ONLINE APPENDIX FOR *INEFFICIENT CONCESSIONS AND MEDIATION*

Buzard and Horne (2024)

Pooling equilibrium

Lemma 5 *From Round 1 on, playing (Distrust,Distrust) in all rounds is a subgame perfect Nash equilibrium of the continuation game for both types regardless of beliefs.*

Proof: As $D > 0$ and $W > 0$, the dominant stage game action for all types is to play Distrust. Thus, the stage game equilibrium is (Distrust,Distrust), and in any repeated game, playing the stage game equilibrium in each round is an equilibrium of the entire game. ■

That is, one high-type equilibrium strategy is to pool with the low types by always playing Distrust. If both the low and high type of player j choose the strategy that gives no concessions and always plays Distrust, then either type of player i is made strictly better off by playing Distrust in each round given player j's behavior. Given that the stage-game equilibrium will be (Distrust,Distrust), players have no incentive to give a concession in Period 0 because concessions are costly, and signaling one's type can give no benefit in this equilibrium. Types pool by never giving concessions and always playing Distrust.

No-Concessions Separating Equilibrium

We examine the conditions for a separating equilibrium held in place by a grim trigger.

Proof of Proposition 1:

Playing Trust in the first round and thereafter, as long as Trust has always been played, is incentive compatible for high types if the expected payoffs from this strategy are greater than the payoffs from playing Distrust in the first round and then every round thereafter by the one-shot deviation principle. If both countries play Trust in a round, a Cooperation equilibrium takes place, and both countries continue to play Trust in every round with the grim trigger threat of Distrust. Low types play Distrust in the first round, so a high type that observed Distrust being played would respond by playing Distrust in

Round 2 and all future rounds. A country plays Trust if and only if the following condition is satisfied:

$$p\left(\frac{T}{1-\delta}\right) + (1-p)\left(-D + \frac{\delta(W-D)}{1-\delta}\right) \geq p\left(T + W + \frac{\delta}{1-\delta}(W-D)\right) + (1-p)\frac{W-D}{1-\delta}.$$

Simplifying, we have

$$pT \geq p\left((1-\delta)(T+W) + \delta(W-D)\right) + (1-\delta)(1-p)W$$

$$pT \geq pT - p\delta T + pW - p\delta W + p\delta W - p\delta D + W - \delta W - pW + p\delta W$$

$$p\delta T + p\delta D + \delta W - p\delta W \geq W$$

$$\delta \geq \frac{W}{(1-p)W + p(T+D)} = \delta^{nc}. \tag{1}$$

If the potential advantages of Cooperation are large enough to outweigh the risks of encountering a low type in expected utility terms, that is $\delta_h \geq \delta^{nc}$, a county has the incentive to play Trust in the first round rather than Distrust. Any player with $\delta < \delta^{nc}$ has the incentive—by this same inequality—to play Distrust in each round even if the other country plays Trust. Thus, we call these countries low types for purposes of this equilibrium. ■

Concessions Separating Equilibrium

If the condition in Expression 1 is not met, a separating equilibrium cannot be achieved in the absence of concessions in Period 0.[26] If, instead, two negotiating parties have sufficient incentive to give a concession in Period 0 that indicates they are capable of cooperation, these high types could safely play Trust in Round 1, allowing two high types to avoid the trap of the No Cooperation outcome that would occur under pooling.

Below are the details of the *concessions separating equilibrium* analysis. We denote the equilibrium separating gifts as g_h from the high type and g_l from the low type. These equilibrium gifts mean that when Round 1 is reached, countries know which type the other country is and play accordingly. The

[26]Countries need to deliver concessions in Period 0, not just promise them.

concessions phase acts as a coordination device to match countries' actions. For two countries of the same type to play different actions is an off-equilibrium contingency. These off-equilibrium payoffs are nonetheless necessary for calculating the minimum separating concession.

Proof of Lemma 1:

In accord with the Revelation Principle (Myerson 1979), we focus without loss of generality on *concessions separating equilibria* in which countries reveal their types truthfully. This allows high types to only play Trust in Round 1 with other high types and to play Distrust with low types. Low types play Distrust (Lemma 5) and their equilibrium payoff is $U_l = X_{FF} + pg_h - g_l$.

