, since $\boldsymbol{\pi} \cdot \mathbf{v} \leq -\eta$ for all $\mathbf{v} \in \mathcal{F}^\eta(\boldsymbol{\pi})$. By Lemma OA.9, we can find for all such m's a vector $\boldsymbol{\pi}^m \in \Delta_n$ such that $\mathbf{h}^m(\boldsymbol{\pi}^m) \leq 0$. Consider now the sequence $\mathbf{v}^m$ such that $\mathbf{v}^m \in \mathcal{F}^\eta(\boldsymbol{\pi}^m)$ and $||\mathbf{h}^m(\boldsymbol{\pi}^m) - \mathbf{v}^m|| \leq 1/m$, for all m. We can assume without loss that $\boldsymbol{\pi}^m$ converges to some $\boldsymbol{\pi}^*$ and $\mathbf{v}^m$ converges to some $\mathbf{v}$. Hence $\mathbf{h}^m(\boldsymbol{\pi}^m)$ converges to $\mathbf{v}$, and it must be that $\mathbf{v} \leq 0$. Since $\mathcal{F}^\eta$ has a compact graph, we also have that $\mathbf{v} \in \mathcal{F}^\eta(\boldsymbol{\pi}^*)$, a contradiction. □

For each $\boldsymbol{\pi}_2 \in \Delta(\mathcal{T}_1)$, a possible belief for Agent 2, let $S(\boldsymbol{\pi}_2)$ be the set of 2's optimal strategies $\alpha \in [0,1]^{\mathcal{T}_2}$ when deciding whether or not to accept payoffs $-D$ given her belief $\boldsymbol{\pi}_2$: $\alpha(t_2) = 0$ (resp., 1) for all t_2 such that $\boldsymbol{\pi}_2 \cdot D(., t_2) > 0$ (resp., $<$). Let then $\mathcal{P}(\boldsymbol{\pi}_2)$ denote the set of 1's expected payoff vectors from B when 2 picks a strategy in $S(\boldsymbol{\pi}_2)$:

$$\mathcal{P}(\boldsymbol{\pi}_2) = \{(\sum_{t_2} \frac{\alpha(t_2)}{|\mathcal{T}_2|} D(t_1, t_2))_{t_1 \in \mathcal{T}_1} \mid \alpha \in S(\boldsymbol{\pi}_2)\}.$$

The next lemma shows that there exists a belief $\boldsymbol{\pi}_2$ and a best-response strategy for Agent 2 given that belief, α, which implies $E[\alpha D|t_1] \leq 0$ for all $t_1 \in \mathcal{T}_1$, making the deviation to $\hat{\mathbf{x}}$ unprofitable.

Lemma OA. 12. *There exists $\boldsymbol{\pi}_2 \in \Delta(\mathcal{T}_1)$ such that $\mathcal{P}(\boldsymbol{\pi}_2)$ contains a vector with non-positive components.*

Proof. The correspondence $\mathcal{P}$ has nonempty convex values and a compact graph. Notice that, for all $\boldsymbol{\pi}_2$, Agent 2 thinks 1 gets a non-positive payoff on average: $\boldsymbol{\pi}_2 \cdot \mathbf{v} \leq 0$, for all $\mathbf{v} \in \mathcal{P}(\boldsymbol{\pi}_2)$, because $\alpha(t_2)\boldsymbol{\pi}_2 \cdot D(., t_2) \leq 0$ for all t_2. The result then follows from Lemma OA.11. □

The final step of the proof is to ensure that player 1's strategy to always accept is optimal. Given player 2's acceptance strategy, type t_1's expected payoff from accepting $E[V_1|t_1]$ satisfies

$$2E[V_1|t_1] = E[\alpha\hat{x}_1 + y_1(1 + \gamma(1-\alpha))|t_1] \leq E[x_1 + y_1|t_1]$$

and so: $E[\alpha\hat{x}_1|t_1] \leq E[x_1 - \gamma y_1(1-\alpha)|t_1]$. Type t_1's payoff if she rejects is $W_1(t_1) = \delta E[\alpha\hat{x}_1|t_1]$ and so difference in payoff between accepting and rejecting satisfies:

$$\begin{aligned} 2(\mathcal{V}(t_1) - W(t_1)) =& E[-\gamma\alpha\hat{x}_1 + y_1(1 + \gamma(1-\alpha))|t_1] \\ \geq& E[-\gamma(x_1 - \gamma y_1(1-\alpha)) + y_1(1 + \gamma(1-\alpha))|t_1] \\ =& E[y_1 - \gamma x_1 + \gamma(1+\gamma)(1-\alpha)y_1|t_1] \geq 0 \end{aligned}$$

where the final inequality uses the fact that equilibrium demands must always satisfy $E[y_1 - \gamma x_1|t_1] \geq 0$. And so acceptance is always optimal for agent 1. This completes the proof of the Proposition. □

OA.6 Proofs of Technical Lemmas

This section provides proof for a series of technical lemmas, which were stated previously.

Proof of Lemma 1 (Efficiency and Weighted Utilitarianism). For (i), replacing any ex-post dominated contract would improve interim utilities. Sufficient conditions in (ii)-(iii) are easy to check. Necessity follows from the separating hyperplane theorem. It remains to show (iv). Observe that:

$$\sum_{i=1,2}\sum_{t_i \in \mathcal{T}_i} \hat{\lambda}_i(t_i)E[y_i|t_i] = \sum_{i=1,2}\sum_{t_i \in \mathcal{T}_i} \hat{\lambda}_i(t_i) \sum_{t_{-i} \in \mathcal{T}_{-i}} \frac{p(t_i, t_{-i})}{p(t_i)} y_i(\mathbf{t}) = \sum_{\mathbf{t} \in \mathcal{T}} p(\mathbf{t}) \sum_{i=1,2} \frac{\hat{\lambda}_i(t_i)}{p(t_i)} y_i(\mathbf{t}).$$

If $\mathbf{y} = \mathbf{x}$ maximizes the LHS, it also maximizes the RHS. Hence, for each $\mathbf{t} \in \mathcal{T}$, $\mathbf{y}(\mathbf{t}) = \mathbf{x}(\mathbf{t})$ must also maximize $\sum_{i=1,2} \frac{\lambda_i(t_i)}{p(t_i)} y_i(\mathbf{t})$. Similarly, if $\mathbf{y}(\mathbf{t}) = \mathbf{x}(\mathbf{t})$ maximizes $\sum_{i=1,2} \frac{\hat{\lambda}_i(t_i)}{p(t_i)} y_i(\mathbf{t})$ for each $\mathbf{t} = 1, 2$ then it maximizes the LHS. □

Proof of Lemma 2. For (i), observe that in a conciliatory equilibrium, we must have $E[x_2^n|t_2] \geq E[X_2^{bs|\mathbf{y}^n}|t_2] = \gamma^n E[y_2^n|t_2]$. 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 $E[y_2^n|t_2] \geq E[Y_2^{bs|\mathbf{x}^n}|t_2] \geq E[x_2^n|t_2]$. Hence $E[y_2|t_2] \geq E[x_2|t_2]$ and so $E[y_2|t_2] = E[x_2|t_2]$. By identical logic $E[y_1|t_1] = E[x_1|t_1]$.

We now prove (ii). If $f_2(\mathbf{t}', x_1(\mathbf{t}')) > x_2(\mathbf{t}')$ for some $\mathbf{t}'$, then $E[y_2^n|t_2'] \geq E[Y_2^{bs|\mathbf{x}^n}|t_2']$ and $Y_2^{bs|\mathbf{x}^n}(\mathbf{t}) \geq f_2(\mathbf{t}, \gamma^n x_1^n(\mathbf{t})) \geq \gamma^n x_2^n(\mathbf{t})$ for all $\mathbf{t}$, imply that $E[y_2|t_2'] \geq \lim E[Y_2^{bs|\mathbf{x}^n}|t_2'] \geq E[f_2(\mathbf{t}, x_1(\mathbf{t}))|t_2'] > E[x_2|t_2']$, a contradiction to (i). Given $f_2(\mathbf{t}, x_1(\mathbf{t})) = x_2(\mathbf{t})$ for all $\mathbf{t}$, if $\mathbf{x}$ is not weakly ex-post efficient then we must have $x_1(\mathbf{t}') < \underline{u}_1(\mathbf{t}')$ for some $\mathbf{t}'$ and $\overline{u}_2(\mathbf{t}') > f_2(\mathbf{t}', x_1(\mathbf{t}')) = x_2(\mathbf{t}')$. Then $x_1^n(\mathbf{t}') < \underline{u}_1(\mathbf{t}')$ for large n, and so $Y_2^{bs|\mathbf{x}^n}(\mathbf{t}') = \overline{u}_2(\mathbf{t}') > f_2(\mathbf{t}', x_1(\mathbf{t}')) = x_2(\mathbf{t}')$. Hence $E[y_2|t_2'] \geq \lim E[Y_2^{bs|\mathbf{x}^n}|t_2'] > E[f_2(\mathbf{t}, x_1(\mathbf{t}))|t_2'] = E[x_2|t_2']$, contradicting (i). Given that $\mathbf{x}$ is weakly ex-post efficient and $f_2(\mathbf{t}, x_1(\mathbf{t})) = x_2(\mathbf{t})$, it immediately follows that $x_2(\mathbf{t}) \geq \underline{u}_2(\mathbf{t})$. □

Proof of Lemma OA.1 (Interim domination for some types). Let $\boldsymbol{\lambda}(\mathbf{t}) \in \mathbb{R}^2_{++}$ be the unique strictly positive orthogonal unit vector to $\mathcal{U}(\mathbf{t})$ at $\mathbf{x}(\mathbf{t})$, for each $\mathbf{t} \in \mathcal{T}_1 \times \mathcal{T}_2$. An orthogonal unit vector exists because $\mathbf{x}$ is ex-post efficient, it is strictly positive by the fact that $x_i(\mathbf{t}) \in (\underline{u}_i(\mathbf{t}), \overline{u}_i(\mathbf{t}))$, and it is unique by smoothness and $x_i(\mathbf{t}) \in (\underline{u}_i(\mathbf{t}), \overline{u}_i(\mathbf{t}))$.

