doi:xx.xxxx/xxxxx

Inefficient Concessions and Mediation

KRISTY BUZARD AND BEN HORNE *Syracuse University, University of California, San Diego*

When two parties are engaged in conflict and each distrusts the other's ability to cooperate or make peace, concessions can be used to indicate an interest in cooperation or peacemaking. However, when negotiating parties are concerned that concessions could be used against them in the future, a lack of trust can prevent optimal concessions from being made. It can also reduce the possibility of peace or cooperation. Using a repeated game that is preceded by an opportunity to signal one's commitment to cooperation through the provision of concessions, we formally demonstrate that concerns over the future use of concessions can explain the existence of inefficient concessions. We then use mechanism design to explore ways a third-party mediator can act as a guarantor that promised concessions would be delivered, thereby reducing inefficiencies and increasing the potential for peace and cooperation. In this process, we open up a new rationale for mediation: to increase the efficiency of signaling in a preliminary round of negotiations and to overcome the concern that concessions will be used against the giver in the future.

Word Count:

13889 (errors:4)

In negotiations with its neighbors, Israel is hesitant to give certain concessions. The Golan Heights, captured from Syria during the 1967 War, is a particular point of contention in Israeli-Syrian relations. While the Golan Heights does not have the West Bank's historical or religious importance, Israel values

This is a manuscript submitted for review.

it for its highly strategic location. The Golan Heights could be a useful concession, but Israel senses that were she to give over the land, Syria may ultimately use the concession against her. This same hesitancy exists with the West Bank settlements. Many are located on hilltops that could offer strategic military outposts for forces hostile to Israel. When she occasionally abandons settlements, Israel generally dismantles all infrastructure, including water and electricity, and bulldozes the buildings. Although inefficient, this reduces the possibility of those settlements being used against Israeli interests. Without trust between negotiating parties, the most efficient concessions often are often difficult to arrange.

In both active wars and long-simmering conflicts, settlement negotiations often occur simultaneously with continued fighting. This is also true in non-violent disputes, such as long-running episodes of geo-strategic competition: parties engage in costly conflict while negotiating, foregoing the value of the public goods they could create through cooperation. If settlement can be achieved, additional costs of conflict – including additional death and destruction – are averted. Third-party mediators often participate in such negotiations, but the literature remains divided regarding how and under what conditions mediation can be effective (Bercovitch 2000; Greig 005a; Beardsley et al. 2006; Duursma 2020). This paper sheds new light on the productive role mediators can play by examining their effect on the efficiency of the concessions that conflict participants can offer. In our context, the efficiency of a concession is determined by how much of the concession's material value is intact when the recipient receives it. If the party who gives the concession destroys any of its material value, the concession is deemed to be less than fully efficient.

We introduce a two-sided signaling model followed by a repeated Prisoner's Dilemma in which concessions can provide future value to the recipient and/or serve as a signal that provides the recipient with information about the type of player the conceding state is. We use this model to show that concerns surrounding the possibility that concessions could be used against the giver in the future can undermine cooperation. Using a mechanism design framework, we demonstrate that a mediator can facilitate more efficient concessions by removing uncertainty about the ability of the parties to commit to a settlement. Perhaps more importantly, the mediator can also facilitate settlement sometimes when the negotiating parties cannot achieve it through bilateral interaction. In other words, commitment is essentially a far-sighted preference that allows cooperation to emerge, uncertainty can result in missed opportunities for cooperation, and concessions and mediation can facilitate conflict settlement and the provision of public goods.

Two archetypal situations illustrate the significance and difficulty of efficient concession-making in dispute settlement. The first is the concession of strategically valuable territory in a military dispute, and the second is the export of advanced technology to a strategic rival.

Control of strategically valuable territory is often contested in both civil and interstate wars. Even if transferring control of strategically valuable territory creates the best chance for peace, its transfer often involves a substantial risk that the territory will provide a military advantage in a future conflict. Following the 1992 war between Armenia and Azerbaijan over the disputed region of Ngorno-Karabakh, Armenia retained control of a significant amount of Azerbaijani territory that neither the Armenian government nor the Karabakh government viewed to be part of its homeland. Azerbaijan, however, valued control of the territory greatly as the home of Azeri residents and wished to govern the territory. Thus, it might have been efficient in the mid-1990s for the Armenian/Karabakh side to cede control of that territory to Azerbaijan as part of a deal in which Azerbaijan recognized Karabakh's independence. However, that territory was strategically valuable, including territory on both sides of the Lachin Corridor, which connects Ngorno-Karabakh and Armenia. Azerbaijani control of the territory surrounding the corridor would empower Azerbaijan to seize military advantage in the future and cut off land travel between Ngorno-Karabakh and Armenia. The territory was not conceded, and the conflict persisted in a costly stalemate for almost twenty years until war broke out anew in 2020.

In our second archetypal example, the US has restricted the export of cutting-edge technology to China under both the Trump and Biden administrations. China and the US are engaged in geo-strategic competition. Both countries could benefit from the more rapid innovation and growth that free trade in chips and other technology would enable. However, there is a risk that any technology exported will be used in the military sector to pursue advanced weapons, altering the military balance of power. As of 2023, trade in numerous advanced technologies is restricted, reducing the global pace of innovation and economic growth.

In both archetypal examples, cooperative/peaceful outcomes are extremely challenging to reach.

Interstate wars over territory tend to end on the battlefield rather than through negotiated settlement (Hensel 1996), and the trajectory of competition and cooperation between the US and China remains highly uncertain. Can mediation offer a path to more settlement, peace, and prosperity in these challenging contexts?

This paper explains why efficient concessions are sometimes impossible to make. It also revises our understanding of why inefficient concessions can sometimes be beneficial and how mediators can help increase the efficiency of concessions and make cooperation possible. Weapons, strategically valuable territory, and the sharing of advanced technology will not be conceded if the conceding side fears these items may be used in future attacks against them. Countries are reluctant to make concessions that can be used against them in later rounds of conflict or negotiation. One way to mitigate the risk of dangerous concessions is to destroy the future value of the concession—e.g., decommission weapons, attempt to reduce the strategic value of the territory being conceded, or only allow trade in less advanced technologies or technologies without military applications. Another option, and the one on which we focus in this paper, is third-party mediation. We use this model to demonstrate that a third-party mediator can facilitate negotiated settlement by increasing trust and enabling disputants to make more efficient concessions. The mediator does this by only requiring concessions from negotiating parties who are able to cooperate, thereby removing the risk that concessions might be used against the giver in the future.

We model a party's ability to cooperate as private information. A high type can cooperate; a low type has incentives not to cooperate. These types map onto the commitment problem, with a high type one who is able to commit. A low type, though perhaps willing, is unable to do so. A low type may posture like the high type but lacks the political muscle or will to follow through on a peace agreement. Democratic leaders may more often be high types because they are less likely to face coups or other hostilities in reaction to short-term decisions. This is not always the case, however. Mahmoud Abbas, President of the Palestinian National Authority, may be a low type with respect to his negotiations with Israel not because he does not want to make peace, but because he cannot credibly act on behalf of the entire Palestinian population.

In our model, concessions may act as signals, as an end in themselves, or as both simultaneously.

One contribution of this paper is that it allows concessions to have future material value, signaling value, and both. We will show that when concessions have both signaling and future material value, some of the material value of the concessions may be purposefully destroyed, and thus inefficiencies can emerge. Another contribution is that our model of concessions does not require countries to make an agreement that requires credible commitment, as is common in bargaining models. Instead, countries willingly provide a concession if they expect the signal it sends to provide a benefit in terms of future cooperation and public good provision.[1]

In analyzing both the signaling and material value of concessions, an extensive literature on gift giving is particularly helpful. Like concessions, costly gifts have been shown to have a valuable role in relationship building. Furthermore, the gift-giving literature—more so than the concessions literature—focuses specifically on the issue of efficiency. For the purposes of this paper, we define inefficiency as giving a gift (i.e., a concession) that is more costly to the giver than it is valuable to the recipient.