Though low types might be willing to give a non-zero gift in a separating equilibrium, they will only do this if it gives them some advantage. Here, high types only cooperate with other high types, so giving a concession cannot help low types to achieve a higher payoff in the repeated game. Because their concession g_l enters negatively in the payoff function, low types' optimal concession is, therefore, 0.

High types have the incentive to play Trust only in the Cooperation equilibrium, which can only be sustained by a pair of high types. Both types benefit from the other country playing Trust, so all countries have an incentive to signal that they are high types if concessions are costless (cheap talk). Thus, costless announcements cannot lead to truthful revelation. If concessions are to lead to a separating equilibrium, they must be costly. ■

Proof of Proposition 2:

In the separating equilibrium constructed here, players believe that any concession not equal to the equilibrium concession of the high type g_h is a low-type gift. In this case, the binding incentive compatibility constraint is the low type IC constraint. When we set $g_l = 0$ following Lemma 1, the low-type IC constraint is

$$pg_h + X^l_{FF} \geq pX^l_{FT} + (1-p)X^l_{FF} - g_h + pg_h.$$

This expression represents the low types' incentives for truth telling as opposed to posing as a high type and giving the corresponding equilibrium gift. Simplifying this inequality, we see that the high types need to give a gift

$$g_h \geq p(X_{FT}^l - X_{FF}^l) \tag{2}$$

in order to separate from the low type.

High type separating equilibrium utility is $U_h = pX_{TT} + (1-p)X_{FF} - g_h + pg_h$. Because utility is decreasing in g_h, it is optimal for high types to make g_h as small as possible while still achieving separation. This occurs when Expression 2 holds with equality. Substituting in from the definitions of X_{FT}^l and X_{FF}^l yields the minimum separating gift[27]

$$g_h^* = p(T + D). \tag{3}$$

To get the high types to send nonzero concessions, the high type incentive compatibility constraint must be satisfied. That is, the payoff from separating must be higher than the payoff from pooling with the low types:

$$pX_{TT}^h + (1-p)X_{FF}^h - g_h + pg_h \geq X_{FF}^h + pg_h.$$

In terms of fundamentals, we have

$$p\frac{T}{1-\delta_h} + (1-p)\frac{W-D}{1-\delta_h} - g_h \geq \frac{W-D}{1-\delta_h}. \tag{4}$$

Combining Expressions 3 and 4, we have the condition for a *concessions separating equilibrium* to exist:

$$\delta_h \geq \frac{W}{T+D} = \delta^c. \tag{5}$$

That is, we need the high types to have a sufficiently large discount factor. ■

[27]Although not formally addressed in this paper, inference about sizes of concessions when parties are of unequal force can be made. Note that the constraints determining the requisite size of the separating gift depend on the low types' ability to gather war spoils. Thus, a more powerful country would be more able to plunder and hence has to make a greater concession to convincingly convey that it is a high type.

Proof of Corollary 1:

To show that the patience threshold for the *no-concessions separating equilibrium* is greater than the patience threshold for the *concessions separating equilibrium*, we start with the assumption $T + D > W$.

We multiply both sides by $(1 - p) > 0$ and then subtract $p(T + D)$ from both sides to get

$$(T + D) > (1 - p)W + p(T + D).$$

Because $0 < p < 1$ and W, T, and D are all positive, both sides of the inequality are positive. Thus, we have

$$\delta^{nc} = \frac{W}{(1 - p)W + p(T + D)} > \frac{W}{T + D} = \delta^{c}.$$

In the case that both types of equilibria exist, i.e., $\delta_h \geq \delta^{nc}$, the high types prefer the *no-concessions separating equilibrium* if and only if

$$p\left(\frac{T}{1 - \delta_h}\right) + (1 - p)\left(-D + \frac{\delta_h(W - D)}{1 - \delta_h}\right) \geq p\left(\frac{T}{1 - \delta_h}\right) + (1 - p)\left(\frac{W - D}{1 - \delta_h}\right) - (1 - p)g_h.$$