Suppose there is no contract e^* which is more efficient than $\mathbf{x}$ when restricted to $\mathcal{T}_1' \times \mathcal{T}_2'$ for any $\mathcal{T}_1' \times \mathcal{T}_2'$ such that $|\mathcal{T}_i'| = 2$. We prove this implies $\mathbf{x}$ is interim efficient, a contradiction. To do this we construct $\tilde{\lambda}_i(t_i) > 0$ for all t_i and i such that $(\tilde{\lambda}_1(t_1)/p(t_1), \tilde{\lambda}_2(t_2)/p(t_2))$ is collinear with $\boldsymbol{\lambda}(t_1, t_2)$ which must imply that $\mathbf{x}$ is interim efficient by Lemma 1. Fix $\bar{t}_2 \in \mathcal{T}_2$. Let:

$$\eta(t_1, t_2) = \frac{\lambda_1(t_1, t_2)}{\lambda_2(t_1, t_2)} \frac{p(t_1)}{p(t_2)}, \ \tilde{\lambda}_1(t_1) = \eta(t_1, \bar{t}_2), \text{ and } \tilde{\lambda}_2(t_2) = \frac{\eta(t_1, \bar{t}_2)}{\eta(t_1, t_2)}$$

for all (t_1, t_2). With this definition, $\tilde{\lambda}_1(t_1)/\tilde{\lambda}_2(t_2) = \eta(t_1, t_2) = (\lambda_1(t_1, t_2)/\lambda_2(t_1, t_2))(p(t_1)/p(t_2))$. It remains to show $\tilde{\lambda}_2(t_2)$ is well-defined, that is, $\eta(t_1, \bar{t}_2)/\eta(t_1, t_2)$ is independent of t_1, for all t_2. To establish this, consider arbitrary distinct types t_1', t_1'', $t_2' \neq \bar{t}_2$ and let $\mathcal{T}_1' = \{t_1', t_1''\}$ and $\mathcal{T}_2' = \{t_2', \bar{t}_2\}$. By definition of η, we have:

$$\text{(OA.9)} \qquad \frac{\eta(t_1, \bar{t}_2)}{\eta(t_1, t_2)} = \frac{\lambda_1(t_1, \bar{t}_2)}{\lambda_2(t_1, \bar{t}_2)} \frac{p(t_1)}{p(\bar{t}_2)} \frac{\lambda_2(t_1, t_2)}{\lambda_1(t_1, t_2)} \frac{p(t_2)}{p(t_1)}.$$

Next define:

$$p'(t_i) = p(t_i|\mathcal{T}_i' \times \mathcal{T}_j') = \frac{p(t_i)p(\mathcal{T}_j'|t_i)}{p(\mathcal{T}_i' \times \mathcal{T}_j')}$$

By Lemma 1 $\mathbf{x}$ is interim efficient when restricted to $\mathcal{T}_1' \times \mathcal{T}_2'$ if and only if there exists $\hat{\boldsymbol{\lambda}}(t_i) > 0$ for $t_i \in \mathcal{T}_i'$ such that

$$\frac{\lambda_1(t_1, t_2)}{\lambda_2(t_1, t_2)} = \frac{\hat{\lambda}_1(t_1)p'(t_2)}{\hat{\lambda}_2(t_2)p'(t_1)} = \frac{\hat{\lambda}_1(t_1)p(t_2)p(\mathcal{T}_1'|t_2)}{\hat{\lambda}_2(t_2)p(t_1)p(\mathcal{T}_2'|t_1)},$$

or

$$\frac{\lambda_1(t_1,t_2)p(t_1)}{\lambda_2(t_1,t_2)p(t_2)} = \frac{\hat{\lambda}_1(t_1)p(\mathcal{T}_1'|t_2)}{\hat{\lambda}_2(t_2)p(\mathcal{T}_2'|t_1)},$$

for all $(t_1,t_2) \in \mathcal{T}_1' \times \mathcal{T}_2'$. Plugging this into equation (OA.9) we get

$$\frac{\eta(t_1,\bar{t}_2)}{\eta(t_1,t_2)} = \frac{\hat{\lambda}_1(t_1)p(\mathcal{T}_1'|\bar{t}_2)}{\hat{\lambda}_2(\bar{t}_2)p(\mathcal{T}_2'|t_1)}\frac{\hat{\lambda}_2(t_2)p(\mathcal{T}_2'|t_1)}{\hat{\lambda}_1(t_1)p(\mathcal{T}_1'|t_2)} = \frac{\hat{\lambda}_2(t_2)p(\mathcal{T}_1'|\bar{t}_2)}{\hat{\lambda}_2(\bar{t}_2)p(\mathcal{T}_1'|t_2)}$$

for each $t_1 \in \mathcal{T}_1'$, from which we conclude $\eta(t_1,\bar{t}_2)/\eta(t_1,t_2)$ is independent of t_1. Hence $\tilde{\lambda}_2(t_2)$ is well defined. □

Proof of Lemma OA.2 (Implication of weak interim efficiency for two types). Suppose $\mathbf{x}$ is not interim efficient (but by assumption is ex-post efficient). Then let $\mathbf{z}$ be an ex-post efficient contract that interim dominates $\mathbf{x}$. For any $\alpha \in (0,1)$, $\mathbf{z}^\alpha = \boldsymbol{\lambda}\mathbf{z}+(1-\boldsymbol{\lambda})\mathbf{x}$ interim dominates $\mathbf{x}$.

First suppose that $z_i^\lambda(\mathbf{t}) > \underline{u}_i(\mathbf{t})$ for all i,t. To fix ideas, say that $E[z_1^\lambda|t_1'] > E[x_1|t_1']$ for some $t_1' \in \mathcal{T}_1$ where $\mathcal{T}_i = \{t_i', t_i''\}$. Define $\hat{\mathbf{z}}(\mathbf{t}) = \mathbf{z}^\alpha(\mathbf{t})$ if $\mathbf{t} \notin \mathcal{T}(t_1')$ and $\hat{\mathbf{z}}(\mathbf{t}) = (z_1^\alpha(\mathbf{t}) - \varepsilon, f_2(\mathbf{t}, z_1^\alpha(\mathbf{t}) - \varepsilon))$ otherwise. For $\varepsilon > 0$ small enough we clearly have $\hat{z}_i(\mathbf{t}) > \underline{u}_i(\mathbf{t})$, $E[\hat{z}_1|t_1'] > E[x_1|t_1']$ and $E[\hat{z}_2|t_2] > E[x_2|t_2]$ for $t_2 \in \mathcal{T}_2$. Finally define $\mathbf{z}^*(\mathbf{t}) = (f_1(\mathbf{t}, \hat{z}_2(\mathbf{t}) - \varepsilon'), \hat{z}_2(\mathbf{t}) - \varepsilon')$ otherwise. For $\varepsilon' > 0$ small enough, $E[z_i^*|t_i] > E[x_i|t_i]$ for all i and all t_i. Hence $\mathbf{x}$ is not weakly interim efficient, a contradiction.

It remains to consider the case $z_i^\alpha(\mathbf{t}) = \underline{u}_i(\mathbf{t})$ for some i,t. It must then be that $\mathbf{x}(\mathbf{t}) = \mathbf{z}(\mathbf{t})$ since both $\mathbf{x}$ and $\mathbf{z}$ are ex-post efficient. But then it must be that $\mathbf{x} = \mathbf{z}$,[9] which contradicts the fact that $\mathbf{z}$ interim dominates $\mathbf{x}$, and establishes the lemma. □

Proof of Lemma OA.3 (Remaining case). By Lemma 3 ((SBC) implies (BC)) we have $y_1(\mathbf{t}) > \underline{u}_1(\mathbf{t})$ for all $\mathbf{t}$. Let $\mathcal{T}_i' = \{t_i', t_i''\}$. By assumption we have that $x_1(\mathbf{t}') \leq \underline{u}_1(\mathbf{t}') < y_1(\mathbf{t}')$ for some $\mathbf{t}' = (t_1', t_2')$ and so we also have $y_2(\mathbf{t}') < x_2(\mathbf{t}') = \overline{u}_2(\mathbf{t}')$ (by Lemma 2 we know $\mathbf{x}$ is weakly ex-post efficient and satisfies $E[x_i|t_i] = E[y_i|t_i]$ for all t_i). Because $E[x_i|t_i'] = E[y_i|t_i']$ we must have some states (t_1', t_2'') and (t_1'', t_2') such that $x_1(t_1', t_2'') > y_1(t_1', t_2'') > \underline{u}_1(t_1', t_2'')$ and $x_1(t_1'', t_2') > y_1(t_1'', t_2') > \underline{u}_1(t_1'', t_2')$. Finally, we have $x_1(t_1'', t_2'') < \overline{u}_1(t_1'', t_2'')$, again by Lemma 3.

Recall that $\bar{\mathbf{x}}$ is defined by $\bar{\mathbf{x}}(\mathbf{t}) = (f_1(\mathbf{t}, x_2(\mathbf{t})), x_2(\mathbf{t}))$ and consider the alternative allocation $\mathbf{e}^*$ defined by $e_2^*(t_1', t_2') = \bar{x}_2(t_1', t_2') - \varepsilon$, $e_2^*(t_1'', t_2') = \bar{x}_2(t_1'', t_2') + K\varepsilon$, $e_2^*(t_1'', t_2'') = \bar{x}_2(t_1'', t_2'') - K'\varepsilon$, $e_2^*(t_1', t_2'') = \bar{x}_2(t_1', t_2'') + K''\varepsilon$ and $e_1^*(\mathbf{t}) = f_1(\mathbf{t}, e_2^*(\mathbf{t}))$ for some $\varepsilon, K, K', K'' > 0$. Choosing $K > p(t_1', t_2')/p(t_1'', t_2')$ ensures that $E[e_2^*|t_2', \mathcal{T}_1'] >$

[9] Otherwise, there is a state – say (t_1', t_2') – where an agent – say 1 – gets strictly more under $\mathbf{x}$ than under $\mathbf{z}$. Since $\mathbf{z}$ interim dominates $\mathbf{x}$, it must be that 1 gets strictly more under $\mathbf{z}$ than under $\mathbf{x}$ in state (t_1', t_2''). In that case, 2 is strictly worse under $\mathbf{z}$ than under $\mathbf{x}$ in that state, and the comparison must reverse in state (t_1'', t_2'') for $\mathbf{z}$ to be interim superior to $\mathbf{x}$. Of course the same reasoning also tells that 1 must be strictly better under $\mathbf{z}$ than under $\mathbf{x}$ in state (t_1'', t_2'). Thus if $\mathbf{x}$ and $\mathbf{z}$ differ in one state, they must differ in all states.

