

Inefficient Concessions and Mediation

Kristy Buzard

Syracuse University

kbuzard@syr.edu

Ben Horne

University of California, San Diego

January 21, 2021

Abstract

When two parties are engaged in conflict and each distrusts the other's ability to pursue peace, concessions can be used to indicate an interest in making peace. 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. We use mechanism design to explore ways in which a third party mediator can act as a guarantor that promised concessions will be delivered, thereby reducing inefficiencies and increasing the potential for peace. In this process, we open up a new rationale for mediation: to remove the inefficiencies of signaling in a preliminary round of negotiations.

1 Introduction

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 historic or religious importance, Israel values 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 them. 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 it 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, the most efficient concessions often are not made.

This holds for concessions other than land. Russia is loathe to cede further autonomy to Chechnya, even though it could hypothetically relieve tensions and reduce violence. If Chechen autonomy could in fact lead to lowered violence, it is possible that this move would be in Russia's best interest. But the Russians perceive, perhaps correctly, that any move towards autonomy might only give Chechens further leverage against the Kremlin. From the Russian perspective, granting the concession of increased autonomy might have large future costs.

This paper uses a two-sided signaling model to propose a new theoretical explanation for why inefficiency might occur in a conflict setting: inefficient concessions may be preferred if the most efficient concessions are seen as having additional, future costs.¹ The Golan Heights, for example, will not be conceded if Israel foresees future security risks that outweigh the immediate benefits of peace. We then suggest a new explanation for the effectiveness of third-party mediation. Using a mechanism design framework, we demonstrate that a mediator can

¹The type of inefficient concessions addressed here are those whose value to the recipient are lower than the cost to the giver.

facilitate more efficient concessions by removing uncertainty about the ability of the parties to commit to peace.

While an integral part of the literature on conflict, concessions have been only partially addressed by theoretical research. The broader literature that attempts to model war has included concessions as part of bargaining, but has not explicitly considered the role of concessions in achieving peace. Schelling's *Arms and Influence* (1966) modeled war with a game theoretic framework, and Powell (1999) has also defended the use of modeling in understanding war. Fearon (1996), Powell (1996a and 1996b, 2006) and others consider war to be a bargaining failure, caused by commitment problems and asymmetric information. When the assets that are the object of the bargaining confer future power, a lack of credible commitment, or trust, can create a holdup problem that prevents the optimal level of cooperation (Fearon 1996).

In these models war can be averted by splitting the pie (i.e. the disputed territory) in a way acceptable to both parties. Concessions are modeled as offering part of the pie to the other party, and asymmetric information is about each party's *resolve* to fight and/or military capabilities (Arena 2013). In this context, an example of costly signaling would be military preparations for war (Slantchev 2011).

We will approach the commitment problem through this same general understanding, but we model a party's ability to cooperate as private information. A high type is able to 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 in certain scenarios. 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 may be a low type, not because he does not want to make peace, but because he cannot credibly act on behalf of the entire Palestinian population.

In conflict resolution, concessions may act as signals or 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.

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. And the gift giving literature, more so than the concessions literature, focuses specifically on the issue of efficiency. Inefficiency is defined for the purposes of this paper as giving a gift or concession that is more costly to the giver than it is valuable to the recipient, that is, using a concession as a signal and not taking advantage of its full inherent value.

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 efficient gifts that are specifically requested or selected from a registry. Since inefficient gift giving is more likely involved in nascent relationships and partnerships, 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. Amongst these explanations are altruism, social mores, and egoism. But the most relevant explanation for the field of international relations is that gift giving is strategic. Camerer also uses mechanism design to model gifts as costly signals; inefficiency is useful in pre-play communication in order to form relationships between like types. This strategic giving of gifts in societies looks very much like the strategic giving of concessions to remedy conflict situations. When modeled game-theoretically, these different environments look even more similar.