This simplifies to preferring the *concessions separating equilibrium* if and only if

$$0 \geq (1 - \delta_h)W - p(1 - \delta_h)(T + D).$$

Rearranging, we have

$$p \geq \frac{W}{T + D}. \tag{6}$$

That is, the high types prefer the *concessions separating equilibrium* when p is smaller than $\frac{W}{T+D}$. ■

Concessions with Material Value

Proof of Proposition 3:

In the separating equilibrium constructed in Proposition 3, players believe that any concession not equal to the equilibrium concession of the high type g_h is a low-type gift. In this case, the binding incentive compatibility constraint is the low type IC constraint. Taking account of the result of Lemma 1 that

$g_l = 0$, we have

$$pg_h + X_{FF}^l \geq pX_{FT}^l + (1-p)X_{FF}^l - g_h + pg_h. \tag{7}$$

On the left side, the low type does not give a concession but receives a concession if facing a high type (pg_h). Then (Distrust,Distrust) is played, and the payoffs depend on whether this low type faces a high or a low type because the implications for future material value are different. With probability p, the low type receives a gift and invests in the military. With probability $(1-p)$, neither side gives concessions. The left side of Expression 7 is then $pg_h + \frac{1}{1-\delta_l}\left[p(W + Wg_h - D) + (1-p)(W-D)\right]$.

On the right side, the low type gives and receives a gift from the high type with probability p. With probability p, this low type faces a high type who invests the gift in civil society. With probability $(1-p)$, the low type's gift goes to another low type who invests in the military, and the low type does not receive a gift. So the right hand side of Expression 7 becomes $-g_h + pg_h + p\left(T + Tg_h + W + Wg_h + \frac{\delta_l}{1-\delta_l}(W + Wg_h - D)\right) + (1-p)\left(\frac{1}{1-\delta_l}(W - D - Dg_h)\right)$.

Substituting into Expression 7 and simplifying, we have

$$g_h + p(W + Wg_h - D) \geq p\left(T + Tg_h + W + Wg_h\right) + (1-p)\left(\frac{-Dg_h}{1-\delta_l}\right)$$

$$(1-\delta_l)\, g_h\, (1-pT) + (1-p)Dg_h \geq (1-\delta_l)\, p\, (D+T)\,.$$

We have assumed that $(1-\delta_l)(1-pT) + (1-p)D > 0$ so we can divide to get

$$g_h \geq \frac{(1-\delta_l)\, p\, (D+T)}{(1-\delta_l)(1-pT) + (1-p)D} = g^1 \tag{8}$$

This implies that the high types need to give a gift at least as large as g^1 to separate from the low type; that is, if the high types give a gift smaller than g^1, a low type has the incentive to mimic the high type and the separating equilibrium is destroyed.

High type separating equilibrium utility is $U_h = pX_{TT} + (1-p)X_{FF} - g_h + pg_h$. Because utility is decreasing in g_h, it is optimal for high types to make g_h as small as possible and still have separation. This occurs when Expression 8 holds with equality. That is, the optimal high type gift in this equilibrium is $g^1 = \frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}$.

The high-type incentive compatibility constraint must not be violated for high types to send nonzero concessions. That is, the payoff from separating must be higher than the payoff from not giving a gift and pooling with the low types. Note that if the high type pools with the low type, it knows it will play Distrust and so invests any concession received in military buildup. Therefore, the high-type IC constraint is

$$p\frac{T+Tg^1}{1-\delta_h}+(1-p)\frac{W-D-Dg^1}{1-\delta_h}-(1-p)g^1 \geq p\frac{W+Wg^1-D}{1-\delta_h}+(1-p)\frac{W-D}{1-\delta_h}+pg^1.$$

Simplifying, we have

$$p(T+Tg^1)+(1-p)(W-D-Dg^1))-(1-\delta_h)g^1 \geq p(W+Wg^1-D)+(1-p)(W-D)$$

$$\delta_h g^1 \geq p(W+Wg^1-D)-pT(1+g^1)+(1-p)Dg^1+g^1$$

$$\delta_h \geq \frac{p(W-D-T)}{g^1}+p(W-T)+(1-p)D+1.$$

Substituting the minimum efficient concession from Expression 8, we have the condition for a *concessions separating equilibrium* to exist in terms of fundamentals:

$$\delta_h \geq \frac{p(W-D-T)}{\frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}}+p(W-T)+(1-p)D+1=\delta^1. \tag{9}$$

That is, there is no parameter restriction beyond our definition of what it means for a player to be a high type. ■

Comparative statics

We start by examining the gifts in a *concessions separating equilibrium*, first with no material value (or destroyed material value) and then when concessions have future material value.