Often, inefficient gifts occur in the formative stages of relationships or in immature relationships (Camerer 1988). Gifts given in established relationships, such as wedding presents or spousal gifts, are more likely to be cash gifts or gifts that are specifically requested and, therefore, efficient. Since inefficient gift-giving is more likely in relationships and partnerships in the early stages of development, it is natural to suppose that inefficient gifts play a signaling role in addition to the material value of transferring the gift's inherent worth.

In efforts to explain the existence of gift-giving, Camerer (1988) and Van de Ven (2000) give several anthropological explanations for inefficiency.[2] Amongst these explanations are altruism, social mores, and egoism. However, the most relevant explanation for international relations is that gift-giving is strategic. Camerer uses mechanism design to model gifts as costly signals, where inefficiency is useful in pre-play communication to form relationships between like types. This strategic giving of gifts in societies looks much like the strategic giving of concessions to remedy conflict situations. When

[1]When mediation is required to achieve cooperation in the model, it relies upon a manipulative mediator to enforce the required level of concessions.

[2]Prendergast and Stole (2001) also provide an explanation for inefficient gift-giving focused on the utility of matching.

modeled game-theoretically, these environments look even more similar.

Camerer's model one period in which players can signal before deciding whether to partner. There are two types, high types and low types, and a separating equilibrium is characterized by a threshold gift that is necessary for high types to reveal themselves, while low types do not send a gift. Two high types who have just received each other's gifts will then choose to create a partnership with a positive payoff; other combinations will not find it economical to form a partnership.

Camerer shows that in this general formulation, a costly signaling model will always yield efficient gifts.[3] In order to explain inefficiency, Camerer adds an additional pre-play period where players must pay a cost to enter the game and send and receive gifts. If types separate in the pre-play period with low types not being willing to pay, high types save the cost of sending a signal to low types in the main game. High types would, in some parameterizations, have incentives to give inefficient equilibrium gifts in the main game to lower the low types' expected payoff of playing to below the pay-to-play fee. Such an action can keep out the low types, raising the expected payoff for a high type by saving on the gifts given.

We propose a different conceptual framework rooted in a realistic international relations puzzle to explain this real-life phenomenon. The basic formulation of our mechanism-design approach expands Camerer (1988)'s analysis. The model can achieve the basic 'inefficient gifts' results of Camerer (1988) but does so using a very different explanation for inefficiency. The basic model is a costly signaling model, at its root, similar to Spence (1973)'s and subsequent papers.[4] In essence, the model shows that if concessions can be later used against the party that gives them, then it can be in a negotiating party's best interest to give inefficient concessions. While our model is tailored to an international relations setting, the explanation for the inefficient giving of concessions may be more broadly applicable.

This demonstration of the incentives for inefficient concessions leads to a new explanation for the role of mediators in conflict resolution. Third-party involvement is ubiquitous in conflict, and mediation has been extensively studied in the literature, but still, the question of third-party effectiveness is

[3]In an environment with repeated opportunities for concessions, Watson (1999) shows that a strong relationship can be reached by starting with small "gifts" and increasing their size as the relationship becomes more committed. See Van de Ven (2000) for an overview of this literature.

[4]See Connelly et al. (2011) for a review of the costly signaling literature.

intensely debated. Some authors have concluded that mediation has little impact (Bercovitch 1996; Bercovitch and Langley 1993; Fortna 2003), while others find mediation playing a positive role in resolving conflict (Dixon 1996; Beardsley et al. 2006). However, the term mediation encompasses a broad range of actions and interactions, making any one-size-fits-all conclusion on the efficacy of mediation difficult. The type of mediation affects the negotiation outcome (Bercovitch 2000; Beardsley et al. 2006), as does the legitimacy of the mediator (e.g., Duursma (2020)) and the timing of the mediator's intervention (Greig 005a).

It is essential to focus more precisely on the specific actions and environments that use particular types of mediation. To this end, this paper models an environment where the primary issue is trust, not uncertainty about an opponent's capabilities, and where parties employ a mediator who has manipulative abilities. Parties are more likely to seek mediation when they are entrenched in costly conflict (Bercovitch and Jackson 2001; Greig 2005; Greig and Diehl 2006; Terris and Maoz 2005; Svensson 2006); we thus model parties to a conflict who can benefit from mediation in their efforts to overcome mistrust.[5]

These empirical results have, until recently, lacked strong theoretical foundations. Mediation is not uncommon, but we are only beginning to understand why actors use it. Recently, mediation has begun to be modeled using mechanism design for settings where militarization or the resolve to fight are private information (Bester and Warneryd 2006; Fey and Ramsay 2009, 2010, 2011; Horner et al. 2010; Meirowitz et al. 2012, 2019). In particular, Meirowitz et al. (2019) show that—when asymmetric information is about the level of militarization—a 'Myerson' mediator can circumvent the incentives for militarization that arise in unmediated peace talks and improve the chances for peace. This paper expands the theoretical literature by modeling the role of a manipulative mediator when the information asymmetry is about parties' abilities to commit to cooperation, and concessions have a future material value that could be used to harm the giver. Here, a mediator can achieve peace or the provision of other public goods (1) in situations where bilateral concessions cannot and (2) with concessions that are more efficient when bilateral concessions must be inefficient to achieve peace.

[5]Alternatively, parties may seek a mediator to justify difficult concessions and avoid angering domestic constituencies (Allee and Huth 2006; Beardsley 2010; Beardsley and Lo 2014).

MODEL

Countries[6] are of two possible types, high and low. A high type discounts future period payoffs at rate δ_h and a low type discounts at δ_l where $\delta_h > \delta_l$ and $\delta_h, \delta_l \in [0, 1]$. Before the start of the game, nature independently determines the types of Country 1 and Country 2. The type distinction, as explained in detail later, encompasses the country's willingness to cooperate in solving the conflict; that is, throughout the analysis we will define a low type to be a player whose only sequentially rational strategy is to remain in conflict in all periods.[7] It is not necessary to think of a type as fixed; it can change as regimes and relations between countries change.[8]

The probability of nature selecting the high type for a given country is p. The probability of nature selecting the low type is $1 - p$.[9] Countries are aware of their own type but not the type of the other country. Similar scenarios have been modeled as credible commitment problems in the bargaining literature. Our model can be thought of as a reduced form bargaining model, but it can also be considered somewhat differently. That is, when trust is the issue to be solved, concessions are one way in which countries can signal their commitment. By modeling conflict in such a way, we can also gain insight into how a third party can ameliorate trust issues inherent to the conflict.

In period 0, countries 1 and 2 simultaneously give costly concessions $g_i \in \mathbb{R} \geq 0$, where $i \in \{1, 2\}$. Concessions are given at cost $C(g_i)$ and are valued by the recipient in the amount g_i. For simplicity, we assume throughout that the cost of giving the gift is equal to the value to the recipient, that is, $C(g_i) = g_i$.

[6]Players will be referred to as countries, though it could also be appropriate to think of players as intra-state actors, as in a civil war.

[7]This is in contrast to the bargaining theory of war, where a high(low) type has strong(weak) military capability that then interacts with a symmetric, known level of δ. This accounts for the different impact of δ in the two models.

[8]We can enrich the model to assign a probability (and even different probabilities) of a country's type shifting from high to low or low to high in a given period.

[9]Note that parameters are symmetric across countries, as is consistent with the literature. Qualitative results do not depend on this being the case, and equilibria and equilibrium concessions can easily be calculated with asymmetric parameters between players.