Camerer's model includes players that have one period to signal before they decide

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 (or send a gift of zero cost). 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.² In order to explain inefficiency, Camerer adds in an additional pre-play period where players must pay a cost in order 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 are saved the cost of sending a signal to low types in the main game. High types would, in some parameterizations, be given incentives to give inefficient equilibrium gifts in the main game in order 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 and this can raise the expected payoff for a high type by saving on the gifts given.

The second major explanation for inefficient gifts is expressed in Prendergast and Stole (2001). The authors explain inefficient gift giving by modeling the utility of matching. That is, if players are of a like mindset they get an extra benefit so they are willing to take the risk of giving the wrong gift.

We propose a different conceptual framework to explain this real-life phenomenon that is more rooted in a realistic international relations puzzle. The basic formulation of our mechanism-design approach is an expansion of Camerer’s (1988) analysis. The model can achieve the basic ‘inefficient gifts’ results of Camerer (1988) and Prendergast and Stole (2001) but does so using a very different explanation for inefficiency. The basic model is a costly signaling model, at its root similar to ideas in Spence (1973) and subsequent papers.³ In essence, the model shows that if concessions can be later used against the party that gives,

²In an environment with repeated opportunities for concessions Watson (1999) is able to show 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.

³See Connelly et al. (2011) for a review of the costly signaling literature.

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 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 intensely debated. Some authors have concluded that mediation has little impact, (Bercovitch 1996; Bercovitch and Langley 1993; Fortna 2003) while others insist that a correct analysis finds mediation playing a positive role in resolving conflict (Dixon 1996; Beardsley et al. 2006). But 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) and timing is important (Greig 2005a).

It is essential to focus more precisely on the specific actions and environments that make use of particular types of mediation, which is how we approach the question. This paper models an environment where the primary issue is trust, not uncertainty about an opponent's capabilities, and where a mediator is employed who has manipulative abilities. Mediation rarely occurs in the absence of violent conflict (Bercovitch 1996; Beardsley 2006). In this paper, parties are in conflict and can benefit from seeking mediation in order to overcome mistrust. Parties are more likely to seek mediation when they are entrenched in costly conflict (Bercovitch and Jackson 2001; Greig 2005b; Greig and Diehl 2006; Terris and Maoz 2005; Svensson 2008). Alternatively, parties may seek a mediator to justify difficult concessions and avoid angering domestic constituencies (Allee and Huth 2006, Beardsley 2010).

These empirical results have until recently lacked strong theoretical foundations. Mediation is used, 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 2008,

2010 and 2011; Horner et. al. 2010; Meirowitz et al. 2012, 2015). This paper expands the theoretical literature by modeling the role of mediation when information is asymmetric about parties' abilities to commit to peace. Here, a mediator can remove the incentive to make concessions that are inefficient.

2 Model

Countries⁴ 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 as a player whose only sequentially rational strategy is to remain in conflict in all periods. It is not necessary to think of a type as fixed; it can change as regimes and relations between countries change.⁵

The probability of nature selecting the high type for a given country is p . The probability of nature selecting the low type is $1 - p$.⁶ 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 – when trust is the issue to be solved, there are several solutions to conflict. Concessions are one way in which countries can signal their commitment. By modeling the conflict in such a way we can also gain insight as to 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

⁴Players 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.

⁵We 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.

⁶Note 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.

$i \in \{1, 2\}$. 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 $C(g_i) = g_i$.

Beginning in period 1, countries engage in an infinitely repeated Prisoner's Dilemma once per period with stage game payoffs in Table 1 below. 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, continued conflict. 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.⁷

The countries' stage game actions will be referred to as "Trust" and "Fight" 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	Fight
Trust	T, T	$-D, T + W$
Fight	$T + W, -D$	$W - D, W - D$

Here $T > 0$ represents the benefit from the other country playing Trust. $D > 0$ represents the damages incurred when the other country plays Fight, while $W > 0$ is the additional benefit a country receives when it plays Fight itself. We assume $T > W - D$ in line with the assumption that the countries play a Prisoner's Dilemma in the stage game.