It is direct that the gift in the concessions separating equilibrium with no material value (or destroyed material value) increases in p because the gift is equal to $p(T+D)$. The gift also clearly increases in both T and D.

It is straightforward that the gift in the model with material value, $g^1 = \frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}$, increases in p. Recall that all the variables are assumed to be strictly positive, with $0 < \delta_l < 1$. It follows that the numerator increases in p and the denominator decreases in p. Recall that we assume the denominator is positive. Therefore, g^1 increases in p. In formal terms, $\frac{\partial g^1}{\partial p} = \frac{(1-\delta_l+D)(T+D)}{[(1-\delta_l)(1-pT)+(1-p)D]^2} > 0$.

g^1 does not changed with W because W does not enter the expressing anywhere. The other three parameters require working out the formal result.

For the patience level of the low type:

$$\frac{\partial g^1}{\partial \delta_l} = \frac{((1-\delta_l)(1-pT)+(1-p)D)(-p)(T+D)-(1-\delta_l)\,p\,(D+T)(-1)(1-pT)}{[(1-\delta_l)(1-pT)+(1-p)D]^2}$$

$$= \frac{-((1-\delta_l)(1-pT)\,p(T+D)+(1-p)D)\,p(T+D)+(1-\delta_l)\,p\,(D+T)(1-pT)}{[(1-\delta_l)(1-pT)+(1-p)D]^2}$$

$$= \frac{(1-p)pD(T+D)}{[(1-\delta_l)(1-pT)+(1-p)D]^2}.$$

Every term in the numerator is positive, and the denominator is a squared term, so it is positive. Therefore we have $\frac{\partial g^1}{\partial \delta_l} > 0$.

For the benefit from negotiating partner playing Trust:

$$\frac{\partial g^1}{\partial T} = \frac{((1-\delta_l)(1-pT)+(1-p)D)(1-\delta_l)p-(1-\delta_l)\,p\,(D+T)(\delta_l p-p)}{[(1-\delta_l)(1-pT)+(1-p)D]^2}$$

$$= \frac{[(1-\delta_l)p]\,\{(1-\delta_l)(1-pT+pD+pT)+(1-p)D\}}{[(1-\delta_l)(1-pT)+(1-p)D]^2}$$

$$= \frac{[(1-\delta_l)p]\,\{(1-\delta_l)(1+pD)+(1-p)D\}}{[(1-\delta_l)(1-pT)+(1-p)D]^2}.$$

Every term in the numerator is positive, and the denominator is a squared term, so it is positive. Therefore we have $\frac{\partial g^1}{\partial T} > 0$.

For the damages from negotiating partner playing Distrust:

$$\frac{\partial g^1}{\partial D} = \frac{((1-\delta_l)(1-pT)+(1-p)D)(1-\delta_l)p-(1-\delta_l)\,p\,(D+T)(1-p)}{[(1-\delta_l)(1-pT)+(1-p)D]^2}$$

$$= \frac{(1-\delta_l)p\,\{(1-\delta_l)(1-pT)+(1-p)D-(D+T)(1-p)\}}{[(1-\delta_l)(1-pT)+(1-p)D]^2}$$

$$= \frac{(1-\delta_l)p\,\{(1-\delta_l)\,(1-pT) - T(1-p)\}}{[(1-\delta_l)\,(1-pT) + (1-p)D]^2}$$

$$= \frac{(1-\delta_l)p\,\{1-\delta_l - pT + \delta_l pT - T + pT\}}{[(1-\delta_l)\,(1-pT) + (1-p)D]^2}$$

$$= \frac{(1-\delta_l)p\,\{1-\delta_l + \delta_l pT - T\}}{[(1-\delta_l)\,(1-pT) + (1-p)D]^2}.$$

Because both the denominator and $(1-\delta_l)p$ are strictly positive, the sign of this derivative matches the sign of $1-\delta_l+\delta_l pT - T$. It is positive when $\frac{1-\delta_l}{1-\delta_l p} > T$ and negative otherwise. So, when the benefit from the partner playing Trust is sufficiently small, the gift size increases in D. However, otherwise, the gift size decreases in the damages from the negotiating partner playing Distrust.