Beginning in period 1, countries engage in an infinitely repeated Prisoner's Dilemma once per period with stage game payoffs in Table 1.[10] While not a good representation of all-out war or even a zero-sum short-term war, the Prisoner's Dilemma is appropriate to represent a limited, continuing conflict, military or otherwise. In cases appropriately modeled by a Prisoner's Dilemma, it would be better for both parties to exit the protracted conflict, but no party has the unilateral incentive to do so. The critical results of this paper do not depend on a prisoner's dilemma structure; they require only that a better result can be achieved through trust or credible commitment. Thus, a stag hunt or other game structures are also possible. The Prisoner's Dilemma is used because of its wide exposure in the literature.[11]

The countries' stage game actions are referred to as "Trust" and "Distrust" as represented in Table 1, where Country 1 is the row player and Country 2 is the column player.

TABLE 1. Stage game payoffs.

	Trust	Distrust
Trust	T, T	$-D, T+W$
Distrust	$T+W, -D$	$W-D, W-D$

Here, $T > 0$ represents the benefit from a country's negotiating partner playing Trust. $D > 0$ represents the damages incurred when a country's negotiating partner plays Distrust, while $W > 0$ is the additional benefit a country receives when it plays Distrust. We will refer to the stage-game Nash Equilibrium of (Distrust,Distrust) as No Cooperation. We will refer to (Trust,Trust) as Cooperation. In line with the assumption that the countries play a Prisoner's Dilemma in the stage game, we assume $T > W - D$. Note that this ensures that the joint payoffs from Cooperation are the highest among all the stage-game action profiles.

[10]For a game-theoretic analysis of the repeated Prisoner's Dilemma with uncertainty over discount factors that focuses on studying players' belief structures, see Maor and Solan (2015). Our model differs in that it adds an initial period before the repeated Prisoner's Dilemma begins, in which players make costly concessions to each other. We will see that these concessions that enable signaling between the players make the analysis of their beliefs much simpler than in Maor and Solan (2015).

[11]This payoff matrix does not distinguish between civil and interstate conflict, though in practice, the effects on the payoffs might have a different structure in the different cases.

Payoffs for a country are the sum of the stage game payoffs, discounted by δ_i. For example, if both parties play "Trust" in every period, the payoff for a player i is $\sum_{t=1}^{\infty} \delta_i^{t-1} T = \frac{T}{1-\delta_i}$.

We will use grim trigger punishments throughout. That is, (Distrust,Distrust) will be played forever if any party plays Distrust in any round. Because we use grim trigger punishments, only four paths of play after concessions are given are interesting:

1. Both countries always play Trust;
2. Both always play Distrust;
3. Country 1 plays Distrust in round 1 while Country 2 plays Trust, which is followed by both countries playing Distrust from round 2 onward; and
4. Country 2 plays Distrust in round 1 while Country 1 plays Trust, followed by both countries playing Distrust from round 2 onward.

Equilibrium payoffs beginning from round 1 are now easily described by noting only the collective first round behavior.[12] Let X_{ij} represent the sum of discounted payoffs for Country i where subscripts represent first round strategies of that country. If both countries play T in round 1, the corresponding payoff to Country i is represented as $X^i_{TT} = \frac{T}{1-\delta_i}$. Likewise $X^i_{TF} = (-D + \frac{\delta_i(W-D)}{1-\delta_i})$, $X^i_{FF} = \frac{W-D}{1-\delta_i}$, and $X^i_{FT} = T + W + \frac{\delta_i(W-D)}{1-\delta_i}$. X^h_{FT}, for example, represents payoffs for the high type of Country 1 when Country 1 plays Distrust and Country 2 plays Trust in the first period. Note that the payoffs differ by type because types have different discount factors.

All parameters are common knowledge except δ_i, which is Country i's private information. We take achieving peace, or more broadly, achieving cooperation with the provision of public goods, as the normative goal with respect to advancing social welfare. Only high types are capable of making peace/providing public goods, and thus, we will take the measure of social welfare to be the sum of the participating high types' utilities.

[12]If a parameter of the model changes after the period 0, the equilibrium may be disrupted. Although it is beyond the scope of this paper, one could model this by adding a probability of equilibrium disruption followed by a new round of concessions that would essentially restart the game—albeit with altered parameters.

ANALYSIS OF BENCHMARK CASES

When countries are uncertain about each other's motives, they can proceed in several ways, including giving concessions to signal their ability to Cooperate. We will examine three types of equilibria.[13] The Appendix provides a thorough analysis of the various equilibria as well as most proofs.

Pooling equilibrium

In the *pooling equilibrium*, neither negotiating party gives concessions in period 0; then both parties play Distrust in every period beginning with period 1. Countries do nothing to bridge the credibility gap and thus do not cooperate. This *pooling equilibrium* is likely to happen when countries are in long-set patterns of distrust. Active fighting is unnecessary; remaining outside of a state of peace/cooperation is all that is required. See the Appendix for a formal characterization of this type of equilibrium.

Sri Lanka offers one example. While the government has finally won the war of secession, it still refuses to address the grievances of the minorities and leaves the conflict unresolved. Other prominent examples of this pooling equilibrium include Israel and its neighbors, Cyprus's civil war, North and South Korea, Greece and Turkey, and a multitude of other long-lasting conflicts. The mistrust is perhaps justified; indeed, if one party took the first step and gave a concession, they may be taken advantage of. This is also possible outside the war context, for instance, when U.S. officials block Chinese firms' access to cutting-edge technology to prevent that technology from being used against U.S. interests.

Without a third party, little hope exists for overcoming mistrust. These long-lasting stalemates are the types of conflict which are most appropriate for mediator involvement. In Section Mediation, we will present a mediation mechanism that illustrates how mediation can help achieve cooperation in this context. In some sense, these cases are the most interesting because the international community might play a positive role.

[13] If the discount factor is high enough, many types of equilibria exist. Consider, for instance, oscillating every other period between (Trust,Trust) and (Distrust,Distrust) with a grim trigger Distrust threat. In this paper, we are not interested in these equilibria because they are not as realistic for applications to conflict scenarios. These are also always payoff-dominated by at least one of the other equilibria discussed and, thus, are not attractive.

No-Concessions Separating Equilibrium

The other two equilibria on which we focus are separating equilibria in which low and high types have different strategies. If δ_i is high enough for both countries, the Cooperation outcome can be sustained with no concessions and a grim trigger punishment threat. We call this a *no-concessions separating equilibrium*. In this equilibrium, no concessions are given by either type in period 0. In period 1, high types play Trust while low types play Distrust. So long as the countries are both high types as revealed by their period 1 play, both countries play Trust in all following periods. Otherwise, both countries play Distrust in all following periods.

Proposition 1 No Concessions Separating Equilibrium

Assume $\delta_h \geq \frac{W}{(1-p)W+p(T+D)} = \delta^{nc}$. *A subgame perfect Nash equilibrium exists in which low types give no concessions in period 0 and always play Distrust and the high types give no concessions in period 0 and play Trust unless the other country defects to Distrust.*

Proof: See Appendix.

Here, as in each further iteration of the model, we define a high type as a country with a discount factor at or above the threshold for cooperation. That is, high types prefer to play the Cooperation equilibrium, i.e., $\delta_h \geq \delta^{nc}$. Any country with $\delta_l < \delta^{nc}$ is a low type for the purposes of the *no concessions separating equilibrium*. This means that two high types—and only two high types—can sustain the Cooperation outcome.[14]

In this equilibrium, two mutually distrustful countries *behave* as if they trust each other, taking a risk. This generally will happen when the result of trusting and then being betrayed is not catastrophic or the odds of such a scenario are small. This is representative of a case where countries have suspicions about each other but proceed anyway.

[14]Note that because we define high and low types relative to each type of equilibrium, a country with a given discount factor can be a high type in some equilibria and a low type in others.

Concessions Separating Equilibrium

The other separating equilibrium of interest involves distrustful countries using concessions as a way of building trust. We call this the *concessions separating equilibrium*. In this equilibrium, the cost of giving a concession can be worth the investment if the other country also cooperates. It is also not as risky as the *no-concession separating equilibrium*. In this equilibrium, concessions are given before more substantive actions, thus removing the possibility of playing one of the cheating corners of the Prisoner's Dilemma. Concessions are a costly signal that the other country is serious about cooperating and can be used to avoid continual distrust. High types use concessions to identify themselves as trustworthy partners in hopes of establishing cooperation.

