1 Make mediator work

November 29-30, 2020: In making revisions in April 2020 to respond to the EJPE referee reports, I discovered the IC constraint does not hold for the low type in Section 5: Mediation. Here I will see what adjustments might make the mediator work.

1.1 IC constraints for Mediation

- High type IC:

$$pX^h_{TT} + (1-p)X^h_{FF} - pg_h + peg_h \geq X^h_{FF}$$

  - Relative to Section 4.2, $-g_h$ there becomes $-pg_h$ here (again, only give gift if matched with high type). Also $+peg_h$ in Section 4.2 disappears on RHS.

- Low type IC:

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

  - If a low type reveals truthfully: you get X_{FF} and no gift with probability p; you get the same with probability $(1-p)$ (you need both countries to be high types in order to do anything)
  - If a low type lies: get X_{FT} and give gift (g_h) and receive gift (eg_h) with probability p; X_{FF} again with probability $(1-p)$
  - Relative to Section 4.2, there is no peg_h on the LHS (in Section 4.2, it cancels out with the peg_h on the RHS): no benefit of getting the gift from the high type. Also, $-g_h$ on the RHS becomes $-pg_h$ since you only send the gift if matched with a high type

1.2 Why does low type IC not hold when $e = 1$?

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

- With our restrictions (mostly the linear assumption, IIRC), High type gives gift with efficiency $e = 1$, so gift giving and receiving cancels out. This leaves

$$X^l_{FF} \geq pX^l_{FT} + X^l_{FF} - pX^l_{FF}$$

which simplifies to

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

Expanding, we have

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

Now we have to fill in all the complicated bits in parentheses from Table 3 in Section 4.2

$$\begin{aligned} W\left(m_1 + \left(1-\alpha_1\right)eg_2\right) - D\left(m_2 + \left(1-\alpha_2\right)eg_1\right) \geq \\ \left(1-\delta^l\right)\left[T\left(s_2 + \alpha_2 e g_1\right) + W\left(m_1 + \left(1-\alpha_1\right)eg_2\right)\right] \\ + \delta^l\left(W\left(m_1 + \left(1-\alpha_1\right)eg_2\right) - D\left(m_2 + \left(1-\alpha_2\right)eg_1\right)\right) \end{aligned}$$

We've already assumed $e = 1$ here as given in Lemma 2 for a concessions separating equilibrium

- **Mediator could instruct them to give $e < 1$ gifts without having to disturb the rest of the theory (I think). This I'll come back to.**
- We *should* assume s_1, s_2, m_1, m_2 to all be 1 to be consistent with the theory though. If we implement all that AND $e = 1$, we get

$$\begin{aligned} W\left(1 + \left(1-\alpha_1\right)g_2\right) - D\left(1 + \left(1-\alpha_2\right)g_1\right) \geq \left(1-\delta^l\right)\left[T\left(1 + \alpha_2 g_1\right) + W\left(1 + \left(1-\alpha_1\right)g_2\right)\right] \\ + \delta^l\left(W\left(1 + \left(1-\alpha_1\right)g_2\right) - D\left(1 + \left(1-\alpha_2\right)g_1\right)\right) \end{aligned}$$

- We also know that there are no gifts exchanged for X^l_{FF}, so $g_1 = g_2 = 0$ on the LHS. On the RHS, gifts ARE exchanged because player 2 is a high type and the low type of player 1 lied. So we can set $g_1 = g_2 = g_M$

$$\begin{aligned} W - D \geq \left(1-\delta^l\right)\left[T\left(1 + \alpha_2 g_M\right) + W\left(1 + \left(1-\alpha_1\right)g_M\right)\right] \\ + \delta^l\left(W\left(1 + \left(1-\alpha_1\right)g_M\right) - D\left(1 + \left(1-\alpha_2\right)g_M\right)\right) \end{aligned}$$

- Now for the α's. The low type of player 1 sets $\alpha_1 = 0$ (this is the amount that goes to civil society). Player 2 is a high type here, so sets $\alpha_2 = 1$.

$$W - D \geq \left(1-\delta^l\right)\left[T\left(1+g_M\right) + W\left(1+g_M\right)\right] + \delta^l\left(W\left(1+g_M\right) - D\right)$$