We now turn to the patience threshold that defines the high type. In the concessions separating equilibrium with no material value, we have $\frac{W}{T+D} = \delta^c$, which increases in W and decreases in T and D. Notice it is not a function of δ_l or p.

The results for the patience threshold are more complex when concessions have future material value. First, recall that $\delta^1 = \frac{p(W-D-T)}{\frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}} + p(W-T) + (1-p)D + 1$ where the denominator of the first term is simply g^1.

As in the case of no material value, the patience threshold increases in W: $\frac{\partial \delta^1}{\partial W} = \frac{p}{g^1} + p > 0$ because both p and g^1 are strictly positive.

The patience threshold increases in p. That is,

$$\frac{\partial \delta^1}{\partial p} = \frac{g^1(W-D-T) - p(W-D-T)\frac{\partial g^1}{\partial p}}{\left(g^1\right)^2} + W - T - D$$

$$= (W-D-T)\left\{\frac{g^1 - p\frac{\partial g^1}{\partial p}}{\left(g^1\right)^2} + 1\right\} = \frac{W-D-T}{(g^1)^2}\left\{g^1 - p\frac{\partial g^1}{\partial p} + (g^1)^2\right\}$$

$$= \frac{W-D-T}{(g^1)^2}\,\{p(T+D)\,[-\delta_l(1-\delta_l) - \delta_l D - \delta_l pD(1-\delta_l)]\} > 0.$$

We assume that $T > W - D$, so the fraction's numerator is positive, while the denominator is positive because it's a squared term. $p(T+D)$ is positive because all three variables are positive. Each term in the square brackets is negative. Therefore, the full expression is positive.

The patience threshold also increases in the low type's patience level δ_l. That is,

$$\frac{\partial \delta^1}{\partial \delta_l} = \frac{g^1 \cdot 0 + p(T + D - W)\frac{\partial g^1}{\partial \delta_l}}{\left(g^1\right)^2} = \frac{p(T + D - W)\frac{\partial g^1}{\partial \delta_l}}{\left(g^1\right)^2} > 0.$$

The result holds because each term in the expression is positive.

In contrast, the patience threshold decreases in the benefit from the negotiating partner playing Trust, T. That is,

$$\frac{\partial \delta^1}{\partial T} = \frac{-p \cdot g^1 + p(T + D - W)\frac{\partial g^1}{\partial T}}{(g^1)^2} - p = \frac{p}{(g^1)^2}\left\{p(T + D - W)\frac{\partial g^1}{\partial T} - g^1 - (g^1)^2\right\}$$

$$= \frac{p}{(g^1)^2}\{-(1 - \delta_l)W - (1 - \delta_l P)WD - 2TD(1 - \delta_l)p\} < 0.$$

The result holds because p and g^1 are positive, while each additive term in the curly braces is negative.

The result is more nuanced for D, the cost of the negotiating partner playing Distrust. The change in the patience threshold when D changes is

$$\frac{\partial \delta^1}{\partial D} = \frac{-p \cdot g^1 + p(T + D - W)\frac{\partial g^1}{\partial D}}{(g^1)^2} + (1 - p).$$

If $T < \frac{1-\delta_l}{1-\delta_l p}$ so that $\frac{\partial g^1}{\partial D} > 0$, this expression is always negative. As T grows so that $\frac{\partial g^1}{\partial D} < 0$, the derivative remains negative as long as $T < \frac{1-\delta_l+\delta_l D(1-p)}{1-\delta_l}$. When T grows even larger, the derivative remains negative as long as $W < \frac{(1-p)\delta_l(T+D)^2}{T-1+\delta_l-\delta_l pT}$.