In the *concessions separating equilibrium*, low types do not give concessions because the concession level is set as the smallest amount that will deter the impatient type from mimicking the patient type. In contrast, high types always give a concession in period 0. Thus, the patience level of both countries is fully revealed in equilibrium. If both countries receive concessions, they know the other is a high type, and both play Trust in subsequent periods. Otherwise, if even one player does not receive a concession, there is at least one low type, so both countries play Distrust in subsequent periods. The signaling value of concessions in period 0 means that starting from period 1 in the Trust/Distrust game, both countries play the same strategy in a separating equilibrium. That is, both play Trust if there are two high signals, and otherwise, both play Distrust.

Details of the analysis are in the Appendix. We begin by characterizing the concession (or *gift*) giving behavior of the low type. We then present the patience threshold and gift-giving behavior of the high type.

Lemma 1 *In a separating equilibrium, low types do not give a concession. Cheap talk does not allow for a* concessions separating equilibrium.

Proof: See Appendix.

Low types don't give gifts in this separating equilibrium because the gift would fail to serve any purpose: the low type will be identified through the separating behavior of the high types, and there is thus nothing to gain.[15]

[15]The transmission of information about each country's type is, by definition, essential for the existence of

Proposition 2 Concessions Separating Equilibrium (CSE)

(a) *Assume* $\delta_h \geq \frac{W}{T+D} = \delta^c$*. Then a subgame perfect Nash equilibrium exists in which low types give no concessions in period 0 and always play Distrust and the high types give a concession in period 0 and play Trust unless the other country defects to Distrust.*

(b) *In the best* concessions separating equilibrium*, high types give a concession of* $p(T+D)$*, which is the smallest concession necessary to separate. Equilibria with higher concessions yield strictly lower payoffs.*

Proof: See Appendix.

This 'best' *concessions separating equilibrium*, where 'best' is from the standpoint of high-type welfare, is justified by appeal to the Intuitive Criterion equilibrium refinement, in which out-of-equilibrium beliefs put zero weight on types that can never gain from deviating.

Interestingly, the patience threshold for this equilibrium to exist, $\delta^c = \frac{W}{T+D}$, is the same as the threshold for the complete information repeated game in which deviating from "Trust" is discouraged through grim trigger punishments.[16] Using concessions to separate the types, therefore, can be seen as making up for the imperfect information in this setting.

We complete our analysis of the *concessions separating equilibrium* by comparing it to the *no-concessions separating equilibrium* in Section No-Concessions Separating Equilibrium.

Corollary 1 *The* no-concessions separating equilibrium *requires countries to be more patient than the*

a *concessions separating equilibrium*. Thus, there cannot exist a *concessions semi-separating equilibrium* where the low type gives a concession with strictly positive probability; the concessions would not transmit the information required to build trust. If no information is transmitted, the high type has no incentive to give a concession. Therefore, there cannot be a semi-separating equilibrium with the low type mixing between strictly positive and zero concessions and the players separating in the repeated game as in the *no-concessions separating equilibrium*.

[16]When $\delta = \frac{W}{T+D}$, there is also a semi-separating equilibrium where only some high types give a concession. Here, high types give a smaller expected concession and trade off a lower probability of achieving cooperation. We will focus on the ff *concessions separating equilibrium* in the analysis below because it provides the best chance for Cooperation and is possible over a much larger portion of the parameter space.

concessions separating equilibrium. *If both equilibria exist, the* no-concessions separating equilibrium *is preferred when the probability of encountering a high type is sufficiently large.*

Proof: See Appendix.

A large probability of a high type makes the *no-concessions separating equilibrium* very attractive because the risk of ending up in the cheating corner of the Prisoner's Dilemma is low, and no concessions have to be made. Conversely, for smaller values of p, high types are better off paying the gift to avoid the possibility of being taken advantage of by a low type.

Concessions separating equilibria are a way for countries to solve their disputes when the *no concessions separating equilibrium* is unattractive or unavailable. By giving costly concessions, they signal their intentions in a way that a partner unwilling to cooperate would not be willing to signal.

One example of an equilibrium involving concessions is Israel and Egypt at the Camp David Accords. Both parties gave concessions (Israel a material land concession and Egypt diplomatic recognition of Israel) that signaled their willingness to end the stalemate that had characterized the previous five years. Of course, this result depended on Carter's use of mediation techniques, including use of manipulative mediation—a concept we will explore in Section Mediation below.

MATERIAL VALUE OF CONCESSION AND INEFFICIENCY

Concessions with Material Value

We now model concessions to have real future material value that can either help or harm the giver. To be clear, in Sections Model and Analysis of Benchmark Cases, the gift only benefits the party who receives it in the period in which it is received, that is, Period 0. In this section, we assume that there is an additional value from the gift in each period of the repeated game, that is, Period 1 and onward.

The intuition for this addition to the model is that the country that receives the concessions can use them to build up its military or to build up its civil society, including through the provision of public goods. The recipient using concessions to build military strength hurts the giver of the concessions in war; the recipient using the concessions to strengthen civil society helps the giver in peace. The Golan Heights is such an example: if Syria maintains peaceful relations with Israel, the Golan Heights is a reasonable concession to make, but if Syria renews military conflict with Israel, surrender of the strategic land would be disastrous for Israel.

To incorporate future material value, we add two elements to the benchmark model in Section 2. The first is an additional decision for the negotiating parties. In Period 0, each country decides how to allocate the value of any concession it receives between military and civil society. α_i is the portion of the received concession that Country i chooses to dedicate to civil society; $(1-\alpha_i)$ is the portion Country i dedicates to military buildup. This decision is a simple optimization problem for each country. Parameters are common knowledge, so the result of this decision is well known should the country's type be known.

Concessions are immediately converted into civil society or military gifts and are incorporated into payoff functions starting in Period 1. Periods 1 to ∞ are as before but with the modified stage game payoffs shown in Table 2.

TABLE 2. Stage game payoffs when concessions have material value.

	Trust	Distrust
Trust	$T + T\alpha_2 g_1$, $T + T\alpha_1 g_2$	$-[D + D(1-\alpha_2)g_1]$, $[T + T\alpha_1 g_2] + [W + W(1-\alpha_2)g_1]$
Distrust	$[T + T\alpha_2 g_1] + [W + W(1-\alpha_1)g_2]$, $-[D + D(1-\alpha_1)g_2]$	$[W + W(1-\alpha_1)g_2] - [D + D(1-\alpha_2)g_1]$, $[W + W(1-\alpha_2)g_1] - [D + D(1-\alpha_1)g_2]$

This modification to the payoffs is the second change in the model relative to Sections Model and Analysis of Benchmark Cases. We add a term to each of T, W, and D that represents the additional benefit or harm flowing from the future material value of the concession. When both countries play Trust, they each receive the base benefit T as well as an additional benefit of the concession they *gave* that is scaled by the proportion the recipient invested in civil society as well as the base benefit of T.

If both countries play Distrust, Country 1 receives not only the immediate value of the concession but also $W - D$ in each period. Its welfare now has two additional terms. First, an additional benefit term in each period of the repeated game $W(1-\alpha_1)g_2$ accounts for the value of the received concession as well as how much Country 1 invested in the military. Second, Country 1 also incurs extra damages $D(1-\alpha_2)g_1$ in each period of the repeated game that are proportional to the size of the concession it *gave* and how much of that concession was invested in the military by Country 2. The other payoffs are modified analogously.[17]

Because α_i enters into the payoffs in a linear fashion, the choice that maximizes welfare in equilibrium is $\alpha = 1$ for high types since a concession would only come from another high type in equilibrium. Thus, only the Cooperation outcome will be played. Conversely, in equilibrium, low types only participate in the No Cooperation equilibrium and thus maximize their payoffs by choosing $\alpha = 0$.