- Distributing, we have

$$W - D \geq T + Tg_M + W + Wg_M - \delta^l T - \delta^l T g_M - \delta^l W - \delta^l W g_M + \delta^l W + \delta^l W g_M - \delta^l D$$

- Canceling terms:

$$-D \geq T + Tg_M + Wg_M - \delta^l T - \delta^l Tg_M - \delta^l D$$

- Rearranging:

$$0 \geq \left(1 - \delta^l\right) T + \left(1 - \delta^l\right) Tg_M + Wg_M + \left(1 - \delta^l\right) D$$

  Assuming $g_M > 0$, this inequality can never hold since each of $T, W, D \geq 0$ and $0 \leq \left(1 - \delta^l\right) \leq 1$

  * What would it mean to have negative gifts?

2 Develop result: mediator increases potential for peace

June 6, 2017

- i.e. want to show mediator makes (Trust, Trust) an equilibrium over a larger parameter space than without mediator

What do we know?

- Theorem 4: Under some parameters, optimal concessions aren't made when there isn't trust
  - This is still a separating equilibrium, but completely inefficient concessions are given. So welfare is lower because of inefficient concessions.
  - Where there is a possibility of mediation helping to achieve peace where it otherwise would not be attainable (not just improving welfare) is if this reduction in welfare due to the inefficient concessions means that for some parameters it's not worth separating so we don't get peace at all without the mediator.
    * i.e. when welfare for high type from separating falls below that for pooling, separating is no longer an equilibrium (convo with Jean-Guillaume).
- Theorem 5: Mediator eliminates inefficient concessions
  - Here, we're already in that parameter space where there *are* inefficient concessions

Need to start by assuming that $\delta_h <$ threshold where separating with no concessions works.

- **Check to see whether Theorem 4 proof has to change.**

When $e = 1$ and no material value, $g_h^* = p(T + D)$ is the smallest concession for CSE.
When $e = 1$ and concessions have material value, $g' = \frac{p(1-\delta_l)(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}$ is the minimum separating concession [Theorem 3] I claim in Theorem 5 that the minimum separating concession under mediation is $g^* = \frac{g'}{1-p}$.

With $\alpha = 1$,

- SWOC (separating without concessions) when $\delta_h \geq \frac{W}{(1-p)W+p(T+D)}$
- CSE (concessions separating equilibrium) when $\delta_h \geq \frac{W}{T+D}$

With $\alpha \in [0, 1)$ and $e = 1$,

- SWOC doesn't change because there are no concessions
- WTS there are parameters where can't get STC outcome because of $\alpha < 1$
  - $e < 1$ means some of these *can* get to peace
  - but some can't under *any* value of e [assuming e chosen at same time as g]

Remember that decision between equilibria is made not by what *can* be achieved in terms of δ, but by best response, i.e. "Does high type do better by not giving concession?"

- Should check in my numerical examples

Intermediate step for Theorem 4

- Low type IC constraint to solve for equilibrium high type gift
- Already having shown $g_l = 0$, can set $g_h = g$ for simplicity of notation
- Also, since all the discount factors are for the low type, I'll let $\delta_l = \delta$
- For the case where $e = 1$:

IC constraint:

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

Expanded in basic terms (without efficiency issues)

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

Multiply through by $(1-\delta)$

$$W - D \geq p(1-\delta)[T+W] + p\delta(W-D) + (1-p)[W-D] - (1-\delta)g$$

Now add complexity from Table 2

$$\begin{aligned} &W(1+(1-\alpha_1)g_2) - D(1+(1-\alpha_2)g_1) \geq \\ &\quad p(1-\delta)[T(1+\alpha_2 g_1) + W(1+(1-\alpha_1)g_2)] + p\delta(W(1+(1-\alpha_1)g_2) - D(1+(1-\alpha_2)g_1)) \\ &\quad\quad + (1-p)[W(1+(1-\alpha_1)g_2) - D(1+(1-\alpha_2)g_1)] - (1-\delta)g \end{aligned}$$