We will refer to the stage-game Nash Equilibrium of (Fight, Fight) as "War." We will refer to (Trust, Trust) as "Peace." The assumption $T > W - D$ ensures that the joint payoffs from "Peace" are the highest among all the stage-game action profiles.

⁷This payoff matrix does not distinguish between civil and interstate conflict, though in practice the effects on the payoffs might indeed have a different structure in the different cases.

Payoffs for a country are the sum 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, (Fight, Fight) will be played forever if Fight is played by any party in any round. Because we use grim trigger punishments, there are only 4 interesting paths of play after concessions are given: 1) Both countries play Trust always, 2) Both play Fight always, 3) Country 1 plays Fight in round 1 while country 2 plays Trust, which is followed by both countries playing Fight from round 2 onwards and 4) Country 2 plays Fight in round 1 while country 1 plays Trust, which is followed by both countries playing Fight from round 2 onwards.

Equilibrium payoffs beginning from round 1 are now easily described by noting only the collective first round behavior. 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 $X_{TT} = \frac{T}{1-\delta_i}$. Likewise $X_{TF} = (-D + \frac{\delta_i(W-D)}{1-\delta_i})$, $X_{FF} = \frac{W-D}{1-\delta_i}$, and $X_{FT} = T + W + \frac{\delta_i(W-D)}{1-\delta_i}$. Note that because types have different discount rates, the payoffs differ by type, and when ambiguous will be superscripted to indicate the type (e.g. X_{FT}^h represents payoffs for the high type of country 1 when country 1 fights and country 2 trusts in the first period).

Parameters are common knowledge with the exception of δ_i , which is country i ’s private information. We will take the measure of social welfare, and thus the determinant of the optimal equilibrium, to be the sum of participating high types’ expected utilities.

3 Analysis of Benchmark Cases

When countries are faced with uncertainty about each other’s motives, they can proceed in several different ways, including giving concessions in order to signal their willingness and ability to make peace. We will examine three types of equilibria.⁸ The Appendix provides a

⁸If the discount factor is high enough, there are many types of equilibria. Consider for instance oscillating every other period between (Trust,Trust) and (Fight,Fight) with a grim trigger War threat. In this paper,

thorough analysis of the various equilibria as well as most proofs for this section.

3.1 Pooling equilibrium

In the *pooling equilibrium*, both countries do not give concessions in period 0, then always play Fight in every period beginning with period 1. Countries continue to doubt, doing nothing to bridge the credibility gap, and thus not cooperating. This *pooling equilibrium* is likely to happen when countries are in long-set patterns of distrust. Active fighting is not necessary; remaining outside of a state of peace is all that is required. See the Appendix for a formal characterization of this type of equilibrium.

Cyprus' civil war is a good example of the pooling equilibrium where neither side is taking actions to move things forward. While both sides posture that they want peace, and while there has been some lessening of tensions, very few concrete actions have been taken that move the sides toward settlement. This conflict is a stalemate.

Sri Lanka is another example. While the government has finally won the war of secession, it still refuses to address grievances of the minorities and leaves the conflict still unresolved. Other prominent examples of this pooling equilibrium include Israel and its neighbors, North and South Korea, Greece and Turkey, and a multitude of other long-lasting conflicts. The mistrust of the other party is perhaps justified; indeed if one party did take the first step and give a concession first, they may be taken advantage of.

Without a third party, there is little hope of overcoming mistrust. These long lasting stalemates are the types of conflict which are most appropriate for mediator involvement. A mediator will be added to the model later to illustrate exactly how mediation can help achieve peace in this context. In some sense, these are the cases that are most interesting because the international community might play a positive role.

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.

3.2 No-Concessions Separating Equilibrium

The other two equilibria on which we focus are separating equilibria in which low types and high types have different strategies. If δ_i is high enough for both countries, the “Peace” 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 Fight. So long as there are two high types, as revealed by their period 1 play, both countries continue to play Trust in all following periods. Otherwise, both countries play Fight in all following periods.