We turn next to characterizing a *concessions separating equilibrium* in this environment with future material value.

Proposition 3 Concessions Separating Equilibrium (CSE) with Material Value

(a) *Assume* $(1-\delta_l)(1-pT)+(1-p)D$ *is positive and* $\delta_h \geq \frac{p(W-D-T)}{\frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}} + p(W-T) + (1-p)D + 1 = \delta^1$. *Then a subgame perfect Nash equilibrium exists in which low types give no concessions in period 0 and always play Distrust and the high types give a concession in period 0 and play Trust unless the other country defects to Distrust.*

(b) *In the best* concessions separating equilibrium *when gifts have full material value, high types give a concession of* $\frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}$, *which is the smallest concession necessary to separate. Equilibria with higher concessions yield strictly lower payoffs.*

Proof: See Appendix.

The problem with concessions with material value is that the potential for the low type to invest the concessions in the military makes more patient types less willing to take a chance that it will be matched with a high type when the costs of being matched with a low type are too high.

[17] The results are qualitatively unchanged if the gifts are not scaled by the variables that represent the impacts of cooperation and non-cooperation, if they are scaled linearly by these variables but with a proportion other than 1, or if they are scaled by many plausible non-linear functions of T, W and D. The results are simpler if there is no scaling. However, we believe it is more realistic to assume that the future benefits and costs of the gifts are proportional to the direct benefits and costs of cooperation/non-cooperation. For instance, a county that could impose more damage without the future value of the concessions is likely to be able to leverage a concession better than a country that can impose less damage.

Corollary 2 *Assume that high types are not patient enough to separate without concessions but can separate using concessions when gifts have no material value. If countries use concessions with future material value instead of concessions with no future material value, there are parameters under which the future material value of gifts destroys countries' ability to separate through concessions, i.e., Cooperation cannot be achieved when there are two high types.*

Proof: An example suffices as proof. Let $T = 1, W = 1$ and $D = 1$. The assumption that $T > W - D$ is satisfied, and the threshold for high types to separate when concessions do not have future material value is $\delta = \frac{1}{1+1} = 0.5$. Let $\delta_h = 0.7$ and $\delta_l = 0.3$ and $p = 0.9$.

By Propositions 1 and 2, both the *no-concessions separating equilibrium* and the *concessions separating equilibrium* are possible when gifts have no material value.

The patience threshold to separate through concessions with future material value is much higher under these parameters when the concessions carry future material value: by Proposition 3(a), this threshold is 0.979. Thus, the *concessions separating equilibrium with efficient gifts* is *not* possible.

δ_h is not involved in any of these calculations, so we can see that for any $\delta_h \in [.5, .979)$, separating through concessions is possible when concessions have no material value but is not possible when they do have future material value. That is, future material value can destroy the opportunity for cooperation.

■

The comparative statics for the patience threshold $\delta^c = \frac{W}{T+D}$ are straightforward for the *concessions separating equilibrium* when there's no material value. The threshold increases and cooperation becomes less likely when W increases; the threshold decreases when either T or D increase. These relationships hold with a small caveat for the *concession separating equilibrium* when there *is* material value, where the patience threshold becomes $\delta^1 = \frac{p(W-D-T)}{\frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}} + p(W - T) + (1 - p)D + 1.$[18] The patience threshold increases in W because it makes the benefit from playing Distrust larger. The patience threshold decreases in T and D but for different reasons. T is straightforward: a larger T is a larger benefit from playing Trust. In contrast, a larger D is a higher cost of the negotiating partner playing Distrust. The presence of future material value does not fundamentally change these relationships because the future material value is proportional to the fundamental values of W, T, and

[18]The caveat is that, if both T and W become sufficiently large, the patience threshold will increase in D.

D.

However, two variables do not affect the patience threshold in the no material value case that *do* affect the threshold when there is future material value: δ_l and p. When the low type's patience level δ_l increases, the high type's patience threshold increases because a more patient low type requires a more significant gift to avoid mimicking the high type. In the case of no material value, the low type's discount factor doesn't enter into the gift, and therefore, it doesn't influence how much patience is required for cooperation.

The other variable influencing the patience threshold in the case of future value is p, the proportion of cooperative types in the population. Although the gift in the no-future-material-value case *is* an increasing function of p, it is a linear function of p just like the variables T, W, and D. Thus, the cost of paying the gift and the direct benefits in the repeated game increase at the same rate so that the likelihood of cooperation does not vary in p when gifts have no future material value. In the future-material-value case, the size of the gift required to separate the high types from the low types increases more than linearly in p and thus faster than the other terms. Therefore, the patience threshold in the future-material-value case increases as p increases. That is, the cost of giving the gift weighs more heavily in the decision to cooperate relative to the material costs and benefits, requiring states to be more patient to separate and thus reducing the likelihood of cooperative types.[19]

Material Value of Concession can be Destroyed

We now relax the assumption about the benefit of a concession being equal to its cost and, instead, allow countries to "burn" a portion of the concession they give. For instance, when decommissioning weapons, rebels leave no material value of the weapons intact to benefit the recipient, but it still signals good intentions on behalf of the giver. Here, we allow the giving of efficient concessions along with new options to give a variety of less efficient concessions.

We implement this assumption by allowing countries to choose a scalar $e \in [0, 1]$ by which to multiply their given concession g. $C(g) = g$ as before, but while the receiving country incurs the full cost of the concession g, the future benefit to the recipient is now only eg.

[19]For the comparative statics calculations, see Appendix 6.4.

TABLE 3. Stage game payoffs when concessions have material value that can be destroyed.

	Trust	Distrust
Trust	$T + T\alpha_2 e g_1$, $T + T\alpha_1 e g_2$	$-[D + D(1 - \alpha_2) e g_1]$, $[T + T\alpha_1 e g_2] + [W + W(1 - \alpha_2) e g_1]$
Distrust	$[T + T\alpha_2 e g_1] + [W + W(1 - \alpha_1) e g_2]$, $-[D + D(1 - \alpha_1) e g_2]$	$[W + W(1 - \alpha_1) e g_2] - [D + D(1 - \alpha_2) e g_1]$, $[W + W(1 - \alpha_2) e g_1] - [D + D(1 - \alpha_1) e g_2]$

We have established the results for the *concessions separating equilibrium* when no material value is destroyed, i.e., $e = 1$, in Section Material Value of Concession and Inefficiency. We have also established the results for the *concessions separating equilibrium* when all material value is destroyed, i.e., $e = 0$. Notice that if $e = 0$, the payoffs in Table 3 are the same as those in Table 1, and thus, the results of Section Concessions Separating Equilibrium apply.

Before turning to the question of how the possibility of destroying the value of concessions impacts the ability of the parties to cooperate, we first formally establish the welfare implications.

Lemma 2 *If the* concessions separating equilibrium with efficient concessions *exists, it is preferred to the* concessions separating equilibrium with inefficient gifts *when both the probability of encountering a high type and the benefit from Cooperation are sufficiently large.*

Proof: See Appendix.

The intuition for Lemma 2 is that the future material value of concession helps you when your negotiating partner turns out to be a high type, and this happens with probability p. On the other hand, the material value of your concession *hurts* you when your negotiating partner turns out to be a low type. Similarly, the larger the benefits from Trust, the larger the welfare gain when the negotiating partner is a high type, and the larger the damages from Distrust, the more significant the welfare loss when the negotiating partner turns out to be a low type.

Lemma 2 implies that the high types prefer either $e = 1$ or $e = 0$. As we turn to our result about incentive compatibility when concessions have material value, we focus on these two cases as the most focal and easiest upon which to coordinate.[20]

[20]For some parameter values, we cannot rule out that a *concessions separating equilibrium with* $0 < e < 1$ exists and improves upon high type welfare when high types are patient enough to separate when $e = 0$ but not when $e = 1$.