$E[\bar{x}_2|t_2', \mathcal{T}_1']$, and choosing $K'' > K'p(t_1'', t_2'')/p(t_1', t_2'')$ ensures that $E[e_2^*|t_2'', \mathcal{T}_1'] > E[\bar{x}_2|t_2'', \mathcal{T}_1']$. Notice that $\lim_{\varepsilon\to 0}(e_1^*(t_1', t_2')-\underline{u}_1(t_1', t_2'))/\varepsilon = -f_1'((t_1', t_2'), \overline{u}_2(\mathbf{t})) = \infty$ where $\bar{x}_1(t_1', t_2') \leq \underline{u}_1(t_1', t_2')$,
$\lim_{\varepsilon\to 0}(e_1^*(t_1', t_2'')-\bar{x}_1(t_1', t_2''))/\varepsilon = K''f_1'((t_1', t_2''), \bar{x}_2(t_1', t_2'')) > -\infty$, hence for any K'', for sufficiently small ε we have $E[e_1^*|t_1', \mathcal{T}_2'] > E[\bar{x}_1|t_1', \mathcal{T}_2']$. Also notice $\lim_{\varepsilon\to 0}(e_1^*(t_1'', t_2') - \bar{x}_1(t_1'', t_2'))/\varepsilon = Kf_1'((t_1'', t_2'), \bar{x}_2(t_1'', t_2')) > -\infty$, and
$\lim_{\varepsilon\to 0}(e_1^*(t_1'', t_2'') - \bar{x}_1(t_1'', t_2''))/\varepsilon = -K'f_1'((t_1'', t_2''), \bar{x}_2(t_1'', t_2''))$. So, choosing

$$K' > -\frac{Kp(t_1'', t_2')f_1'((t_1'', t_2'), \bar{x}_2(t_1'', t_2'))}{p(t_1'', t_2'')f_1'((t_1'', t_2''), \bar{x}_2(t_1'', t_2''))},$$

we have $E[e_1^*|t_1'', \mathcal{T}_2'] > E[\bar{x}_1|t_1'', \mathcal{T}_2']$ for all sufficiently small ε, we complete the proof. □

Proof of Lemma OA.4 (smooth 2×2 case). If $\mathbf{x}$ is ex-post efficient (but not interim efficient) so $\mathbf{x} = \bar{\mathbf{x}}$ then Lemma OA.2 shows that it not weakly interim efficient. Hence suppose the limit contract $\mathbf{x}$ is not ex-post efficient. By Lemma 2, any limit equilibrium demands $\mathbf{x}$ and $\mathbf{y}$ are weakly ex-post efficient with $x_2(\mathbf{t}) \geq \underline{u}_2(\mathbf{t})$ and $y_1(\mathbf{t}) \geq \underline{u}_1(\mathbf{t})$, while $E[x_i|t_i] = E[y_i|t_i]$. Hence, without loss of generality, assume that $x_1(t_1', t_2') < \underline{u}_1(t_1', t_2') \leq y_1(t_1', t_2')$ where $\mathcal{T}_i = \{t_i', t_i''\}$. Weak ex-post efficiency then implies $x_2(t_1', t_2') = \overline{u}_2(t_1', t_2')$. Smoothness and $\underline{u}_1(t_1', t_2') > 0$ ensures that (SBC) and hence (BC) is satisfied for agent $i = 2$ in state (t_1', t_2'), and so by Lemma 3 we must have $y_2(t_1', t_2') < \overline{u}_2(t_1', t_2') = x_2(t_1', t_2')$. Combined with $E[x_i|t_i] = E[y_i|t_i]$ we must then have $\overline{u}_2(t_1'', t_2') \geq y_2(t_1'', t_2') > x_2(t_1'', t_2')$. Given the weak ex-post efficiency of $\mathbf{x}$ and $x_2(\mathbf{t}) \geq \underline{u}_2(\mathbf{t})$ for all $\mathbf{t}$, the fact that $\overline{u}_2(t_1'', t_2') > x_2(t_1'', t_2')$ implies $\mathbf{x}(t_1', t_2'')$ must be efficient in $\mathcal{U}(t_1', t_2'')$. Similarly, given $y_1(\mathbf{t}) \geq \underline{u}_1(\mathbf{t})$ for all $\mathbf{t}$, we get $y_2(t_1'', t_2') > x_2(t_1'', t_2') \geq \underline{u}_2(t_1'', t_2')$ and so the weak ex-post efficiency of $\mathbf{y}$ implies $\mathbf{y}(t_1', t_2'')$ must be efficient in $\mathcal{U}(t_1', t_2'')$. The efficiency of $\mathbf{y}(t_1', t_2'')$ and $\mathbf{x}(t_1', t_2'')$ in $\mathcal{U}(t_1'', t_2')$ and $y_2(t_1'', t_2') > x_2(t_1'', t_2')$ then implies $y_1(t_1'', t_2') < x_1(t_1'', t_2')$. Combining that fact with $E[x_i|t_i] = E[y_i|t_i]$ we must then have $\overline{u}_1(t_1'', t_2'') \geq y_1(t_1'', t_2'') > x_1(t_1'', t_2'')$. By a similar argument: $x_1(t_1', t_2') < y_1(t_1', t_2')$ and $E[x_i|t_i] = E[y_i|t_i]$ ensure that $x_1(t_1', t_2'') > y_1(t_1', t_2'') \geq \underline{u}_1(t_1', t_2'')$, and so $x_2(t_1', t_2'') < \overline{u}_2(t_1', t_2'')$.

Recall that $\bar{x}_2 = x_2$ and consider the allocation $\mathbf{e}^*$ as defined in the proof of Lemma OA.3: $e_2^*(t_1', t_2') = x_2(t_1', t_2') - \varepsilon$, $e_2^*(t_1'', t_2') = x_2(t_1'', t_2') + K\varepsilon$, $e_2^*(t_1'', t_2'') = x_2(t_1'', t_2'') - K'\varepsilon$, $e_2^*(t_1', t_2'') = x_2(t_1', t_2'') + K''\varepsilon$ and $e_1^*(\mathbf{t}) = f_1(\mathbf{t}, e_2^*(\mathbf{t}))$. As in that proof, we can then choose K, K' and K'' such that for all $\varepsilon > 0$ small e^* is feasible and $E[e_i^*|t_i] > E[\bar{x}_i|t_i]$ given that $f_1'((t_1', t_2'), \overline{u}_2(t_1', t_2')) = 1/f_2'((t_1', t_2'), \underline{u}_1(t_1', t_2')) = -\infty$ where $x_2(t_1', t_2') = \overline{u}_2(t_1', t_2')$ and (SBC) is satisfied for agent $i = 2$ in state (t_1', t_2'). □

Proof of Lemma OA.7 (Properties of $\mathcal{F}$). We only prove that $\mathcal{F}$ is lower hemi-continuous, as other properties are straightforward to check. Let $\mathbf{v}^n \to \mathbf{v}$ be a sequence in $\mathcal{U}$, $\mathbf{u} \in \mathcal{F}(\mathbf{v})$, and $\varepsilon > 0$. We have to show that there exists an integer N large enough that $\mathcal{F}(\mathbf{v}^n)$ intersects the ε ball centered at $\mathbf{u}$, $\mathcal{B}(\mathbf{u}, \varepsilon)$, for all $n \geq N$. Let $\alpha < 1$ be large enough that $\mathbf{v} + \alpha(\mathbf{u} - \mathbf{v}) \in \mathcal{B}(\mathbf{u}, \varepsilon/2)$. Notice that there exists

N' large enough that $\mathbf{w}^n = \mathbf{v}^n + \alpha(\mathbf{u} - \mathbf{v}) \in \mathcal{U}$, for all $n \geq N'$. Indeed, consider $\mathbf{t} \in \mathcal{T}$, and assume that $\mathbf{u}(\mathbf{t}) \neq \mathbf{v}(\mathbf{t})$. If $\mathbf{v}(\mathbf{t}) + \alpha(\mathbf{u}(\mathbf{t}) - \mathbf{v}(\mathbf{t})) \in \text{int}(\mathcal{U}(\mathbf{t}))$, then clearly $\mathbf{w}^n(\mathbf{t}) \in \mathcal{U}(\mathbf{t})$ for all n large enough. Otherwise, the boundary of $\mathcal{U}(\mathbf{t})$ contains both $\mathbf{u}(\mathbf{t})$ and $\mathbf{v}(\mathbf{t})$, and is flat in between. It is easy to check then that $\mathbf{w}^n(\mathbf{t}) \in Conv(\{0, \mathbf{u}(\mathbf{t}), \mathbf{v}(\mathbf{t}), (\underline{u}_1(\mathbf{t}), \overline{u}_2(\mathbf{t})), (\overline{u}_1(\mathbf{t}), \underline{u}_2(\mathbf{t}))\}) \subseteq \mathcal{U}(\mathbf{t})$, for all n large enough, as desired. Next, $E[w_i^n|t_i] = E[v_i^n|t_i] + \alpha(E[u_i|t_i] - E[v_i|t_i]) \geq E[v_i^n|t_i]$, for all i, t_i, since $\mathbf{u} \in \mathcal{F}(\mathbf{v})$. Hence $\mathbf{w}^n \in \mathcal{F}(\mathbf{v}^n)$, for all $n \geq N'$. As $\mathbf{w}^n \to \mathbf{u} + \alpha(\mathbf{u} - \mathbf{v})$, we can find $N \geq N'$ such that $\mathbf{w}^n \in \mathcal{B}(\mathbf{u} + \alpha(\mathbf{u} - \mathbf{v}), \varepsilon/2)$ for all $n \geq N$. Since $\mathbf{u} + \alpha(\mathbf{u} - \mathbf{v}) \in \mathbf{B}(\mathbf{u}, \varepsilon/2)$, $\mathbf{w}^n$ is within distance ε of $\mathbf{u}$, as desired. $\square$

Proof of Lemma OA.8 (Set of interim-efficient contracts is closed). Suppose that $\mathbf{z} \in \mathcal{U}$ is such that $E[z_i|t_i] \geq E[x_i|t_i]$ for all i, t_i, with at least one of the inequalities being strict. Let then $\mathbf{z}^n = \mathbf{x}^n + \alpha(\mathbf{z} - \mathbf{x})$, where α is say 1/2. As established in the proof of the previous lemma, there exists N' large enough that $\mathbf{z}^n \in \mathcal{U}$ for all $n \geq N'$. Notice that $E[\mathbf{z}_i^n|t_i] = E[x_i^n|t_i] + \alpha(E[z_i|t_i] - E[x_i|t_i]) \geq E[x_i^n|t_i]$ for all i, t_i, with at least one of the inequalities being strict. This contradicts the fact that $\mathbf{x}^n$ is interim efficient, which concludes the proof. $\square$

