

## 0.1 Review of "Inefficient Concessions and Mediation-JCR-21-0096"

The paper considers a mechanism design approach to reduce conflict in trust situations. The basic idea is that opportunities to end conflict go unrealized because nations don't trust the other side to not exploit them. The authors introduce the idea of gifts or concessions that nations can give in order to signal that they are the type that can cooperate. The authors then allow the gift to be used to strengthen a nation, such that by giving the gift a nation will be worse off in future conflict. Finally the authors introduce a mediator and show that a mediator can facilitate cooperation under conditions that nations can not reach cooperation by themselves.

The basic model is in the flavor of an infinitely repeated prisoners dilemma. Both sides are better off if they can agree to cooperate rather than fight but each side has a myopic incentive to fight. The authors assume that there are two types of nations: patience and impatient. The patience type prefers to cooperate forever rather than exploit the other side and remain at war forever. The impatient type prefers to exploit. Cooperation is not possible if there is an impatient type. The strategic problem for nations is that don't know if they can trust the other side (ie is the other side patient or not).

The solution is for nations to give a gift. The basic equilibrium is that patient types are willing to give a gift to the other side that signals that they are patient. If both sides give the gift then they learn that they are both of the cooperative type and in equilibrium they then cooperate indefinitely.

Theorem 3 deals with the cases that the gift can be converted into future military power. For instance, I give you some of my guns, then you might turn around and use them to shoot me. Not surprisingly, the conditions for cooperation are harder to establish when gifts are convertible.

Corollary 2 simply says that under some circumstances the nations could start cooperation via a separating strategy but when gift can be converted to future force then such cooperation is impossible. **This argument is straightforward; what might be nice to know is some comparative statics.** For instance as  $p$  changes how does the relative parameter space for which cooperative equilibria exist in case vary. Such comparative statics might identify the conditions under which the threat of using gift as future weapons creates an especial problem.

My understanding of the game is that nations with convertible resources is that nations always pick to either consume gift completely ( $\alpha = 1$ ) if they are a high type or invest them completely in the conflict ( $\alpha = 0$ ) if they are a low type. The choice is made at the beginning of the game. However, **I am interested to know what changes if nations can vary their use of the gift over time.** For instance, Israel might give the Golan Heights to Syria. In the initial period Syria uses it for housing. However, if conflict starts then it would switch how it uses the Golan Heights. Presumably this results in a different constraint on the equilibrium condition.

Theorem 4 deals with inefficient gifts. That is to say, given the problem

created by a gift being usable in future conflict, nations might want to make a costly offer to signal their type but that the signal does not increase the other side's strength. For instance, a nation might simply burn money. Doing so signals that it is a high type but does not give the other side money with which to buy future weapon. This setting slightly reduces the problem of gifts being used as future resources for conflict.

The meat of the paper is to introduce a mediator in the form of a mechanism design problem. The idea is that the nations decide whether to participate with the mediator and if they do so they must be incentivized to honestly reveal whether they are high or low types. If both sides say they are high then M tells each side how much to transfer and how efficiently to do so (ie. how much of the gift is burnt rather than possibly empower the other side).

I have concerns about this mechanism, M, section. The idea is that each nation sends a report to the mechanism (High or Low) and then based on these inputs, M says what to do. In particular M specifies a gift size and efficiency. **My concern is that the nations are bound to follow the mandated concessions.** This seems to be a massive cop out. If the mechanism proposed concessions and then in equilibrium the nations wanted to follow the recommendations then that would be a far more impressive set of results. **Given the substantive application, the assumption that the nations must do what they are told seems inconsistent with national sovereignty.** We don't expect a mediator to invade one of the nations to enforce its recommendation. The results seem to lack subgame perfection in the final stage. In a domestic context with contracts I might buy the assumption but in the international setting, such a commitment to follow the required concessions seems untenable. It is the phrase "The countries are bound to send the mandated concessions" that bothers me.

The incentive compatibility conditions state a size of concession that a nation will be told to give a certain sized gift if both sides say 'high' such that low types don't want to say high and high types want to say high. These conditions seem a key first step, but in the context of an international setting, the constraints only work because if told to give a concession then the nation must do so. **Why is it that a low type does not say 'high' and then when told to give a concession does not do so and instead just consumes the concession made by the other side.** **The paper would be very impressive if the mechanism was designed such that all types wanted to follow the instruction that they were given by the mediator.** In the current design, nations follow the instructions because they are compelled. The mechanism works ex ante, in that nations must decide voluntarily whether to participate but the mechanism does not work without the compulsion. In the substantive setting of international relations it is hard to see how such compulsion can be credible.

I have mixed feelings about this paper. It is extremely clear. The model is setup well. The math is very clear and there is the right amount of technical material in the main text so the reader can see what is going on without having to see every messy detail. The results are clear and presented in a straightforward way. However, at the end of the day, I find the main findings about the mechanism design is a bit let down because all the results are driven off

the nations being compelled to do what the mediator says. If the authors had a mechanism in which the nations wanted to report type honestly and then voluntarily comply with the mediator's instructions then the paper would be magnificent. Such a setting would be consistent with national sovereignty. In the substantive setting I don't believe that the results are useful unless they involve voluntary compliance.