Material Value of Concessions can be Destroyed

Proof of Lemma 2:

The expected separation utility for the high type in the game with material value that can be destroyed is

$$p\frac{T + Teg_h}{1 - \delta_h} + (1 - p)\frac{W - D - Deg_h}{1 - \delta_h} - (1 - p)g_h.$$

The derivative with respect to e is

$$\frac{1}{1-\delta_h}\left\{pTg_h-(1-p)Dg_h\right\}.$$

This is non-negative when $pT \geq (1-p)D$. High-type welfare is thus non-decreasing in e when this condition is met. Thus, the largest admissible value of e, i.e., $e = 1$, thus maximizes welfare when $pT \geq (1-p)D$.[28] ■

Proof of Lemma 3:

The proof is by contradiction. Suppose the patience threshold for the *concessions separating equilibrium with inefficient gifts* is weakly larger than the patience threshold for the *concessions separating equilibrium with efficient gifts*. Because low types have patience levels below the threshold, we have that

$$\delta^c = \frac{W}{T+D} \geq \frac{p(W-D-T)}{\frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}} + p(W-T)+(1-p)D+1 = \delta^1 > \delta_l. \tag{10}$$

It follows directly that $W > \delta_l(T+D)$. Because $\delta_l < 1$, it is also true that $\frac{1-p}{1-\delta_l p} < 1$. Combining this with the previous fact, we have

$$W > \delta_l(T+D) > \frac{\delta_l(T+D)(1-p)}{1-\delta_l p}. \tag{11}$$

We invoke Expression 11 below. First we go back to Expression 10, simplifying and creating a common denominator for δ^1, we have

$$\frac{W}{T+D} \geq \frac{(W-(D+T))(1-\delta_l)-(W-(D+T))(1-\delta_l)pT+(W-(D+T))(1-p)D}{(1-\delta_l)(D+T)}$$
$$+\frac{p(W-T)(1-\delta_l)(D+T)+(1-p)D(1-\delta_l)(D+T)+(1-\delta_l)(D+T)}{(1-\delta_l)(D+T)}$$

[28] If $pT = (1-p)D$, $e = 1$ is not the unique maximizer but is still a maximizer.

Subtract $\frac{W}{T+D}$ from both sides to get

$$0 \geq \frac{-W(1-\delta_l) + W(1-\delta_l) - (D+T)(1-\delta_l) - (W-(D+T))(1-\delta_l)\, pT}{(1-\delta_l)(D+T)}$$
$$+\frac{(W-(D+T))(1-p)D + p(W-T)(1-\delta_l)(D+T) + (1-p)D(1-\delta_l)(D+T) + (1-\delta_l)(D+T)}{(1-\delta_l)(D+T)}$$

Canceling the first two terms as well as the third and last terms and then expanding and canceling the terms with pT as well as $(1-p)D\,(D+T)$, we have

$$0 \geq \frac{-W(1-\delta_l)\,pT + W(1-p)D + pW(1-\delta_l)(D+T)}{(1-\delta_l)(D+T)} - \frac{\delta_l(1-p)D\,(D+T)}{(1-\delta_l)(D+T)}$$

Expanding once more and then canceling out the $W(1-\delta_l)\,pT$ and pWD terms

$$0 \geq \frac{WD - \delta_l pWD}{(1-\delta_l)(D+T)} - \frac{\delta_l(1-p)D\,(D+T)}{(1-\delta_l)(D+T)}$$

Combining this with the inequality in Expression 11, we arrive at

$$0 \geq \frac{WD(1-\delta_l p)}{(1-\delta_l)(D+T)} - \frac{\delta_l(1-p)D\,(D+T)}{(1-\delta_l)(D+T)} > \frac{\frac{\delta_l(T+D)(1-p)}{1-\delta_l p}\left[D(1-\delta_l p)\right]}{(1-\delta_l)(D+T)} - \frac{\delta_l(1-p)D\,(D+T)}{(1-\delta_l)(D+T)}$$

To simplify, we have

$$0 > \frac{\delta_l(T+D)(1-p)D}{(1-\delta_l)(D+T)} - \frac{\delta_l(1-p)D\,(D+T)}{(1-\delta_l)(D+T)} = 0$$

We have arrived at a contradiction from Expression 10, and therefore, it must be that

$$\frac{p(W-D-T)}{\frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}} + p(W-T) + (1-p)D + 1 > \frac{W}{D+T} \tag{12}$$