Lemma 3 *The* concessions separating equilibrium with efficient concessions *requires countries to be more patient than the* concessions separating equilibrium with inefficient concessions.

Proof: See Appendix.

This means that there are parameters under which Cooperation cannot be achieved in a *concessions separating equilibrium with efficient gifts* but Cooperation can be achieved in a *concessions separating equilibrium with inefficient gifts*. This is part of the substance of Proposition 4.

Proposition 4 *Assume the high types are not patient enough to separate without concessions and that concessions have material value that can be destroyed. There are parameters under which the optimal equilibrium is a separating equilibrium in which concessions are inefficient.*

Proof: We invoke Lemmas 2 and 3 as two different reasons for which the parties would choose the *concessions separating equilibrium with inefficient gifts*. First, they may be patient enough to separate when gifts are inefficient but not when gifts are efficient (Lemma 3). In this case, the *concessions separating equilibrium with efficient gifts* is not available to them.

However, even if the negotiating parties are patient enough to separate when gifts are efficient, Lemma 2 tells us that welfare will be higher with efficient gifts only when the probability of encountering a high type and the benefit from Cooperation are sufficiently large. Otherwise, welfare is higher under the *concessions separating equilibrium with inefficient gifts*, and this equilibrium is thus optimal. ■

Inefficient concessions are optimal when there is a good chance of facing a low type who will use the concession against you (i.e., p is small) as well as when the benefits from Cooperation are small relative to the damages from No Cooperation. This finding of inefficiency is for profoundly different reasons than the literature related to Camerer (1988) or Prendergast and Stole (2001).[21] Instead of being in reaction to a behavioral regret for mismatching or serving as a way to discourage low types

[21] Since our initial setup is essentially identical to Camerer's, the inefficiency of pre-play communication would also manifest in our model if a stage of pre-play communication were added. With a slight behavioral modification, the reasoning of Prendergast and Stole (2001) can also be had: if a country suffers regret after efficiently giving to a type different than their own, there is a premium for guessing correctly. In Prendergast and Stole, the two types are equal except in labeling; in our model, there is an incentive for low types to try to pose as high types.

from entering the game altogether, here, making concessions inefficient is a device to prevent low types from using those concessions against the giver of the concessions in the future.

MEDIATION

The results in this paper thus far have relied on bilateral engagement. We will now consider the involvement of a mediator when concessions have future material value. We model the mediator as solving a mechanism design problem to create a plan for concessions that elicits truthful revelation of the parties' abilities to commit to cooperation.

Because many conflicts occur in an international arena with no clear enforceable rule of law (Waltz 2018), much of the literature has focused on self-enforcing mechanisms. If, however, a credible third party exists, then results beyond self-enforcing agreements can be relevant. A mediator with such enforcement ability is said to use manipulative mediation, specifically if she "offered to verify compliance with the agreement" or "took responsibility for concessions" (Bercovitch 1996). We will now introduce a mediator who does the latter.[22]

The model setup is as before, but instead of parties being free to choose their concession levels, the mediator solicits type reports and then enforces incentive-compatible concessions. This mediator will be modeled as a mechanism M.

After learning its type, each country chooses to participate in the mechanism or not. By the Revelation Principle (Myerson 1979), it is without loss of generality to restrict attention to direct mechanisms[23] and to focus on equilibria in which both countries participate and send truthful reports. Participation in the mechanism, therefore, involves sending a message that declares one's type to be High or Low and then being bound to send the mandated concessions. Formally, the mechanism

However, putting this difference aside, the result of inefficiency can be viewed as having the same flavor as that in Prendergast and Stole (2001). If countries suffer a "mismatching" cost due to giving to the wrong type, they will be hesitant to do so and may offer inefficient concessions.

[22]Here, the parties' actions in the stage game, as well as their choices of whether to spend concessions on the civil versus the military sector, are assumed to derive from their type, so there is no need for the mediator to enforce these decisions once the type is revealed. Thus, assuming that the mediator can do both is unnecessary.

[23]Direct mechanisms are those in which countries simply declare their types.

M inputs the reports $t_i \in \{High, Low\}$ for $i \in \{1,2\}$ and outputs the required concessions g_i and efficiency levels e_i so that $M : (t_1, t_2) \to (e_1, g_1; e_2, g_2)$.

In Sections Analysis of Benchmark Cases and Material Value of Concession and Inefficiency, the same concession must be sent to both types since types are private information. Importantly, here, the mechanism M can be differentiated by recipient type. Otherwise, the setup is the same as in Section Material Value of Concession can be Destroyed, where concessions have future material value that can optionally be destroyed. In the mechanism design context, we focus on the incentive compatibility constraints that must be satisfied for each type to reveal its type truthfully.

We find that some value of the concessions must be destroyed to prevent the low type from being tempted to misrepresent its type in order to receive the gift. That is, $e < 1$, implying that the gifts mandated by the mediator must be inefficient.

Lemma 4 *A manipulative mediator cannot facilitate the most efficient concessions, i.e., $e = 1$ is not incentive compatible under mechanism M.*

Proof: See Appendix.

The range of efficiency that allows a mediator to make mechanism M incentive compatible and individually rational for both types is given in Proposition 5.

Proposition 5 Mediation

(a) A mediator using mechanism M must ensure that M is incentive-compatible and individually rational for the low type. This happens when the specified concession is greater than or equal to $\frac{(1-\delta_l)(T+D)}{p-e(p+(1-\delta_l)T+W)}$, with the denominator of this expression strictly positive.

(b) A mediator using mechanism M must ensure that M is incentive-compatible and individually rational for the high type. This happens when either $Te < (1-\delta_h)(1-e)$ or the specified concession is less than or equal to $\frac{T+D-W}{Te-(1-\delta_h)(1-e)}$.

Proof: See Appendix.

In part (a), we see that the low-type incentive compatibility constraint determines a lower bound on the size of the concession specified by the mechanism. This puts an upper bound on how efficient concessions can be. In part (b), we see that, in some cases, the high-type incentive constraint also provides an upper bound on the concession size.

Ultimately, mediation is desirable if it can improve welfare or improve the chances for cooperation. Proposition 6 shows that a mediator who implements mechanism M can do both.

Proposition 6 *When concessions have future material value, a mediator can use mechanism M to*

(a) *improve welfare by increasing the efficiency of concessions when a* concessions separating equilibrium with efficient gifts *is not possible; or*

(b) *improve welfare despite reducing the efficiency of concessions when a* concessions separating equilibrium with efficient gifts *is possible; or*

(c) *achieve cooperation when it cannot be achieved either in a* concessions separating equilibrium with efficient gifts *or a* concessions separating equilibrium with inefficient gifts.

Proof: See Appendix.

To understand part (a), remember that concessions are not exchanged in the mechanism if there is even one report of "Low." Since a concession is only given to the high types and high types invest in civil society and public goods, the efficiency of the concession can be increased without creating extra damages when matched with a low type; this is of course subject to satisfying the low type's constraints. High-type welfare is improved because no transfer is made to the low types, and there is no threat of future damage from the concessions since they only go to a high type.

Here, inefficiency can be mitigated because the mediator is able to mandate concessions only between declared high types. The inefficiency is no longer needed to prevent low types from using concessions against high types. However, some inefficiency is required to incentivize low types to reveal their type truthfully.

Part (b) may be somewhat surprising. Here, a mediator would insert inefficiency when it is not necessary to achieve a cooperative outcome. This happens when the possibility of the low type using concessions against their negotiating partner and the magnitude of the damages that can be imposed are small enough that efficient gifts are possible but large enough that the full efficiency of the gifts has a large, negative impact on welfare. The mediator removes the possibility of future damages but requires some inefficiency to incentivize low types to be revealed truthfully.