Theorem 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 Fight and the high types give no concessions in period 0 and play Trust unless the other country defects to Fight.

Proof: See Appendix.

Here, as in each further iteration of the model, we define a high type as a country who prefers to play the “Peace” equilibrium, i.e. $\delta_h \geq \delta^{nc}$. Any country with $\delta_l < \delta^{nc}$ is a low type. This means that two high types—and only two high types—can sustain the Peace outcome in a simple repeated game.

In this equilibrium, two mutually distrustful countries behave as if they trust anyway, 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.

Examples of a *no-concessions separating equilibrium* are most likely to be those where countries have traditionally had good relations but something has recently come in the way. Certain relationships between the USA and countries that its wikileaks alienated – such as Italy – would be an example. Italy essentially laughed off the criticisms and reassured the

USA of its commitment to their relationship. Warsaw Pact countries generally allowed the USSR similar leeway; it was better to have occasionally imperfect cooperation than conflict. Some events can disrupt trust without completely derailing cooperation.

Of course, it is possible that one country can take advantage of the other in a *no-concessions separating equilibrium*. It can be argued that indeed that did happen in the Warsaw Pact cases; the Eastern Bloc countries were first exploited before the fall of the iron curtain; eventually this led to a steely relationship between many of those countries and Russia that exists to this day in Poland, Lithuania and other countries.

3.3 Concessions Separating Equilibrium

The other separating equilibrium of interest involves distrustful countries to 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*. The concessions are given before more substantive actions, thus concessions are a costly signal that the other country is serious about cooperating. Most agreements come with concessions, and in many cases, though not all, they are used as signals. The goal is to use this option to avoid continual distrust. That is, high types make use of concessions to identify themselves as trustworthy partners in hopes of establishing Peace.

In the *concessions separating equilibrium*, low types do not give concessions while high types give a concession in period 0. 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 Fight in subsequent periods. Because of the signaling value of the concessions given in period 0, starting from period 1 in the Trust/Fight game, both countries play the same strategy in a separating equilibrium, that is both play Trust if there were two high signals and otherwise both play Fight.

Details of the analysis, which uses standard incentive compatibility constraints, 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.

Theorem 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 Fight and the high types give a concession in period 0 and play Trust unless the other country defects to Fight.*
- (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. 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 3.2.

Corollary 1 *The no-concessions separating equilibrium requires countries to be more patient than the 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 cheater’s 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 unavailable or unattractive. By giving costly concessions, they signal their intentions in a way that a partner unwilling to make peace would not be willing to signal.

One example of such an equilibrium 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 5 below.

4 Material Value of Concession and Inefficiency

4.1 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 2 and 3, 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 this is an additional value from the gift in each period that the repeated game is played, that is, Period 1 and onward.

The interpretation is that the country which receives the concessions can use them to build up their military or to build up their civil society. Using concessions to build military strength hurts the giver of the concessions in war; using the concessions to strengthen civil society helps the giver in peace. The Golan heights is such an example: if Syria is an ally of Israel, the Golan Heights is a reasonable concession to make, but if Syria later turns out to betray a trust, surrender of the strategic land would be disastrous for Israel.

We add two elements to the benchmark model in Section 2 in order to incorporate future material value. The first is an additional decision for the negotiating parties. In Period 0, after receiving a concession, each country makes a decision about how to allocate the value of the concession 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. They 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	Fight
Trust	$T + T\alpha_2g_1,$ $T + T\alpha_1g_2$	$-[D + D(1 - \alpha_2)g_1],$ $[T + T\alpha_1g_2] + [W + W(1 - \alpha_2)g_1]$
Fight	$[T + T\alpha_2g_1] + [W + W(1 - \alpha_1)g_2],$ $-[D + D(1 - \alpha_1)g_2]$	$[W + W(1 - \alpha_1)g_2] - [D + D(1 - \alpha_2)g_1],$ $[W + W(1 - \alpha_2)g_1] - [D + D(1 - \alpha_1)g_2]$

This modification to the payoffs is the second change in the model relative to Sections 2 and 3. 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 that was invested in civil society as well as the base benefit of T .