■

Mediation

Proof of Lemma 4:

We show that the low type's incentive compatibility constraint cannot hold when the mediator specifies fully efficient gifts. The low type's incentive compatibility constraint in the mechanism M is

$$X^l_{FF} \geq pX^l_{FT} + (1-p)X^l_{FF} - pg_h + peg_h.$$

Here we specify that the mediator chooses $e = 1$, which means that gift-giving and receiving cancel out. The left-hand side has zero gifts because the low-type player under consideration has truthfully revealed itself to be a low type, and the mediator only specifies strictly positive gifts if two countries report that they are High types. Likewise, the $(1-p)X^l_{FF}$ term on the right does not contain a gift because the negotiating partner has made a report of Low. This leaves

$$X^l_{FF} \geq X^l_{FT}.$$

Expanding, we have

$$\frac{W-D}{1-\delta_l} \geq T + (\cdot) + W + (\cdot) + \frac{\delta_l}{1-\delta_l}\left(W + (\cdot) - D - (\cdot)\right)$$

The left-hand side has no gifts, so the payoffs from Table 1 apply. The right-hand side has placeholders $(\cdot)$ for the impact of the material value of gifts. On the right-hand side, we only have cases where both parties declare themselves to be high types. Both parties are instructed to give efficient gifts. The negotiating partner really *is* a high type and thus invests entirely in civil society because they expect to play the Cooperation equilibrium. The low-type player knows it will defect and be in a No Cooperation equilibrium and so invests entirely in the military.

$$W - D \geq (1-\delta_l)\left[T + Tg_h + W + Wg_h\right] + \delta_l\left(W + Wg_h - D\right) \tag{13}$$

Canceling like terms and rearranging, we have

$$0 \geq (1-\delta_l)\,T + (1-\delta_l)\,Tg_h + Wg_h + (1-\delta_l)\,D$$

Because we assume throughout that gifts must be non-negative, each of T, W and D are strictly positive and $0 < \delta_l < 1$, this inequality can never hold. ■

Proof of Proposition 5:

(a) The low type's individual rationality constraint is satisfied trivially: the low type gets the No Cooperation payoff forever outside the mechanism; it also gets the No Cooperation payoff forever *inside* the mechanism while not giving or receiving any concession.

For the general form of the low type incentive compatibility constraint, we start from Expression 13 and add back in the e's that were removed when assuming $e = 1$ in the proof of Lemma 4. So, we start from

$$W - D \geq (1-\delta_l)\,[T + Teg_h + W + Weg_h] + \delta_l\,(W + Weg_h - D) + peg_h - pg_h.$$

Simplifying, we have

$$g_h\,[p\,(1-e) - (1-\delta_l)eT - eW] \geq (T - \delta_l T) + (1-\delta_l)\,D.$$

The right-hand side is positive. If the left-hand side were negative, we would have a negative upper bound on the size of the gift. Given the requirement that the gift be non-negative, the low type's incentive compatibility constraint cannot be satisfied when the left-hand side is negative, and we must have that $p > e\,(p + (1-\delta_l)\,T + W)$. Using this to isolate g_h on the left-hand side and gathering terms, we have

$$g_h \geq \frac{(1-\delta_l)\,(T+D)}{p - e\,(p + (1-\delta_l)T - W)}. \tag{14}$$

If Condition 14 is met, the mechanism is incentive compatible for the low type.

(b) The high-type incentive compatibility constraint is

$$pX_{TT}^h + (1-p)X_{FF}^h - pg_h + peg_h \geq X_{FF}^h. \tag{15}$$

It is always optimal for the high type to invest in civil society when they truthfully reveal. When a high type misrepresents itself as a low type, the mechanism specifies that gifts are exchanged only when both countries are high types. This means that there are no concessions when a high type lies. This implies that Expression 15 is also the individual rationality constraint because outside the mechanism they get the No Cooperation payoff and exchange no concessions.