OA.7 More general beliefs

We assumed that types are independent. Although this might seem like a restriction, in fact an even stronger assumption that each player's belief is derived from a common uniform prior (i.e. $p^*(\mathbf{t}|t_i) = 1/|\mathcal{T}_{-i}|$) is without loss of generality. Myerson (1984) highlights that any Bayesian game with state-dependent utility is strategically equivalent to a Bayesian game with such a uniform prior, as probabilities cannot be determined independently of utilities. In fact, this is the justification for the probability invariance axiom. To see this equivalence in our setting, suppose there is a game with states $\mathcal{T}$ and ex-post utility sets $\mathcal{U}$ and the prior probability that agent i's type t_i believes she faces t_{-i} is $p_i(t_{-i}|t_i)$ (which need not be independent of t_i, and potentially cannot be derived from any common prior), and let this game be described by $\mathbf{B} = (\mathcal{T}, \mathcal{U}, \mathbf{p})$. We can now define ex-post utility sets $\mathcal{V}(\mathbf{t}) = \{\mathbf{v} \in \mathbb{R}_+^2 : v_i = K_i(\mathbf{t})u_i(\mathbf{t}), \mathbf{u} \in \mathcal{U}(\mathbf{t})\}$ where $K_i(\mathbf{t}) = |\mathcal{T}_{-i}|p_i(t_{-i}|t_i)$, so any contract $\mathbf{u} \in \mathcal{U}$ corresponds to a contract $\mathbf{v}(\mathbf{u}) \in \mathcal{V} = \times_{\mathbf{t} \in \mathcal{T}} \mathcal{V}(\mathbf{t})$ where $v_i(\mathbf{t}, \mathbf{u}) = K_i(\mathbf{t})u_i(\mathbf{t})$. The bargaining game $\mathbf{B}^* = (\mathcal{T}, \mathcal{V}, \mathbf{p}^*)$ with beliefs derived from the uniform prior, $p^*(\mathbf{t}|t_i) = 1/|\mathcal{T}_{-i}|$, is then strategically identical to game $\mathbf{B} = (\mathcal{T}, \mathcal{U}, \mathbf{p})$. To illustrate this, suppose that agents mix over only a finite set of demands $\underline{\mathcal{U}} \subseteq \mathcal{U}$, and the probability that type t_i demands $\mathbf{u} \in \underline{\mathcal{U}}$ in $\mathbf{B}$ (respectively $\mathbf{v}(\mathbf{u})$ in $\mathbf{B}^*$) is $\sigma_i(t_i, \mathbf{u})$, and the probability that t_i accepts after demands $\mathbf{x}, \mathbf{y}$ in $\mathbf{B}$ (respectively $\mathbf{v}(\mathbf{x}), \mathbf{v}(\mathbf{y})$ in $\mathbf{B}^*$) is $\hat{\sigma}_i(t_i, \mathbf{x}, \mathbf{y})$ then type t_i's payoff is identical in $\mathbf{B}$ and $\mathbf{B}^*$. In particular:

$$\sum_{t_2 \in \mathcal{T}_2} \sum_{\mathbf{x}, \mathbf{y} \in \underline{\mathcal{U}}} \sigma_1(t_1, \mathbf{x})\sigma_2(t_2, \mathbf{y})p_1(t_2|t_1)Q_1(\mathbf{t}) = \sum_{t_2 \in \mathcal{T}_2} \sum_{\mathbf{x}, \mathbf{y} \in \underline{\mathcal{U}}} \sigma_1(t_1, \mathbf{x})\sigma_2(t_2, \mathbf{y})p^*(t_2|t_1)K_1(\mathbf{t})Q_1(\mathbf{t})$$

where agent 1's expected payoff in state $\mathbf{t}$ is $Q_1(\mathbf{t})$ in $\mathbf{B}$ and $K_1(\mathbf{t})Q_1(\mathbf{t})$ in $\mathbf{B}^*$:

$$Q_1(\mathbf{t}) = \hat{\sigma}_1(t_1, \mathbf{x}, \mathbf{y})\hat{\sigma}_2(t_2, \mathbf{x}, \mathbf{y})\frac{x_1(\mathbf{t}) + y_1(\mathbf{t})}{2} + \delta(\hat{\sigma}_1(t_1, \mathbf{x}, \mathbf{y})y_1(\mathbf{t}) + \hat{\sigma}_2(t_2, \mathbf{x}, \mathbf{y})x_1(\mathbf{t})).$$

Fudenberg and Tirole (1991) take advantage of exactly this equivalence between games, in order to extend the reach of their perfect Bayesian equilibrium definition: they transform any game without independent types into one with independent types and then require that beliefs in the transformed game satisfy NSWYDK.

Finally, it is easily verified that (BC), (NLB), (SBC) and smoothness are all preserved under this transformation.

OA.8 Generic interim inefficiency of ex-post Nash

In the text we stated that the ex-post Nash solution is generically not interim efficient. What we mean by this is: if the ex-post Nash solution is interim efficient in some bargaining problem where both players have at least two types, then the solution is not interim efficient when one player's utility is rescaled (in any way) in some state.

A more concrete way to highlight the inefficiency is to specialize to the case of risk averse players with *CRRA* utility functions but players have different coefficients of relative risk aversion, have at least two types, and players divide $\$M(\mathbf{t}) > 0$ in state $\mathbf{t}$. If the ex-post Nash solution is interim efficient, then any change in the money available in some state implies the solution is no longer interim efficient (the simple proof follows similar arguments to the result below and is left to the reader).

Lemma OA. 13. *Suppose for a smooth bargaining problem $\mathbf{B} = (\mathcal{T}, \mathcal{U}, \mathbf{p})$ with $|\mathcal{T}_i| \geq 2$ for $i = 1, 2$ that the ex-post Nash solution $\mathbf{u}^N \in \mathcal{U}$ is interim efficient. Then for any $\mathbf{t}^* \in \mathcal{T}$, and $K \in (0, 1) \cup (1, \infty)$, in the bargaining problem $\tilde{\mathbf{B}} = (\mathcal{T}, \tilde{\mathcal{U}}, \mathbf{p})$, with $\tilde{\mathcal{U}}(\mathbf{t}) = \mathcal{U}(\mathbf{t})$ for $\mathbf{t} \neq \mathbf{t}^*$ and $\tilde{\mathcal{U}}(\mathbf{t}^*) = \{(u_1, Ku_2) : \mathbf{u} \in \mathcal{U}(\mathbf{t}^*)\}$, the ex-post Nash solution is not interim efficient.*

Proof. For the smooth problem $\mathbf{B}$, the ex-post Nash bargaining solution $\mathbf{u}^N$, must satisfy $f_1'(\mathbf{t}, u_2^N(\mathbf{t}))u_2^N(\mathbf{t}) + u_1^N(\mathbf{t}) = 0$ in state $\mathbf{t}$ and have $u_i^N(\mathbf{t}) > \underline{u}_i(\mathbf{t})$. This in turn implies that there is a unique positive unit vector $\mathbf{w}(\mathbf{t})$ which is ex-post orthogonal to $\mathcal{U}(\mathbf{t})$ at $\mathbf{u}^N(\mathbf{t})$, which satisfies $w_2(\mathbf{t})/w_1(\mathbf{t}) = -f_1'(\mathbf{t}, u_2^N(\mathbf{t})) = u_1^N(\mathbf{t})/u_2^N(\mathbf{t})$. Fix $t_1' \neq t_1^*$. The characterization of interim efficiency in Lemma 1 implies there is a unique vector $\hat{\boldsymbol{\lambda}} \in \mathbb{R}_{++}^{\mathcal{T}_1} \times \mathbb{R}_{++}^{\mathcal{T}_2}$ normalized so that $\hat{\lambda}_1(t_1') = 1$, which is interim orthogonal to $\mathcal{U}(\mathbf{B})$ at $\mathbf{u}^N$. Moreover, this interim orthogonal vector must satisfy $\hat{\lambda}_1(t_1)p(t_2)/(\hat{\lambda}_2(t_2)p(t_1)) = w_1(t_1, t_2)/w_2(t_1, t_2)$.

Now consider the ex-post Nash solution $\tilde{\mathbf{u}}^N$ for bargaining problem $\tilde{\mathbf{B}}$. The Nash solution does not change for $\mathbf{t} \neq \mathbf{t}^*$ and so neither do the associated ex-post orthogonal unit vectors, $\tilde{\mathbf{u}}^N(\mathbf{t}) = \mathbf{u}^N(\mathbf{t})$ and $\tilde{\mathbf{w}}(\mathbf{t}) = \mathbf{w}(\mathbf{t})$. This means $\hat{\boldsymbol{\lambda}}$ remains the *unique* vector in $\mathbb{R}_{++}^{\mathcal{T}_1} \times \mathbb{R}_{++}^{\mathcal{T}_2}$ such that $\hat{\lambda}_1(t_1)p(t_2)/(\hat{\lambda}_2(t_2)p(t_1)) = w_1(t_1, t_2)/w_2(t_1, t_2)$ for $(t_1, t_2) \neq \mathbf{t}^*$ and normalized so that $\hat{\lambda}_1(t_1') = 1$. However, the Nash solution in state $\mathbf{t}^*$ must satisfy $\tilde{\mathbf{u}}^N(\mathbf{t}^*) = (u_1^N(\mathbf{t}^*), Ku_2^N(\mathbf{t}^*))$ by invariance. So the unique ex-post orthogonal unit vector $\tilde{\mathbf{w}}(\mathbf{t}^*)$ satisfies $\tilde{w}_2(\mathbf{t}^*)/\tilde{w}_1(\mathbf{t}^*) = u_1^N(\mathbf{t}^*)/Ku_2^N(\mathbf{t}^*) \neq \hat{\lambda}_1(t_1^*)p(t_2^*)/\hat{\lambda}_2(t_2^*)p(t_1^*)$, which implies $\tilde{\mathbf{u}}^N$ cannot be interim efficient. □

OA.9 War of attrition

We show here how our results can be extended to the war of attrition bargaining game outlined in the text. We are interested in characterizing the set of conciliatory equilibria (stationary equilibria with deterministic demands and initial concession on path). Following on path demands in such equilibria, players must accept in every future period (as beliefs must match those in period 1). For the same reasons as in our simple one period model, it is without loss of generality to focus on stationary *pooling* equilibria with initial acceptance (the proof is identical to Proposition 2). The expected outcome of such an equilibrium is $\mathbf{c} = \overline{\delta}(\mathbf{x}+\mathbf{y})/2$ where $\overline{\delta} = \frac{1-\varepsilon}{1-\varepsilon\delta}$ and $\mathbf{x}$ and $\mathbf{y}$ are the pooling demands. Thus given equivalent demands $\mathbf{x}$ and $\mathbf{y}$, payoffs are simply discounted by $\overline{\delta}$ compared to the our one period model.