Substitute in $\alpha_1 = 0$ everywhere since this is the IC for a low-type of player 1:

$$\begin{aligned} &W(1+g_2) - D(1+(1-\alpha_2)g_1) \geq \\ &\quad p(1-\delta)[T(1+\alpha_2 g_1) + W(1+g_2)] + p\delta(W(1+g_2) - D(1+(1-\alpha_2)g_1)) \\ &\quad\quad + (1-p)[W(1+g_2) - D(1+(1-\alpha_2)g_1)] - (1-\delta)g \end{aligned}$$

Set $g_1 = 0$ on the LHS and $g_1 = g$ on the RHS:

$$\begin{aligned} &W(1+g_2) - D \geq \\ &\quad p(1-\delta)[T(1+\alpha_2 g) + W(1+g_2)] + p\delta(W(1+g_2) - D(1+(1-\alpha_2)g)) \\ &\quad\quad + (1-p)[W(1+g_2) - D(1+(1-\alpha_2)g)] - (1-\delta)g \end{aligned}$$

Set $g_2 = 0$ and $\alpha_2 = 0$ wherever there is a $(1-p)$ and $g_2 = g$ and $\alpha_2 = 1$ wherever there is a p:

$$\begin{aligned} &pW(1+g) + (1-p)W - D \geq \\ &\quad p(1-\delta)[T(1+g) + W(1+g)] + p\delta(W(1+g) - D) \\ &\quad\quad + (1-p)[W - D(1+g)] - (1-\delta)g \end{aligned}$$

Expand

$$\begin{aligned} &pW + pWg + W - pW - D \geq \\ &\quad (p-p\delta)[T + Tg + W + Wg] + p\delta(W + Wg - D) \\ &\quad\quad + (1-p)[W - D - Dg] - (1-\delta)g \end{aligned}$$

Cancel some like terms and move $(1-\delta)\,g$ to LHS

$$\begin{aligned}(1-\delta)\,g + pWg + W - D \geq \\ &p\,[T + Tg + W + Wg] - p\delta\,[T + Tg] - p\delta D \\ &\qquad + (1-p)\,[W - D - Dg]\end{aligned}$$

Do some more canceling and expanding

$$\begin{aligned}(1-\delta)\,g + W - D \geq \\ &p\,[T + Tg + W] - p\delta T - p\delta Tg - p\delta D \\ &\qquad + W - D - Dg - pW + pD + pDg\end{aligned}$$

$$(1-\delta)\,g \geq pT + pTg + pW - p\delta T - p\delta Tg - p\delta D - Dg - pW + pD + pDg$$

$$(1-\delta)\,g \geq pT + pTg - p\delta T - p\delta Tg - p\delta D - Dg + pD + pDg$$

Now just rearrange to get all the g terms on the left

$$(1-\delta)\,g - pTg + p\delta Tg + Dg - pDg \geq pT - p\delta T - p\delta D + pD$$

$$[(1-\delta) - p(1-\delta)T + (1-p)D]\,g \geq p(1-\delta)(T+D)$$

$$g \geq \frac{p(1-\delta)(T+D)}{(1-\delta) - p(1-\delta)T + (1-p)D}$$

In comparing to the minimum separating gift when $e = 0$, which is

$$g^* \geq \frac{p(1-\delta)(T+D)}{(1-\delta)}$$

We can simplify to see that the minimum gift when $e = 1$ is larger when

$$(1-\delta)pT > (1-p)D$$

Next we get the threshold for δ_h that is necessary for a concessions separating eqm to exist:

- Use the high-type IC constraint:

$$pX^h_{TT} + (1-p)X^h_{FF} - g_h \geq X^h_{FF}$$

- We'll need to expand for material effects but then solve for both $e = 0$ and $e = 1$

- For $e = 1$:

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

$$pT + pTg + (1-p)(W - D - Dg) - g_h(1-\delta_h) \geq W - D$$

$$pT + pTg + W - D - Dg - pW + pD + pDg - g_h + g_h\delta_h \geq W - D$$

$$pT + pTg - Dg - pW + pD + pDg - g_h + g_h\delta_h \geq 0$$

$$pT + pTg - Dg - pW + pD + pDg - g_h + g_h\delta_h \geq 0$$

  Note that all the g's should be g_h's in the five lines above:

$$g_h\delta_h \geq -pT - pTg_h + Dg_h + pW - pD - pDg_h + g_h$$

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

- For $e = 0$, you're back in the original case with no material value, i.e. enforceable for same δ_h as no material value case

$$\frac{p}{1-\delta_h}T + \frac{1-p}{1-\delta_h}(W - D) - g_0 \geq \frac{1}{1-\delta_h}(W - D)$$

$$pT + (1-p)(W - D) - (1-\delta_h)g_0 \geq (W - D)$$

$$pT - p(W - D) + \delta_h g_0 \geq g_0$$

$$\delta_h g_0 \geq g_0 + p(W - D - T)$$

$$\delta_h \geq 1 + \frac{p(W - D - T)}{g_0}$$

  Optimal separating gift in this setting is $g_0 = p(D + T)$. Substituting this in, we have:

$$\delta_h \geq 1 + \frac{p(W - D - T)}{p(D + T)} = 1 + \frac{pW - p(D - T)}{p(D + T)} = 1 + \frac{pW}{p(D + T)} - 1 = \frac{W}{D + T}$$

  - But check: when $e = 0$, also wipe out immediate monetary value? YES, in terms of benefit, but NOT cost, which is the only part that doesn't cancel out of IC constraints (same for high type and low type; benefit is on both sides; cost is on only one side)
    * Not from utility standpoint, but from IC standpoint because direct benefit of gift cancels out

- But high type utility goes down

– Interestingly, if $e = 0$, if high types would have separated without material value, they will here with material value that is destroyed (they don't do better in pooling). We see this by calculating the high type utility for the case where $e = 0$, which is

$$\frac{1}{1-\delta_h}\left[pT + (1-p)(W-D)\right] - g_0$$

Substituting in the value of the g_0 (the optimal separating gift with $e = 0$, which I have as $p(T+D)$ but I should check again to make sure it's correct for this situation) and the threshold value to separate when $e = 0$ (i.e. $\delta = \frac{W}{T+D}$), we have

$$\frac{T+D}{T+D-W}\left[pT + (1-p)(W-D)\right] - p(T+D)$$

$$\frac{T+D}{T+D-W}\left[p(T+D-W) + (W-D)\right] - p(T+D)$$

$$p(T+D) + \frac{T+D}{T+D-W}\left[W-D\right] - p(T+D)$$

$$\frac{T+D}{T+D-W}\left[W-D\right]$$

This is exactly the payoff to the pooling equilibrium (that is, both types pool on fight with no gift giving) when $\delta = \frac{W}{T+D}$, since $U_{pooling} = \frac{W-D}{1-\delta}$.
So $U_{g_0} = U_{pooling}$ at the cutoff for separating, and U_{g_0} is larger for any discount factor higher than the cutoff.

  * This can be seen, for instance, by decomposing $U_{pooling}$ along the same lines as U_{g_0} above into components for the cooperation vs. punishment and cost of the gift (even though that's not what's going on in pooling). The whole thing increases, including the cost of the gift portion for the pooling when δ increases. But the cost of the gift does not scale up in U_{g_0}.

We can also do this more generally. Take

$$U_{g_0} = \frac{1}{1-\delta}(W-D)$$

$$U_{pooling} = \frac{1}{1-\delta_h}\left[pT + (1-p)(W-D)\right] - p(T+D)$$

$$= pT + pD + \frac{\delta}{1-\delta}(pT+pD) + \frac{1}{1-\delta}(W-D-pW) - pT - pD$$

$$= \frac{\delta}{1-\delta}(pT+pD) + \frac{1}{1-\delta}(W-pW-D)$$

Now compare the two, $U_{g_0} \lesseqgtr U_{pooling}$:

$$\frac{\delta}{1-\delta}(pT+pD)+\frac{1}{1-\delta}(W-pW-D) \lesseqgtr \frac{1}{1-\delta}(W-D)$$
$$\frac{\delta}{1-\delta}(pT+pD)-\frac{1}{1-\delta}(pW) \lesseqgtr 0$$
$$\delta(pT+pD) \lesseqgtr pW$$
$$\delta \lesseqgtr \frac{pW}{pT+pD}=\frac{W}{T+D}$$

So that the utility from U_{g_0} is greater exactly when the discount factor is above the threshold for concessions separating equilibrium to be possible without material value.