Part (c) is relevant when the impacts of future material value are too strong, i.e., when the probability of a low type is too high and/or the damages from investing in military power are too large. In this case, the parties cannot cooperate on their own, even with the possibility of signaling their type through concessions. By ensuring that only high types exchange gifts, the mediator removes the threat of damaging behavior by one's negotiating partner as well as the cost of giving the concession to a low type. Here, the mediator's help allows types with lower patience levels to separate when those same types cannot separate in the bilateral game with concessions with future material value.

Thus, a mediator is able to both improve welfare and bring willing parties to the table to give the concessions necessary to achieve the cooperative outcome. The improvements by the mediator here do not rely on the mediator's private information as in the models of Kydd (2003), Rauchhaus (2006), and Smith and Stam (2003). Therefore, the mediator's preferences need not be an important factor in her participation. So long as the costs are small enough to be worth her participation, the mediator can effectively enter and assist in solving the conflict.

One might worry about the requirement that the mediator be willing and able to ensure that the specified concessions are delivered. While it is not guaranteed that a mediator who will play this role exists for every conflict, it is not difficult to believe that such a mediator exists for many conflicts. We do not require a formal international enforcement mechanism but rather that the mediator has, and is willing to use, sufficient leverage to punish a deviating party. Examples of such punishments are financial or trade sanctions or the withdrawal of aid or military support. Furthermore, the mechanism can work if the mediator is either (1) only able to incentivize some portion of the concessions or (2) enforces the concessions with some probability less than 1. This would tighten the low type's incentive constraint and loosen the high type's incentive constraint, creating a different level of optimal efficiency of the concessions.[24]

The value of truth telling by the mediator has been examined in the literature. Kydd (2003) studies

[24]To see this, in the incentive constraints in the proof of Proposition 5, replace g_h with $q \cdot g_h$, where q represents the proportion of the concession that can be enforced or the probability with which the full concession is enforced.

bias in mediation and how a mediator can credibly communicate information to the parties. Conversely, Horner et al. (2010) rely on the mediator to "hoard" information in order to improve on unmediated interaction.[25] In our model, a successful mediator relies on the parties trusting her. However, here, the mediator has no incentive to lie: lying can only potentially hurt the high types, the only type the mediator can help.

CONCLUSION

In situations of perpetual conflict, concessions have been shown in the empirical literature to ease distrust. This paper introduces a mechanism by which this might work. In our model's setup, there are three classes of interesting equilibria. First, if prior trust is low, a no-concessions pooling equilibrium exists, which leaves all types in a state of non-cooperation. Second, if prior trust is sufficiently high, a no-concessions pooling equilibrium exists where high types trust in the first period and can maintain that trust if both are high types. Third, a concessions-giving separating equilibrium exists where high types offer concessions to eliminate distrust. This is the most interesting class of equilibria for our purposes.

While parties generally are better off when concessions are efficient, we have shown that inefficiency may be useful. In particular, if concessions may be used to harm the giver materially, it may be better to give inefficient concessions. This explanation captures the hesitancy of nations to fully engage in moving forward with a peace process or other cooperation when there is mutual distrust.

At first glance, it may seem somewhat abstract to model a concession as a number. However, the magnitude of concessions can indeed vary: often, concessions do come in sizes, for example, the number of prisoners released, the amount of land ceded, or the amount of time given to carry out actions. But while magnitudes often do vary, this points to a very real caveat. The nature of the concessions is sometimes non-negotiable. For instance, if one country demands the release of hostages, it is not possible—-or at least not realistic—-for the other country to give the equivalent value in land, cash, or weapons.

[25]Horner et al. (2010) take cheap talk communication as their benchmark of unmediated communication, whereas we take the unmediated signaling game as our benchmark.

This points out that when we think about concessions, we should think about their nature as being fixed. With this in mind, we can then accurately consider whether concessions are of sufficient size. These observations do not undermine the model. The nature of the desired concessions helps determine whether inefficient concessions can help achieve cooperation when efficient concessions cannot and thus also when a mediator can improve the efficiency of concessions.

The effectiveness of mediation is debated in the literature; this paper shows that a third-party mediator can remedy a particular kind of inefficiency if she is trusted. One contribution of this paper is to show a new mechanism through which a mediator can help.

This model focuses on manipulative mediation, which is indeed used in resolving conflicts and achieving cooperation. However, the usefulness of mediation in this case relies on the mediator being both perfectly trustworthy and perfectly able to enforce concessions. Such an all-powerful mediator is a strong assumption, and while useful as a benchmark, it is unlikely to exist in all scenarios. It would be useful as a further exercise to examine the effectiveness of mediators with different strengths of enforcement capabilities.

While it is important theoretically to show how a mediator can help in conflict situations, it should be noted that mediation is not free, at least to the providing (third) party. Therefore, mediation should not be undertaken if the cost of providing the mediation is higher than the benefits of increased cooperation and/or efficiency of concessions. If the model accurately captures the costs, the expected savings of mediation can be calculated, and therefore, the value of mediation in a specific scenario can be determined. This will allow the mediator to analyze whether her involvement is worthwhile.

This paper shows that inefficient concessions may be used to achieve peace or cooperation to provide other public goods when efficient concessions are insufficient; it also demonstrates how a mediator can achieve cooperation in these situations when the parties to the conflict cannot do so on their own. Opportunities for further research include looking at the roles of different types of mediation. Because of the many humanitarian, security, and diplomatic implications of this work, both deeper and broader analysis is required.

APPENDIX

Pooling equilibrium

Lemma 5 *From period 1 on, playing (Distrust,Distrust) in all periods 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 period 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" by definition—and always play 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 period given player j's behavior. Given that the stage-game equilibrium will be (Distrust,Distrust), there is 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 period and thereafter, as long as Trust has always been played, will be incentive compatible for high types if the expected payoffs from this strategy are greater than the payoffs from playing Distrust in the first period and then every period thereafter by the one-shot deviation principle. If both countries play Trust in a period, a Cooperation equilibrium takes place, and both countries continue to play Trust in every period with the grim trigger threat of Distrust. Low types play Distrust in the first period, so a high type that observed Distrust being played would respond by playing Distrust in period 2 and all future periods. 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 period 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 with no concessions in period 0.[26] If, instead, high-type negotiating parties have sufficient incentive to give a concession that reveals their type in period 0, the high types could safely play Trust in period 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 refer to the equilibrium separating gifts as g_h from the high type and g_l from the low type. These equilibrium gifts will mean

[26]Countries need to deliver concessions in period 0, not just promise them. A concession that is promised but not delivered will count as cheap talk.

that when period 1 is reached, countries know which type the other country is and play accordingly. The concessions phase acts as a coordination device to match countries' actions. To play different actions from each other 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 period 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 will 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, which, when simplified to allow $g_l = 0$ following Lemma 1 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

$$g_h \geq p(X^l_{FT} - X^l_{FF}) \tag{2}$$

in order to separate from the low type.

High type separating equilibrium utility is $U_h = pX_{TT} + (1-p)X_{FF} - g_h + pg_h$. Since 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 the definitions of X^l_{FT} and X^l_{FF} 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.

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

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]While 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 its type as being high.

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

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

Since $0 < p < 1$ 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 will 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

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

Rearranging,

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

i.e., 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 here, players will 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, which, when simplified to allow $g_l = 0$ following Lemma 1 is:

$$pg_h + X^l_{FF} \geq pX^l_{FT} + (1-p)X^l_{FF} - g_h + pg_h \tag{7}$$

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, the low types will have an incentive to mimic the high type and the separating equilibrium will be destroyed.

High type separating equilibrium utility is $U_h = pX_{TT} + (1-p)X_{FF} - g_h + pg_h$. Since 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 if the high types are 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.

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

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:

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

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 there is future material value.