As noted in the main text, for the war of attrition model we define $\gamma = \delta(1-\varepsilon)/(1-\varepsilon\delta^2)$. We then define best-safe contracts and payoffs exactly as in the main text but using this new γ (e.g. $Y_1^{bs|\mathbf{x}}(\mathbf{t}) = \gamma x_1(\mathbf{t})$ and $Y_2^{bs|\mathbf{x}}(\mathbf{t}) = \max\{u_2 \mid \mathbf{u} \in \mathcal{U},\ u_1 \geq \gamma x_1(\mathbf{t})\}$). The result below then shows that conciliatory equilibrium in the war of attrition are characterized by the same necessary and sufficient conditions as conciliatory equilibrium in our single period model. The proof follows a similar structure to Proposition 3. Given this result, Proposition 5 establishes conditions for convergence to the Myerson solution. We have not extended our sequential equilibrium results to this model.

Proposition OA. 3. *Let $\mathbf{x}, \mathbf{y}$ be contingent contracts in $\mathcal{U}$. There is a conciliatory pooling equilibrium where all types of player 1 propose $\mathbf{x}$, and all types of player 2 propose $\mathbf{y}$, if and only if for all $t_i \in \mathcal{T}_i$ and all $i = 1, 2$:*

$$E[x_i|t_i] \geq E[X_i^{bs|\mathbf{y}}|t_i] \text{ and } E[y_i|t_i] \geq E[Y_i^{bs|\mathbf{x}}|t_i].$$

Proof. To establish the necessity of $E[y_1|t_1] \geq E[Y_1^{bs|\mathbf{x}}|t_1]$ notice that following equilibrium demands $\mathbf{x}$ and $\mathbf{y}$, if player 1's type t_1 rejects in period s and then returns to his equilibrium strategy (of always accepting) he gets:

$$(1-\varepsilon)\delta \sum_{j=1}^{\infty} \left((\varepsilon\delta)^{2j-2} E[x_1|t_1] + (\varepsilon\delta)^{2j-1} E[y_1|t_1]\right) = \frac{(1-\varepsilon)\delta}{1-(\varepsilon\delta)^2} \left(E[x_1|t_1] + \varepsilon\delta E[y_1|t_1]\right).$$

For this to be less than his payoff of $E[y_1|t_1]$ from accepting, we need $E[y_1|t_1] \geq \gamma E[x_1|t_1] = E[Y_1^{bs|\mathbf{x}}|t_1]$. By identical logic $E[x_2|t_2] \geq E[X_2^{bs|\mathbf{y}}|t_2]$

To establish the necessity of $E[y_2|t_2] \geq E[Y_2^{bs|\mathbf{x}}|t_2]$, temporarily suppose that there is a single state of the world so players 1 and 2 make demands $\mathbf{x} \in \mathbb{R}_+^2$ and $\mathbf{y} \in \mathbb{R}_+^2$. Clearly if $x_2 > y_2$ then it cannot be optimal for player 2 to reject $\mathbf{x}$ in any period s, so suppose that $y_2 \geq x_2$, then 2's best possible continuation payoff after rejecting $\mathbf{x}$ clearly requires that 1 always accepts. Let V_2 be player 2's maximum expected utility when he gets to accept in period $r \geq s$ assuming player 1 always accepts then:

$$V_2 = \max\left\{x_2, \frac{\delta(1-\varepsilon)}{1-\delta^2\varepsilon^2}y_2 + \frac{\delta^2\varepsilon(1-\varepsilon)}{1-\delta^2\varepsilon^2}V_2\right\} = \max\{x_2, \gamma y_2\}.$$

If $\gamma y_2 < x_2$, therefore, player 2's maximum possible payoff from rejecting an offer is strictly less than x_2, and so player 2 must certainly accept whenever he gets the chance. Returning now to bargaining problems with multiple states of the world. Player 1 can ensure that player 2 accepts whenever possible when 1 deviates to proposing (arbitrarily close to) his best-safe contract in every state. Player 1's expected payoff from making this deviation and then always accepting $\mathbf{y}$ is $\overline{\delta}(E[X_1^{bs|\mathbf{y}}|t_1] + E[y_1|t_1])/2$ and so we clearly need $E[x_1|t_1] \geq E[X_1^{bs|\mathbf{y}}|t_1]$ for that deviation to be unprofitable. By identical logic, $E[Y_2^{bs|\mathbf{x}}|t_2] \leq E[y_2|t_2]$.

We now turn to establishing sufficiency and so consider two pooling contingent contracts $\mathbf{x}$ and $\mathbf{y}$ satisfying our equilibrium inequalities. After receiving offer $\mathbf{x}$, player 2's updated belief over player 1's type coincides with his interim belief, and acceptance of $\mathbf{x}$ is a best response since $E[x_2|t_2] \geq E[X_2^{bs|\mathbf{y}}|t_2]$, for all $t_2 \in \mathcal{T}_2$. For identical reasons, player 1 optimally accepts $\mathbf{y}$.

We now define beliefs and strategies after a unilateral deviation where player 1 proposed $\mathbf{x}' \neq \mathbf{x}$, but 2 proposed $\mathbf{y}$. Unilateral deviations $\mathbf{y}'$ by player 2 are deterred analogously. As in Proposition 3's proof, let $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) = \{t_1 \in \mathcal{T}_1 : x_2'(t_1, t_2) < \gamma y_2(t_1, t_2)\}$. If $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) \neq \emptyset$, then the probability type t_2 believes he faces type t_1 is $\mu_2(t_1|t_2, \mathbf{x}', \mathbf{y}) = 1$ for some $t_1 \in \mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y})$ and he always rejects $\mathbf{x}'$. If $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) = \emptyset$ then type t_2 believes $\mu_2(t_1|t_2, \mathbf{x}', \mathbf{y}) = 1$ for some arbitrary $t_1 \in \mathcal{T}_1$, and always accepts. Player 1's belief after $\mathbf{y}$ coincides with his interim belief and he always accepts.

We next check that this behavior is sequentially rational. If type t_2 expects that 1 always accepts $\mathbf{y}$, then it is certainly optimal to reject $\mathbf{x}'$ when $\mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) \neq \emptyset$ (as $x_2(t_1, t_2) < \gamma y_2(t_1, t_2)$ for $t_1 \in \mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y})$) and to accept $\mathbf{x}'$ otherwise. To check player 1's incentives, let $\mathcal{T}_2(\mathbf{x}', \mathbf{y}) = \{t_2 \in \mathcal{T}_2 : \mathcal{T}_1(t_2, \mathbf{x}', \mathbf{y}) = \emptyset\}$ be the set of player 2's types who accept $\mathbf{x}'$. Let the probability that type t_1 believes the state is $\mathbf{t}$ in period s following $\mathbf{x}'$ and $\mathbf{y}$ be denoted $\mu_1^s(\mathbf{t}|t_1, \mathbf{x}', \mathbf{y})$. Given player 2's strategy, this satisfies:

$$\mu_1^{2k+d}(\mathcal{T}_2(\mathbf{x}', \mathbf{y})|t_1) = \frac{p(\mathcal{T}_2(\mathbf{x}', \mathbf{y})|t_1)\varepsilon^k}{p(\mathcal{T}_2(\mathbf{x}', \mathbf{y})|t_1)\varepsilon^k + p(\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})|t_1)}$$

where $d \in \{1, 2\}$. This decreases in k. Beliefs about opponent types are:

$$\mu_1^{2k+d}(t_2|t_1) = \mu_1^{2k+d}(\mathcal{T}_2(\mathbf{x}', \mathbf{y})|t_1)p(t_2|\mathcal{T}_2(\mathbf{x}', \mathbf{y}), t_1) + \mu_1^{2k+d}(\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y})|t_1)p(t_2|\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}', \mathbf{y}), t_1).$$

Type t_1's expected payoff from accepting $\mathbf{y}$ in period $s = 2k + d$ is then:

$$\begin{aligned}\mathcal{U}^A_{t_1}(2k+d) =& \mu_1^{2k+d}(\mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1) \sum_{t_2 \in \mathcal{T}_2(\mathbf{x}',\mathbf{y})} p(t_2|\mathcal{T}_2(\mathbf{x}',\mathbf{y}), t_1) y_1(t_1,t_2) \\ &+ \mu_1^{2k+d}(\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1) \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y})} p(t_2|\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y}), t_1) y_1(t_1,t_2).\end{aligned}$$

By contrast his payoff from a one-step deviation of rejecting in period s is:

$$\begin{aligned}\mathcal{U}^R_{t_1}(2k+d) =& \mu_1^{2k+d}(\mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1) \sum_{t_2 \in \mathcal{T}_2(\mathbf{x}',\mathbf{y})} p(t_2|\mathcal{T}_2(\mathbf{x}',\mathbf{y}), t_1) \frac{\delta(1-\varepsilon)}{1-(\varepsilon\delta)^2}(x_1'(t_1,t_2) + \delta\varepsilon y_1(t_1,t_2)) \\ &+ \mu_1^{2k+d}(\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1) \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y})} p(t_2|\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y}), t_1) \frac{\delta^2(1-\varepsilon)}{1-\varepsilon\delta^2} y_1(t_1,t_2)\end{aligned}$$

These payoffs are linear in $\mu_1^{2k+d}(\mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1) = 1 - \mu_1^{2k+d}(\mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1)$, and hence so is their difference. Given that $\mu_1^{2k+d}(\mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1)$ is decreasing in k, it is therefore sufficient to check player 2's incentive to accept when $k = 0$ and when $k \to \infty$. In the latter case, if $\mathcal{T}_2(\mathbf{x}',\mathbf{y}) \neq \mathcal{T}_2$ then $\lim_{k\to\infty} \mu_1^{2k+d}(\mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1) = 0$ and accepting is certainly a best response (as remaining opponents always reject). Of course, if $\mathcal{T}_2(\mathbf{x}',\mathbf{y}) = \mathcal{T}_2$ then beliefs are stationary, and we only need to check the former case of $k = 0$, where $\mu_1^d(\mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1) = p(\mathcal{T}_2(\mathbf{x}',\mathbf{y})|t_1)$. The payoff to rejecting then satisfies:

$$\mathcal{U}^R_{t_1}(d) \leq \frac{\delta(1-\varepsilon)}{1-(\varepsilon\delta)^2} E[X_1^{bs|\mathbf{y}} + \varepsilon\delta y_1|t_1] \leq \frac{\delta(1-\varepsilon)}{1-(\varepsilon\delta)^2} E[x_1 + \varepsilon\delta y_1|t_1],$$

where the first inequality follows from $x_1'(t_1,t_2) \leq X_1^{bs|\mathbf{y}}(t_1,t_2)$ for $t_2 \in \mathcal{T}_2(\mathbf{x}',\mathbf{y})$ and from $X_1^{bs|\mathbf{y}}(\mathbf{t}) \geq y_1(\mathbf{t})$ (so that $y_1(\mathbf{t})\delta^2(1-\varepsilon)/(1-\varepsilon\delta^2) \leq (X_1^{bs|\mathbf{y}}(\mathbf{t})\delta(1-\varepsilon)/(1-\varepsilon^2\delta^2) + \varepsilon\delta y_1(\mathbf{t}))$. The second follows from $E[X_1^{bs|\mathbf{y}}|\mathcal{T}_1] \leq E[x_1|t_1]$. Hence, we have $\mathcal{U}^A_{t_1}(d) = E[y_1|t_1] \geq \mathcal{U}^R_{t_1}(d)$ when:

$$\frac{\delta(1-\varepsilon)}{1-(\varepsilon\delta)^2} E[x_1 + \varepsilon\delta y_1|t_1] \leq E[y_1|t_1],$$

which rearranges to give the (assumed) equilibrium condition $E[y_1|t_1] \geq \gamma E[x_1|t_1]$.