If both countries play Fight, Country 1 (symmetric for Country 2) receives an additional benefit term $W(1 - \alpha_1)g_2$, which accounts for the value of the received concession as well as how much Country 1 invested in the military. However, it also incurs extra damages $D(1 - \alpha_2)g_1$ 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 (Fight, Trust) and (Trust, Fight) payoffs are modified analogously.

Because α_i enters into the payoffs in a linear fashion, the choice that maximizes welfare is $\alpha = 1$ for high types since a concession would only come from another high type in equilibrium, and thus only the “Peace” outcome will be played. Conversely, in equilibrium, low types only participate in the “War” equilibrium and thus maximize their payoffs by choosing $\alpha = 0$.

Theorem 3 Concessions Separating Equilibrium (CSE) with Material Value

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

Corollary 2 Assume the high types are not patient enough to separate without concessions but can separate using concessions when concessions have no material value. When concessions instead have future value that the low-type country can use to reduce the future payoffs

of its negotiating partner, there are parameters under which the negotiating partners' ability to separate through concessions is destroyed, i.e. peace cannot be achieved when there are two high types.

Proof: An example suffices as proof. Let $T = 10$, $W = 8$ and $D = 5$. 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 $\delta = \frac{8}{10+5} = .5\bar{3}$. Let $\delta_h = .6$ and $\delta_l = .3$. Assume $p = .2$. By Theorem 2, separation through concessions is possible.

The patience threshold to separate through concessions is much higher under these parameters when the concessions carry future material value. The minimum patience level required for separation is $\delta_h^1 = \frac{p(W-D-T)}{(1-\delta_l)p(D+T)} + p(W-T) + (1-p)D + 1 = 2.4$, with a minimum separating concession of $g^1 = \frac{p(1-\delta_l)(D+T)}{(1-\delta_l)(1-pT)+(1-p)D} = .6\bar{3}$. Since discount factors can be no larger than 1, we have shown that under this set of parameters, the concessions separating equilibrium when concessions have material value cannot be sustained.

δ_h is not involved in any of these calculations, so we can see that for any $\delta_h \in [.5\bar{3}, 1)^9$ Thus, separating through concessions is possible when concessions have no material value but is not possible when they do have future material value. ■

4.2 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 example, an expensive display of pageantry to a dignitary might benefit the receiving country very little, if at all, but it still signals good intentions on behalf of the giving country. The option to give efficient concessions is still available along with new options to give a variety of less efficient concessions.

We implement this by allowing countries to choose a scalar e to multiply by their given concession g . $0 \leq e \leq 1$. As before, $C(g) = g$ but while the receiving country observes the

⁹Recall that we restrict $p < 1$ since there is no imperfect information when only one type is possible.

full concession g , its benefit is now only eg .¹⁰

Table 3: Stage game payoffs when concessions have material value that can be destroyed

	Trust	Fight
Trust	$T + T\alpha_2eg_1,$ $T + T\alpha_1eg_2$	$-[D + D(1 - \alpha_2)eg_1],$ $[T + T\alpha_1eg_2] + [W + W(1 - \alpha_2)eg_1]$
Fight	$[T + T\alpha_2eg_1] + [W + W(1 - \alpha_1)eg_2],$ $-[D + D(1 - \alpha_1)eg_2]$	$[W + W(1 - \alpha_1)eg_2] - [D + D(1 - \alpha_2)eg_1],$ $[W + W(1 - \alpha_2)eg_1] - [D + D(1 - \alpha_1)eg_2]$

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

Lemma 2 *Assuming that a separating equilibrium exists when concessions have material value and are efficient (where $e = 1$), it is optimal to give efficient concessions in a separating equilibrium.*