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

That the gift in the model with material value, $g^1 = \frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}$, is increasing in p is straightforward. Recall that all the variables are assumed to be strictly positive, with $0 < \delta_l < 1$. It follows that the numerator is increasing in p and the denominator is decreasing in p and assumed positive. Therefore, g^1 is increasing 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 is unchanged in W since W does not enter the expressing anywhere. The other three parameters require working out the formal result.

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 $\frac{\partial g^1}{\partial \delta_l} > 0$.

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 $\frac{\partial g^1}{\partial T} > 0$.

Damages from negotiating partner playing Distrust:

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

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

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

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

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

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 there is 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 δ^1.

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

The patience threshold is increasing in p. That is,

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

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

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

We assume that $T > W - D$, so the fraction's numerator is positive, while the denominator is positive since 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,

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

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

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.

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

If $T < \frac{1-\delta_l}{1-\delta_l p}$ so that $\frac{\partial g^1}{\partial D} > 0$, this expression will always be 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 under 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. 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 proceeds 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*. Since 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 $W>\delta_l(T+D)$. Since $\delta_l<1$, 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 will invoke Expression 11 below. First we go back to Expression 10, simplifying and creating a common denominator for δ^1, we have

$$\begin{aligned}\frac{W}{T+D} \geq\; & \frac{(W-(D+T))\,(1-\delta_l)-(W-(D+T))\,(1-\delta_l)\,pT+(W-(D+T))(1-p)D}{(1-\delta_l)\,(D+T)}\\ &+\frac{p(W-T)\,(1-\delta_l)\,(D+T)+(1-p)D\,(1-\delta_l)\,(D+T)+(1-\delta_l)\,(D+T)}{(1-\delta_l)\,(D+T)}\end{aligned}$$

[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

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

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

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

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

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

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

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

To simplify, we have

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

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 proceed by showing 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 there are two reports of 'High.' Likewise, there are no gifts in the $(1-p)X^l_{FF}$ term on the right because the negotiating partner has made a report of 'Low.' This leaves

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

Expanding, we have

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

There are no gifts on the left-hand side, 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

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

Since we assume throughout that gifts must be non-negative, each of T, W and D are positive and $0 \leq \delta_l \leq 1$, 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

$$pX^h_{TT} + (1-p)X^h_{FF} - pg_h + peg_h \geq X^h_{FF} \tag{15}$$

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 there are two high types. This means that there will be no concessions when a high type lies. This implies that Expression 15 is also the individual rationality constraint since outside the mechanism they get the No Cooperation payoff and exchange no concessions.

Thus the constraint for both incentive compatibility and individual rationality is

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

This simplifies to

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

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

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

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 there is no 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.

■

REFERENCES

Allee, Todd L and Paul K Huth. 2006. “Legitimizing dispute settlement: International legal rulings as domestic political cover”. *American Political Science Review* 100 (2): 219–234.

Beardsley, Kyle. 2010. “Pain, pressure, and political cover: Explaining mediation incidence”. *Journal of Peace Research* 47 (2): 395–406.

Beardsley, Kyle and Nigel Lo. 2014. “Third-party conflict management and the willingness to make concessions”. *The Journal of Conflict Resolution* 58 (2): 363–392.

Beardsley, Kyle, David M. Quinn, Bidisha Biswas, and Jonathan Wilkenfeld. 2006. “Mediation style and crisis outcomes”. *The Journal of Conflict Resolution* 50 (1): 58–86.

Bercovitch, Jacob. 1996. *Resolving international conflicts: The theory and practice of mediation*. Lynne Rienner Publishers.

Bercovitch, Jacob. 2000. “Why do they do it like this? an analysis of factors influencing mediation in international conflicts”. *Journal of Conflict Resolution* 44 (2): 170–202.

Bercovitch, Jacob and Richard Jackson. 2001. “Negotiation or mediation? an exploration of factors affecting the choice of conflict management in international conflict”. *Negotiation Journal* 17 (1): 59–77.

Bercovitch, Jacob and Jeff Langley. 1993. “The nature of dispute and the effectiveness of international mediation”. *Journal of Conflict Resolution* 37 : 670–691.

Bester, Helmut and Karl Warneryd. 2006. “Conflict and the social contract”. *Scandinavian Journal of Economics* 108 (2): 231–249.

Camerer, Colin. 1988. “Gifts as economic signals and social symbols”. *The American Journal of Sociology* 94 : S180–S214. Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure.

Connelly, Brian L., S. Trevis Certo, R. Duane Ireland, and Christopher R. Reutzel. 2011. “Signaling theory: A review and assessment”. *Journal of Management* 37 (1): 39–67.

Dixon, William. 1996. “Third-party techniques for preventing conflict escalation and promoting peaceful settlement”. *International Organization* 50 : 653–681.

Duursma, Allard. 2020. “African solutions to african challenges: The role of legitimacy in mediating civil wars in africa”. *International Organization* 74 (2): 295–330.

Fey, Mark and Kristopher W. Ramsay. 2009. “Mechanism design goes to war: Peaceful outcomes with interdependent and correlated types”. *Review of Economic Design* 13 (3): 233–250.

Fey, Mark and Kristopher W. Ramsay. 2010. “When is shuttle diplomacy worth the commute?”. *World Politics* 62 (4): 529–560.

Fey, Mark and Kristopher W. Ramsay. 2011. “Uncertainty and incentives in crisis bargaining: Game-free analysis of international conflict”. *American Journal of Political Science* 55 : 149–169.

Fortna, Virginia Page. 2003. *Peace Time: Cease-Fire Agreements and the Durability of Peace*. Princeton: Princeton University Press.

Greig, J. Michael. 2005. “Stepping into the fray: When do mediators mediate?”. *American Journal of Political Science* 49 (2): 249–66.

Greig, J. Michael. 2005a. “Moments of opportunity: Recognizing conditions of ripeness for international mediation between enduring rivals”. *Journal of Conflict Resolution* 45 : 691–718.

Greig, J. Michael and Paul F. Diehl. 2006. “Softening up: Making conflicts more amenable to diplomacy”. *International Interactions* 32 : 355–84.

Hensel, Paul. 1996. “Charting a course to conflict: Territorial issues and interstate conflict, 1816-1992”. *Journal of Conflict Resolution* 15 (1): 43–73.

Horner, Johannes, Massimo Morelli, and Francesco Squintani. 2010. “Mediation and peace”. *The Review of Economic Studies* 84 : 1483–1501.

Kydd, Andrew. 2003. “Which side are you on? bias, credibility, and mediation”. *American Journal of Political Science* 47 (4): 597–611.

Maor, Cy and Eilon Solan. 2015. “Cooperation under incomplete information on the discount factors”. *International Journal of Game Theory* 44 (2): 321–346.

Meirowitz, Adam, Massimo Morelli, Kristopher W. Ramsey, and Francesco Squintani. 2012. "Mediation and strategic militarization". Unpublished manuscript.

Meirowitz, Adam, Massimo Morelli, Kristopher W. Ramsey, and Francesco Squintani. 2019. "Dispute resolution institutions and strategic militarization". *Journal of Political Economy* 127 (1): 378–418.

Prendergast, Canice and Lars Stole. 2001. "The non-monetary nature of gifts". *European Economic Review* 45 (10): 1793–810.

Spence, Michael. 1973. "Job market signaling". *The Quarterly Journal of Economics* 87 (3): 355–74.

Svensson, Isak. 2006. "Do mediators go where they are needed the most? mediation selection in civil wars: paper presented at the 2nd national conference on peace research, göteborg, september 7-8, 2006".

Terris, Lesley G and Zeev Maoz. 2005. "Rational mediation: A theory and a test". *Journal of Peace research* 42 (5): 563–583.

Van de Ven, Jeroen. 2000. "The economics of the gift". *Available at SSRN 244683*.

Waltz, Kenneth. 2018. *Man, the state, and war: A theoretical analysis*. Columbia University Press.

Watson, Joel. 1999. "Starting small and renegotiation". *Journal of economic Theory* 85 (1): 52–90.