We show deviating to $\mathbf{x}'$ is unprofitable. Type t_1's payoff from this is:

$$\begin{aligned}\mathcal{U}_{t_1} = \frac{(1-\varepsilon)}{2}\Big(&\sum_{t_2 \in \mathcal{T}_2(\mathbf{x}',\mathbf{y})} p(t_2|t_1) \frac{1}{1-\delta\varepsilon}(x_1'(t_1,t_2) + y_1(t_1,t_2)) \\ &+ \sum_{t_2 \in \mathcal{T}_2 \setminus \mathcal{T}_2(\mathbf{x}',\mathbf{y})} p(t_2|t_1) \frac{1+\delta}{1-\delta^2\varepsilon} y_1(t_1,t_2)\Big)\end{aligned}$$

Again, we can bound this from above:

$$\mathcal{U}_{t_1} \leq \frac{1-\varepsilon}{1-\delta\varepsilon} \frac{E[X_1^{bs|\mathbf{y}} + y_1|t_1]}{2} \leq \frac{1-\varepsilon}{1-\delta\varepsilon} \frac{E[x_1 + y_1|t_1]}{2},$$

where the first inequality again follows from $x_1'(t_1,t_2) \le X_1^{bs|\mathbf{y}}(t_1,t_2)$ for $t_2 \in \mathcal{T}_2(\mathbf{x}',\mathbf{y})$ and from $X_1^{bs|\mathbf{y}}(\mathbf{t}) \ge y_1(\mathbf{t})$ (so that $y_1(\mathbf{t})(1+\delta)/(1-\delta^2\varepsilon) \le (X_1^{bs|\mathbf{y}} + y_1(\mathbf{t}))/(1-\delta\varepsilon)$. The second inequality again follows from $E[X_1^{bs|\mathbf{y}}|\mathcal{T}_1] \le E[x_1|t_1]$. The RHS is exactly type t_1's equilibrium payoff of $\overline{\delta}E[x_1+y_1|t_1]/2$, and so the deviation is not profitable.

It remains to ensure there exist mutually optimal, stationary continuation strategies given the players' beliefs after the joint deviation to $\mathbf{x}'$ and $\mathbf{y}'$. We define beliefs consistent with those after unilateral deviations: the probability type t_2's believes that he faces type t_1 is $\mu_2(t_1|t_2,\mathbf{x}',\mathbf{y}') = \mu_2(t_1|t_2,\mathbf{x}',\mathbf{y})$, and similarly, $\mu_1(t_2|t_1,\mathbf{x}',\mathbf{y}') = \mu_1(t_1|t_2,\mathbf{x},\mathbf{y}')$. As these beliefs are degenerate they are not be updated over time. Let $t_j(t_i) = t_j$ if $\mu_i(t_j|t_i,\mathbf{x}',\mathbf{y}') = 1$. Define an auxiliary game with players $\mathcal{T}_1 \cup \mathcal{T}_2$ where type t_i chooses a "mixed" strategy $\sigma_{t_i} \in [0,1]$. Type t_1's expected payoff given σ is:

$$U_{t_1}(\sigma) = \sigma_{t_1} y_1'(t_1,t_2(t_1)) + (1-\sigma_{t_1}) \frac{\delta(1-\varepsilon)\sigma_{t_2(t_1)}}{1-\delta^2(1-(1-\varepsilon)\sigma_{t_2(t_1)})} x_1'(t_1,t_2(t_1))$$

The utility of t_2 is defined similarly. This game has a Nash equilibrium σ^* in "mixed" strategies by standard reasoning (e.g., Kakutani). Type t_i's payoffs are linear in σ_{t_i}, so if $\sigma_{t_i} \in (0,1)$ is a best response, then so is $\sigma_{t_i} \in [0,1]$. Denote by $\sigma_{t_i}^*$ type t_i's stationary acceptance probability in each period of the war of attrition. It is easy to check optimality of $\sigma_{t_i}^*$ in the auxiliary game entails no profitable one-shot deviations in the war of attrition. □

OA.10 Computing Myerson solutions

In this section we will highlight a simple procedure which can be used to easily identify Myerson solutions numerically. It effectively involves solving for the zeros of one-dimensional, continuous monotonic functions, and can be easily executed using standard numerical minimizers available in Matlab, R etc.

We limit attention to smooth bargaining problems satisfying (SBC), and which have a *Non Linear Frontier* (NLF). A smooth bargaining problem has a non-linear frontier if $f_i'(\mathbf{t},.)$ is strictly decreasing on $(\underline{u}_{-i}(\mathbf{t}),\overline{u}_{-i}(\mathbf{t}))$, for all $\mathbf{t}$, $i=1,2$. For such a bargaining problem, for any $\alpha > 0$, there is a unique $\mathbf{u}(\alpha,\mathbf{t}) \in \mathcal{U}(\mathbf{t})$ such that $\mathbf{u}(\alpha,\mathbf{t}) \in \arg\max_{\mathbf{v}\in\mathcal{U}(\mathbf{t})}(\alpha v_1 + v_2)$; it satisfies $-f_2'(u_1(\alpha,\mathbf{t}),\mathbf{t}) = \alpha$.[10] We also assume, without loss of generality as described in OA.7, that there is a uniform prior.

In what follows, we slightly adapt the notation from Section OA.2. For any $\alpha > 0$, there is a unique $\mathbf{u}(\alpha,\mathbf{t}) \in \mathcal{U}(\mathbf{t})$ such that $\mathbf{u}(\alpha,\mathbf{t}) \in \arg\max_{\mathbf{v}\in\mathcal{U}(\mathbf{t})}(\alpha v_1 + v_2)$; it satisfies $-f_2'(u_1(\alpha,\mathbf{t}),\mathbf{t}) = \alpha$. For such an $\alpha > 0$, also let $S_1(\alpha,\mathbf{t}) = \max_{\mathbf{v}\in\mathcal{U}(\mathbf{t})}(\alpha v_1 + v_2)/(2\alpha)$ and $S_1(\alpha,\mathbf{t}) = \alpha S_2(\alpha,\mathbf{t})$ define the mid-point of the associated linearized

[10] Moreover, $\mathbf{u}(.,\mathbf{t})$ is continuous in α, and $u_1(.,\mathbf{t})$ is strictly increasing (conversely $u_2(.,\mathbf{t})$ is strictly decreasing), and approaches $(\overline{u}_1(\mathbf{t}),\underline{u}_2(\mathbf{t}))$ as $\alpha \to \infty$ and $(\underline{u}_1(\mathbf{t}),\overline{u}_2(\mathbf{t}))$ as $\alpha \to 0$.

bargaining set.[11] Bargainer i's excess utility in state $\mathbf{t}$, when comparing her payoff to the midpoint of the linearized bargaining set, is then: $\Delta_i(\mathbf{t},\alpha) = u_i(\alpha,\mathbf{t}) - S_i(\alpha,\mathbf{t})$ for $i = 1,2$,[12] where it is immediately verified that:

$$\alpha\Delta_1(\mathbf{t},\alpha) + \Delta_2(\mathbf{t},\alpha) = 0. \tag{OA.10}$$

For interim weights $\hat{\boldsymbol{\lambda}} \in \mathbb{R}^{\mathcal{T}_1}_{++} \times \mathbb{R}^{\mathcal{T}_2}_{++}$, we then define $\alpha(\mathbf{t},\hat{\boldsymbol{\lambda}}) = \hat{\lambda}_1(t_1)/\hat{\lambda}_2(t_2)$ and $\bar{\Delta}_i(t_i,\hat{\boldsymbol{\lambda}}) = \sum_{\mathbf{t}\in\mathcal{T}(t_i)} \Delta_i(\mathbf{t},\alpha(\mathbf{t},\hat{\boldsymbol{\lambda}}))$. With $\mathbf{t} = (t_i,t_j)$, notice that $\Delta_i(\mathbf{t},\alpha(\mathbf{t},\hat{\boldsymbol{\lambda}}))$ depends only the weight $\lambda_j(t_j)$ for type t_j of bargainer j and the weight $\lambda_i(t_i)$ for bargainer i. Similarly, $\bar{\Delta}_i(\mathbf{t},\boldsymbol{\lambda})$ depends on $\lambda_j(t_j)$ *for all* $t_j \in \mathcal{T}_j$ but only on $\lambda_i(t_i)$ for bargainer i. Thus, in a slight abuse of notation, we will henceforth use $\Delta_i(\mathbf{t},\alpha(\mathbf{t},\hat{\boldsymbol{\lambda}}))$ and $\bar{\Delta}_i(\mathbf{t},\hat{\boldsymbol{\lambda}})$ for subvectors $\hat{\boldsymbol{\lambda}}$ that may not contain all the interim weights, but do specify the weights needed to define the object.

To simplify notation in the description of the computational process and the proof that it identifies the Myerson solution, we present it when $\mathcal{T}_1$ has three states, arbitrarily enumerated as $\mathcal{T}_1 = \{t_1^1, t_1^2, t_1^3\}$ (the cardinality of $\mathcal{T}_2$ is irrelevant for the notation). The generalization of this process and the proof to arbitrary $\mathcal{T}_1$ are both straightforward. The process finds the interim weights $\hat{\lambda}_i^*(t_i)$ for each $t_i \in \mathcal{T}_i$ and each $i = 1,2$ associated with the Myerson solution. Aside from $\hat{\lambda}_1^*(t_1^3) = 1$, which is normalized to equal 1, each interim weight $\hat{\lambda}_i^*(t_i)$ will be chosen as an element in some interval $[\underline{\lambda}_i(t_i), \overline{\lambda}_i(t_i)]$. Defining $g_i(\mathbf{t}) = \Delta_i^{-1}(\mathbf{t}, -\sum_{\mathbf{t}'\in\mathcal{T}(t_i)\setminus\{t\}} \overline{u}_i(\mathbf{t}')/2)$,

$$\underline{\lambda}_2(t_2) = 1/g_2(t_1^3,t_2) \text{ and } \overline{\lambda}_2(t_2) = 1/g_1(t_1^3,t_2)$$

for each $t_2 \in \mathcal{T}_2$,

$$\underline{\lambda}_1(t_1^k) = \max_{t_2\in\mathcal{T}_2} g_1(t_1^k,t_2)/g_2(t_1^3,t_2) \text{ and } \overline{\lambda}_1(t_1^k) = \min_{t_2\in\mathcal{T}_2} g_2(t_1^k,t_2)/g_1(t_1^3,t_2),$$

for each $k < 3$.