Thus the constraint for both incentive compatibility and individual rationality is

$$\frac{p}{1-\delta_h}\left(T+Teg_h\right)-pg_h+peg_h+\frac{1-p}{1-\delta_h}\left(W-D\right)\geq\frac{1}{1-\delta_h}\left(W-D\right).$$

This simplifies to

$$\frac{1}{1-\delta_h}\left(T+Teg_h\right)-g_h+eg_h\geq\frac{1}{1-\delta_h}\left(W-D\right)$$

$$T+Teg_h-(1-\delta_h)(1-e)g_h\geq W-D$$

$$T+D-W\geq\left[Te-(1-\delta_h)(1-e)\right]g_h.$$

If $Te<(1-\delta_h)(1-e)$, this is a negative lower bound on the size of the concession and thus implies no additional restriction. If $Te=(1-\delta_h)(1-e)$, there is again no restriction because $T+D>W$ by assumption. If $Te>(1-\delta_h)(1-e)$, the mechanism is both individually rational and incentive compatible for the high type when $\frac{T+D-W}{Te-(1-\delta_h)(1-e)}>g_h$. ■

Proof of Proposition 6:

An example suffices to prove each part.

(a) Examine the parameterization in Corollary 2. $\delta_h=0.7$ and the threshold for high types to separate through concessions when the future material value of concessions is destroyed is the same as when concessions have no future material value, i.e., 0.5. By Corollary 2, the *concessions separating equilibrium with inefficient gifts* is possible. Corollary 2 also showed that the *concessions separating equilibrium with efficient gifts* is not possible because the relevant patience threshold of 0.979 is too high.

Using the mediation mechanism M with $e=0.29$ and $g_h=12.98$, we satisfy both the low type constraint that $g_h \geq \frac{(1-\delta_l)(T+D)}{p-e(p+(1-\delta_l)T+W)} = 9.5890411$ and the high type constraint $g_h \leq \frac{T+D-W}{Te-(1-\delta_h)(1-e)} = 12.987013$. Welfare is 5.0768, compared to 2.82 in the *concessions separating equilibrium with inefficient gifts*.

(b) Let $T = 1$, $W = 1$ and $D = 1$. The assumption that $T > W - D$ is satisfied, and the threshold for high types to separate through concessions when the future material value of concessions is destroyed is $\delta = \frac{1}{1+1} = .5$. Let $\delta_h = \frac{15}{16} = 0.9375$ and $\delta_l = .5$ and $p = .5$. By Proposition 2, the *concessions separating equilibrium with efficient gifts* is possible.

By Proposition 3(a), the patience threshold to separate through concessions with efficient concessions is 0.75. Using Lemma 3, we see that both the *concessions separating equilibrium with efficient gifts* and the *concessions separating equilibrium with inefficient gifts* are possible. Using the mediation mechanism M with $e = 0.18$ and $g_h = 7.7669$, we satisfy both the low type constraint the $g_h \geq \frac{(1-\delta_l)(T+D)}{p-e(p+(1-\delta_l)T+W)} = 7.143$ and the high type constraint $g_h \leq \frac{T+D-W}{Te-(1-\delta_h)(1-e)} = 7.7669$. Welfare is 12.82, compared to 7.5 in the *concessions separating equilibrium with inefficient gifts* and 7.67 in the *concessions separating equilibrium with inefficient gifts*.

(c) Let $T = 1$, $W = 1$ and $D = 1$. The assumption that $T > W - D$ is satisfied, and the threshold for high types to separate through concessions when concessions have no future material value is $\delta = \frac{1}{1+1} = 0.5$. Let $\delta_h = 0.49$ and $\delta_l = 0.25$ and $p = \frac{15}{16} = 0.9375$.

We use Proposition 2(a) to find that the patience threshold to separate through concessions with inefficient concessions is 0.5. This, combined with Lemma 3, shows that neither the *concessions separating equilibrium with efficient gifts* nor the *concessions separating equilibrium with inefficient gifts* is possible.

Using the mediation mechanism M with $e = 0$ and $g_h = 1.6$, we satisfy both the low type constraint $g_h \geq \frac{(1-\delta_l)(T+D)}{p-e(p+(1-\delta_l)T+W)} = 1.6$ and there is no high type constraint because $0 = Te < (1 - \delta_h)(1 - e) = 0.0625$. Welfare is approximately 0.24, compared to 0 outside the mechanism.

■