How does utility compare under $e=1$ and $e=0$? Compare the two, $U_{g_1} \lesseqgtr U_{g_0}$:

$$\frac{p}{1-\delta_h}T(1+g)+\frac{1-p}{1-\delta_h}(W-D(1+g))-g_h \lesseqgtr \frac{1}{1-\delta_h}\left[pT+(1-p)(W-D)\right]-g_0$$
$$pT(1+g_1)+(1-p)(W-D(1+g_1))-(1-\delta_h)g_1 \lesseqgtr pT+(1-p)(W-D)-(1-\delta_h)g_0$$
$$pTg_1+(1-p)(W-D-Dg_1)-(1-\delta_h)g_1 \lesseqgtr (1-p)(W-D)-(1-\delta_h)g_0$$
$$pTg_1-(1-p)Dg_1-(1-\delta_h)g_1 \lesseqgtr -(1-\delta_h)g_0$$
$$pTg_1-(1-p)Dg_1 \lesseqgtr (1-\delta_h)g_1-(1-\delta_h)g_0$$

$$pT-(1-p)D \lesseqgtr \frac{(1-\delta_h)(g_1-g_0)}{g_1} \tag{1}$$

We know (from above) that the minimum gift when $e=1$ is larger than the minimum gift when $e=0$ when

$$(1-\delta)pT > (1-p)D$$

Let's case this out.

1. Assume $g_0=g_1$; then $(1-\delta)pT=(1-p)D$. The right hand side of 1 is 0 and the left hand side is

   $$pT-(1-\delta)pT=pT-pT+\delta pT=\delta pT$$

   And utility under $e=1$ is higher.

2. When $g_0>g_1$; then $(1-\delta)pT<(1-p)D$. The right hand side of 1 is negative and the left hand side is larger (more positive) than in case 1, so utility under $e=1$ is also higher in this case.

3. When $g_1 > g_0$; then $(1-\delta)pT > (1-p)D$. The right hand side of 1 becomes positive and the left hand side becomes smaller (less positive, and possibly negative) than in case 1, so utility under $e = 1$ is higher when g_1 is only a little larger than g_0, but eventually as g_1 gets much larger than g_0, utility under $e = 0$ becomes larger than utility under $e = 1$.

3 To-do list after Venice CesIfo 06/15/17

- Jim Fearon: maybe get rid of repeated game and just parameterize a one-shot game (doens't work because concessions required for separation are too large; require future to recoup cost)
  - he has a paper about concessions being used against you; he also gave me the reference for another one but I've forgotten it

Notes from CPEG, Oct. 25, 2019 presentation:

- Arnaud
  - δ are not exogenous (look at Brexit)
  - Proposes a one-shot game to replace the repeated game. Simple bayesian game with difference preferences depening on type.
- Raphael Godefroy
  - Can δ_i be a function of g? That is, likelihood of defection decreases as concession gets better

4 Review 1 from EJPE (Email from Toke Aidt, August 19, 2018)

While clearly these are interesting issues, I have difficulties reading this paper. Everything is presented in a very unclear manner. Game-theoretical concepts are invoked, but typically not correctly. For instance, an equilibrium is a strategy profile (one strategy for each type of each player). Hence it does not make sense to say, as the authors do in Theorem 1 on page 10, that something is an equilibrium strategy for one type without also specifying what others do. As another example, on page 18, the authors say that "The revelation principle is used to attain truthful self-identification." But in reality, the revelation principle is just an argument for why it is without loss of generality to restrict attention to direct mechanisms. The truthfulness of messages is guaranteed by requiring the mechanism to be incentive compatible, something the authors do not address.

Another thing. It is not considered a good idea to introduce new assumptions, which have not been mentioned before, in the proof of a proposition, as happens in the alleged proof of Theorem 4 on page 27.