Step 0: For any inputted interim weights $(\hat{\lambda}_1(t_1^1), \hat{\lambda}_1(t_1^2)) \in [\underline{\lambda}_1(t_1^1), \overline{\lambda}_1(t_1^1)] \times [\underline{\lambda}_1(t_1^2), \overline{\lambda}_1(t_1^2)]$ and $\hat{\lambda}_1(t_1^3) = 1$, define for each $t_2 \in \mathcal{T}_2$ the function $\check{\lambda}_2(t_2|\cdot)$ that sets

$$\check{\lambda}_2(t_2|\hat{\lambda}_1(t_1^1),\hat{\lambda}_1(t_1^2)) = \arg\min_{\hat{\lambda}_2'(t_2)\in[\underline{\lambda}_2(t_2),\overline{\lambda}_2(t_2)]} \left(\bar{\Delta}_2\left(t_2,\hat{\lambda}_1(t_1^1),\hat{\lambda}_1(t_1^2),\hat{\lambda}_1(t_1^3),\hat{\lambda}_2'(t_2)\right)\right)^2,$$

This function is simple to identify because $\bar{\Delta}_2\big(t_2,\hat{\lambda}_1(t_1^1),\hat{\lambda}_1(t_1^2),\hat{\lambda}_1(t_1^3),\hat{\lambda}_2'(t_2)\big)$ is continuous in the weights and strictly increasing in $\hat{\lambda}_2'(t_2)$. Let $\check{\boldsymbol{\lambda}}^2(\hat{\lambda}_1(t_1^1),\hat{\lambda}_1(t_1^2)) \in \mathbb{R}^{\mathcal{T}_2}_{++}$

[11] Lemma OA.5 shows that $S_1(.,\mathbf{t})$ is decreasing in α and $S_2(.,\mathbf{t})$ is increasing. Moreover, $S_1(\alpha,\mathbf{t}) \to \overline{u}_1(\mathbf{t})/2$ and $S_1(\alpha,\mathbf{t}) \to -\infty$ as $\alpha\to\infty$ (likewise $S_2(\alpha,\mathbf{t}) \to \overline{u}_2(\mathbf{t})/2$ and $S_1(\alpha,\mathbf{t}) \to -\infty$ as $\alpha \to 0$).

[12] For $\mathbf{t} = (t_1,t_2)$, $\Delta_1(\mathbf{t},\alpha)$ is clearly continuous and strictly increasing in α, converging to $\overline{u}_1(\mathbf{t})/2$ as $\alpha\to\infty$ and to $-\infty$ as $\alpha \to 0$ (conversely $\Delta_2(\mathbf{t},\alpha)$ is continuous and strictly decreasing, converging to $-\infty$ and $\overline{u}_2(\mathbf{t})/2$ respectively).

be the *vector* of interim weights for bargainer 2 which uses, for each $t_2 \in \mathcal{T}_2$, the interim weight $\check{\lambda}_2(t_2|\hat{\lambda}_1(t_1^1), \hat{\lambda}_1(t_1^2))$ defined above.

Step 1: For any inputted interim weight $\hat{\lambda}_1(t_1^2) \in [\underline{\lambda}_1(t_1^2), \overline{\lambda}_1(t_1^2)]$, define the function $\check{\lambda}_1(t_1^1|\cdot)$ that sets

$$\check{\lambda}_1(t_1^1|\hat{\lambda}_1(t_1^2)) = \arg \min_{\hat{\lambda}_1'(t_1^1) \in [\underline{\lambda}_1(t_1^1), \overline{\lambda}_1(t_1^1)]} \left(\bar{\Delta}_1 \left(t_1^1, \hat{\lambda}_1'(t_1^1), \check{\boldsymbol{\lambda}}^2(\lambda_1'(t_1^1), \hat{\lambda}_1(t_1^2)) \right) \right)^2 .$$

This is simple to identify because $\bar{\Delta}_1\big(t_1^1, \hat{\lambda}_1'(t_1^1), \check{\boldsymbol{\lambda}}^2(\hat{\lambda}_1'(t_1^1), \hat{\lambda}_1(t_1^2))\big)$ is continuous in the weights and strictly increasing in $\hat{\lambda}_1'(t_1^1)$.

Step 2 (final): Given the initialization $\hat{\lambda}_1(t_1^3) = 1$, select $\check{\lambda}_1(t_1^2)$ by

$$\check{\lambda}_1(t_1^2) = \arg \min_{\hat{\lambda}_1'(t_1^2) \in [\underline{\lambda}_1(t_1^2), \overline{\lambda}_1(t_1^2)]} \left(\bar{\Delta}_1 \left(t_1^2, \hat{\lambda}_1'(t_1^2), \check{\boldsymbol{\lambda}}^2(\check{\lambda}_1(t_1^1|\hat{\lambda}_1'(t_1^2)), \hat{\lambda}_1'(t_1^2)) \right) \right)^2 .$$

This is simple to identify because $\bar{\Delta}_1\big(t_1^2, \hat{\lambda}_1'(t_1^2), \check{\boldsymbol{\lambda}}^2(\check{\lambda}_1(t_1^1|\hat{\lambda}_1'(t_1^2)), \hat{\lambda}_1'(t_1^2))\big)$ is continuous and strictly increasing in $\hat{\lambda}_1'(t_1^1)$. The interim weights in the Myerson solution are then:

$$\hat{\boldsymbol{\lambda}}^* = \left(\check{\lambda}_1(t_1^1|\check{\lambda}_1(t_1^2)), \check{\lambda}_1(t_1^2), 1, \check{\boldsymbol{\lambda}}^2(\check{\lambda}_1(t_1^1|\check{\lambda}_1(t_1^2)), \check{\lambda}_1(t_1^2)) \right),$$

and so, ex-post utilities are $\mathbf{u}(\alpha(\boldsymbol{\lambda}^*, \mathbf{t}), \mathbf{t})$.

In the general case with $\mathcal{T}_1 = \{t_1^1, ..., t_1^K\}$, we normalize $\hat{\lambda}_1(t_1^K) = 1$ and identify $\check{\lambda}_1(t_1^k|.)$ for $k < K$ as the function of $\hat{\boldsymbol{\lambda}}(t_1^{k+1}), ..., \hat{\lambda}_1(t_1^{K-1})$ which minimizes $(\bar{\Delta}(t_1^k, .))^2$ where interim vectors for each t_2 and t_1^m with $m < k$ have been previously identified by the functions $\check{\lambda}_2(t_2|.)$ and $\check{\lambda}_1(t_1^m|.)$. When executing these steps numerically, it is clear that increasing agent 1's types adds more to the computational load than increasing agent 2's types (as we must then execute more nested minimization operations). While using a numerical approach identifies a solution approximately (i.e. with $|\bar{\Delta}_i(t_i, \hat{\boldsymbol{\lambda}})| < \varepsilon \approx 0$ for all i, t_i), it is readily verified by continuity that the identified utilities are arbitrarily close to Myerson's solution when there is sufficient numerical precision (i.e. when ε is small). We next justify the steps above, showing they indeed identify the solution.

Proof of the above claims

We begin by pointing out we can work on the compact set of interim weights above instead of $\mathbb{R}_{++}^{\mathcal{T}_1} \times \mathbb{R}_{++}^{\mathcal{T}_2}$. Notice that if $\Delta_i(\mathbf{t}, \alpha(\mathbf{t}, \hat{\boldsymbol{\lambda}})) < -\sum_{\mathbf{t}' \in \mathcal{T}(t_i)\setminus\{t\}} \overline{u}_i(\mathbf{t}')/2$ then $\bar{\Delta}_i(t_i, \hat{\boldsymbol{\lambda}}) < 0$. This implies that in the Myerson solution, $\alpha(\mathbf{t}, \hat{\boldsymbol{\lambda}}^*) \in [g_1(\mathbf{t}), g_2(\mathbf{t})]$, and so given $\hat{\lambda}_1^*(t_1^3) = 1$, we must have $\hat{\lambda}_i^*(t_i) \in [\underline{\lambda}_i(t_i), \overline{\lambda}_i(t_i)]$. More generally, given $\hat{\lambda}_1(t_1^3) = 1$, notice that if $\hat{\lambda}_i(t_i) = \underline{\lambda}_i(t_i)$ then $\bar{\Delta}_i(t_i, \hat{\boldsymbol{\lambda}}) \leq 0$.

Step 0: The fact that $\bar{\Delta}_2\big(t_2, \hat{\lambda}_1(t_1^1), \hat{\lambda}_1(t_1^2), \hat{\lambda}_1(t_1^3), \hat{\lambda}_2'(t_2)\big)$ is continuous in the weights and strictly increasing in $\hat{\lambda}_2'(t_2)$ is an immediate implication of the properties of $\Delta_i(\mathbf{t}, .)$. Given that and the fact that $\bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}) \leq 0$ when $\hat{\lambda}_2(t_2) = \underline{\lambda}_2(t_2)$, we must have that the value of $\bar{\Delta}_2(t_2,.)$ at $\check{\lambda}_2(t_2|.)$ is weakly negative, and is strictly negative only if $\check{\lambda}_2(t_2|.) = \overline{\lambda}_2(t_2)$. It is immediate the $\check{\lambda}_2(t_2|.)$ is continuous.

Step 1: We first claim that $\bar{\Delta}_1\big(t_1^1, \hat{\lambda}'_1(t_1^1), \check{\boldsymbol{\lambda}}^2(\lambda'_1(t_1^1), \hat{\lambda}_1(t_1^2))\big)$ is strictly increasing in $\hat{\lambda}'_1(t_1^1)$; it is clearly continuous in the interim weights. This is an implication of the fact that $\hat{\lambda}'_1(\mathbf{t}_1^1)/\check{\lambda}_2(t_2|\hat{\lambda}'_1(\mathbf{t}_1^1), \hat{\lambda}_1(t_1^2))$ is strictly increasing in $\hat{\lambda}'_1(t_1^1)$. Suppose not, so that there is some $\hat{\lambda}'_1(t_1^1) > \hat{\lambda}_1(t_1^1)$ and some t_2 with

$$1 < \frac{\hat{\lambda}'_1(t_1^1)}{\hat{\lambda}_1(t_1^1)} \le \frac{\check{\lambda}_2(t'_2|\hat{\lambda}'_1(\mathbf{t}_1^1), \hat{\lambda}_1(t_1^2))}{\check{\lambda}_2(t_2|\hat{\lambda}_1(\mathbf{t}_1^1), \hat{\lambda}_1(t_1^2))}. \tag{OA.11}$$

For that t_2 we must have $\check{\lambda}_2(t_2|\hat{\lambda}_1(\mathbf{t}_1^1), \hat{\lambda}_1(t_1^2)) < \overline{\lambda}_2(t_2)$ and so (as argued in Step 0) $\bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}) = 0 \ge \bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}')$ for the interim vectors given by

$$\hat{\boldsymbol{\lambda}} = \left(\hat{\lambda}_1(t_1^1), \hat{\lambda}_1(t_1^2), \lambda_1(t_1^3), \check{\boldsymbol{\lambda}}^2(\hat{\lambda}_1(\mathbf{t}_1^1), \lambda_1(\mathbf{t}_1^2))\right),$$

$$\hat{\boldsymbol{\lambda}}' = \left(\hat{\lambda}'_1(t_1^1), \hat{\lambda}_1(t_1^2), \lambda_1(t_1^3), \check{\boldsymbol{\lambda}}^2(\hat{\lambda}'_1(\mathbf{t}_1^1), \lambda_1(\mathbf{t}_1^2))\right).$$

But given (OA.11), we must have $\Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}}') \ge \Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}})$ for $\mathbf{t} = (t_1^1, t_2)$, and since $\hat{\lambda}_1(\mathbf{t}_1^m) = \hat{\lambda}'_1(\mathbf{t}_1^m)$ for $m > 1$ we must also have $\Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}}') > \Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}})$ for all $\mathbf{t} = (t_1^m, t_2)$ with $m > 1$ as well. This contradicts $\bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}) = 0 \ge \bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}')$.

As in step 0, it is immediate that $\check{\lambda}_1(t_1^1|.)$ is continuous and that the value of $\bar{\Delta}_1(t_1^1,.)$ at $\check{\lambda}_1(t_1^1|.)$ is weakly negative and can only be strictly negative when $\check{\lambda}_1(t_1^1|.) = \overline{\lambda}_1(t_1^1)$ (since $\bar{\Delta}_1(t_1^1, \hat{\boldsymbol{\lambda}}) \le 0$ when $\hat{\lambda}_1(t_1^1) = \underline{\lambda}_1(t_1^1)$).

Step 2: A nearly analogous argument as in Step 1 shows that the function $\bar{\Delta}_1\big(t_1^2, \check{\lambda}_1(t_1^2|\hat{\lambda}'_1(t_1^2)), \hat{\lambda}'_1(t_1^2), \check{\boldsymbol{\lambda}}^2(\check{\lambda}_1(t_1^1|\hat{\lambda}'_1(t_1^2)), \hat{\lambda}'_1(t_1^2))\big)$ is strictly increasing in $\hat{\lambda}'_1(t_1^1)$; again, it is clearly continuous in the interim weights. Suppose not, then for some t_2 and $\hat{\lambda}'_1(t_1^1) > \hat{\lambda}_1(t_1^1)$ we must have

$$1 < \frac{\hat{\lambda}'_1(t_1^2)}{\hat{\lambda}_1(t_1^2)} \le \frac{\check{\lambda}_2(t_2|\check{\lambda}_1(\mathbf{t}_1^1|\hat{\lambda}'_1(t_1^2)), \hat{\lambda}'_1(t_1^2)}{\check{\lambda}_2(t_2|\check{\lambda}_1(\mathbf{t}_1^1|\hat{\lambda}_1(t_1^2)), \hat{\lambda}_1(t_1^2)} = \max_{t'_2 \in \mathcal{T}_2} \frac{\check{\lambda}_2(t'_2|\check{\lambda}_1(\mathbf{t}_1^1|\hat{\lambda}'_1(t_1^2)), \hat{\lambda}'_1(t_1^2)}{\check{\lambda}_2(t'_2|\check{\lambda}_1(\mathbf{t}_1^1|\hat{\lambda}_1(t_1^2)), \hat{\lambda}_1(t_1^2)}. \tag{OA.12}$$

Letting

$$\hat{\boldsymbol{\lambda}} = \left(\check{\lambda}_1(t_1^1|\hat{\lambda}_1(t_1^2)), \hat{\lambda}_1(t_1^2), \lambda_1(t_1^3), \check{\boldsymbol{\lambda}}^2(t_2|\check{\lambda}_1(\mathbf{t}_1^1|\hat{\boldsymbol{\lambda}}(t_1^2)), \hat{\boldsymbol{\lambda}}(t_1^2))\right)$$
$$\hat{\boldsymbol{\lambda}}' = \left(\check{\lambda}_1(t_1^1|\hat{\lambda}'_1(t_1^2)), \hat{\lambda}'_1(t_1^2), \lambda_1(t_1^3), \check{\boldsymbol{\lambda}}^2(t_2|\check{\lambda}_1(\mathbf{t}_1^1|\hat{\boldsymbol{\lambda}}'(t_1^2)), \hat{\boldsymbol{\lambda}}'(t_1^2))\right),$$

we must have $\Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}}') \ge \Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}})$ for $\mathbf{t} = (t_1^m, t_2)$ and $m \ge 2$, with strict inequality for $m > 2$. The difference with the previous step is that to avoid contradicting $\bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}) = 0 \ge \bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}')$, it must be that $\Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}}') < \Delta_2(\mathbf{t}, \hat{\boldsymbol{\lambda}})$ for $\mathbf{t} = (\mathbf{t}_1^1, t_2)$.[13] This

[13]More broadly, when generalizing beyond three types for bargainer 1, the inequality must hold for some t_1^m where m is smaller than the index of the current type examined.

implies $\hat{\lambda}_1'(t_1^1)/\hat{\lambda}_1(t_1^1) > \hat{\lambda}_2'(t_2)/\hat{\lambda}_2(t_2) > 1$ and $\Delta_1(\mathbf{t}, \hat{\boldsymbol{\lambda}}') > \Delta_1(\mathbf{t}, \hat{\boldsymbol{\lambda}})$. Furthermore, since $\hat{\lambda}_1(t_1^1) < \overline{\lambda}_1(t_1^1)$ we must have $\bar{\Delta}_1(t_1^1, \hat{\boldsymbol{\lambda}}') \leq \bar{\Delta}_1(t_1^1, \hat{\boldsymbol{\lambda}}) = 0$, and so there must exist t_2' with $\Delta_1(\mathbf{t}, \hat{\boldsymbol{\lambda}}') < \Delta_1(\mathbf{t}, \hat{\boldsymbol{\lambda}})$ for $\mathbf{t} = (t_1^1, t_2')$. That is possible, only if $\hat{\lambda}_2'(t_2')/\hat{\lambda}_2(t_2') > \lambda_1'(t_1^1)/\hat{\lambda}_1(t_1^1) > \hat{\lambda}_2'(t_2)/\hat{\lambda}_2(t_2)$. However, this contradicts the equality in (OA.12), which establishes the claim.

As in step 0 and 1, it is immediate that the value of $\bar{\Delta}_1(t_1^2, .)$ at $\check{\lambda}_1(t_1^2)$ is weakly negative and can only be strictly negative if $\check{\lambda}_1(t_1^2) = \overline{\lambda}_1(t_1^2)$ (since $\bar{\Delta}_1(t_1^2, \hat{\boldsymbol{\lambda}}) \leq 0$ when $\hat{\lambda}_1(t_1^2) = \underline{\lambda}_1(t_1^2)$).

Why this leads to Myerson For the vector $\hat{\boldsymbol{\lambda}}^*$ to be interim weights in the Myerson solution, we need $\bar{\Delta}_i(t_i, \hat{\boldsymbol{\lambda}}^*) = 0$ for all t_i. To establish this, first notice that (OA.10) implies

$$\sum_{t_1} \lambda_1^*(t_1)\bar{\Delta}_1(\mathbf{t}, \hat{\boldsymbol{\lambda}}^*) + \sum_{t_2} \lambda_2^*(t_2)\bar{\Delta}_2(\mathbf{t}, \hat{\boldsymbol{\lambda}}^*) = 0.$$

We established above that $\bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}^*) \leq 0$ with a strict inequality only if $\hat{\boldsymbol{\lambda}}_2^*(t_2) = \overline{\lambda}_2(t_2)$ and $\bar{\Delta}_1(t_1^k, \hat{\boldsymbol{\lambda}}^*) \leq 0$ for $k < 3$ with a strict inequality only if $\hat{\boldsymbol{\lambda}}_1^*(t_1^k) = \overline{\lambda}_1(t_1^k)$. If $\hat{\lambda}_2^*(t_2) = \overline{\lambda}_2(t_2)$ for some t_2, however, then $\bar{\Delta}_1(t_1^3, \hat{\boldsymbol{\lambda}}^*) \leq 0$, and so our interim weights would be those of the Myerson solution (all inequalities would need to be equalities). On the other hand suppose that $\hat{\lambda}_2^*(t_2) < \overline{\lambda}_2(t_2)$ for all t_2. If $\hat{\lambda}_1^*(t_1) = \overline{\lambda}_1(t_1)$ for some t_1 then we would have $\bar{\Delta}_2(t_2, \hat{\boldsymbol{\lambda}}^*) < 0$ for some t_2 (by the definition of $\overline{\lambda}_1(t_1)$ and $\hat{\lambda}_2^*(t_2) < \overline{\lambda}_2(t_2)$), which implies $\hat{\lambda}_2^*(t_2) = \overline{\lambda}_2(t_2)$. This is a contradiction, and so we must also have $\hat{\lambda}_1^*(t_1^k) < \overline{\lambda}_1(t_1^k)$ for $k < 1$ and hence $\bar{\Delta}_i(t_i, \hat{\boldsymbol{\lambda}}^*) = 0$ for $t_i \in \mathcal{T}_2 \cup \{t_1^1, t_1^2\}$. Combined with (OA.10), we must then additionally have $\bar{\Delta}_1(t_1^3, \hat{\boldsymbol{\lambda}}^*) = 0$, and so again our interim weights must be those of the Myerson solution.

References

Border, K. C. (1985). *Fixed Point Theorems with Applications to Economics and Game Theory*. Cambridge University Press.

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

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

von Neumann, J. (1937). Über ein ökonomisches Gleichungssystem und eine Verallgemeinerung des Brouwerschen Fixpunktsatzes. Ergebn. math. Kolloqu. Wien 8, 73-83 (1937).