Proof: The benefit of a gift appears on both sides of the IC for both types and thus cancels out in the low type IC ($peg_h + X_{FF}^l \geq pX_{FT}^l + (1 - p)X_{FF}^l - g_h + peg_h$) and in the high type IC: ($pX_{TT}^h + (1 - p)X_{FF}^h - g_h + peg_h \geq X_{FF}^h + peg_h$).¹¹ Therefore, receiving a gift where $e < 1$ does not affect incentives to truth tell about the countries' types. However, the expected separation utilities under this game are now

$$U_h = pX_{TT} + (1 - p)X_{FF} - g_h + peg_h \quad (1)$$

and

$$U_l = X_{FF} + peg_h \quad (2)$$

¹⁰In a more general model, we could also think of e as a function $E(\cdot)$. The argument is still g , but instead of a scalar, $E(\cdot)$ can transform g into any concession, not only a numerical value. These gifts will be converted into utilities by cost and benefit functions, which then allow the possibility of a gift having efficiency greater than 1. The transformation function $E(\cdot)$ can be thought of as the type of gift, and has a cost to the sender and a direct benefit to the receiver in addition to its signaling value.

¹¹Elaboration and explanation of the IC constraints is in the proof of Theorem 2.

$\frac{dU}{de} > 0$ for both countries so it is never optimal have $e < 1$.¹² ■

We next turn to our main result for the case where the material value of concessions can be destroyed...

Theorem 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: See Appendix.

Inefficient concessions are optimal under low p (there's a good chance of facing a low type who will use the concession against you) as well as when δ_l is low (g^* is decreasing in the patience of the low type, as a larger concession is required to deter an impatient low-type from being untruthful).

This finding of inefficiency is for profoundly different reasons than the literature related to Camerer (1988) or Prendergast and Stole (2001).¹³ Instead of being in reaction to a behavioral regret for mismatching or serving as a way to discourage low types for entering the game altogether, here making concessions inefficient is a device to prevent low types from using those concessions against oneself in the future.

5 Mediation

All results in this paper thus far have relied on bilateral engagement. We will now consider the involvement of a mediator.

¹²This result does not depend on the benefit so it holds for any monotonic benefit function. It also holds for any monotone cost function.

¹³Since our initial setup is essentially identical to Camerer's, the desired inefficiency of pre-play communication can also be found in our model. 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, high types and low types are used so there is an incentive for low types to try to pose as high types. Aside from this difference, essentially the same result is found: 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.

Because many conflicts occur in an international arena with no clear enforceable rule of law (Waltz, 2002), much of the literature has focused on self-enforcing mechanisms. If however a credible third party does exist, 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 1997). We will now introduce a mediator who does the latter.¹⁴

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. The Revelation Principle tells us that it is without loss of generality to restrict attention to direct mechanisms, that is, mechanisms in which countries simply declare their types (Myerson 1979). This mediator will be modeled as a mechanism M .

After learning their type, each country chooses to participate in the mechanism or not; that is, whether to send a message that declares its type to be High or Low. By the revelation principle, it is again without loss of generality to consider equilibria in which both countries participate and send truthful reports. Therefore we have a mechanism M that 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 $M : (t_1, t_2) \rightarrow (e_1, g_1; e_2, g_2)$. The countries are bound to send the mandated concessions.

In Sections 3 and 4, the same concession must be sent to both types since types are private information. Importantly, here the mechanism M can differentiate based on recipient’s type. Otherwise, the setup is the same as in Section 4.2 where concessions have future material value.

[both individual rationality and incentive compatibility constraints must be satisfied, leading to truthful self-identification.]

[Further, it can be shown that restricting attention to symmetric recommendations is

¹⁴Here 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, we do not assume that the mediator can do either because it is not necessary.

without loss of generality, because the optimal arbitration programme is linear.]

low type IC

$$X_{FF}^l \geq pX_{FT}^l + (1-p)X_{FF}^l - pg_h + peg_h$$

high type IC:

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

The welfare target is again the sum of the high type utilities. Qualitatively, similar results would hold if we also included low types in the analysis with weighting based on p .

Theorem 5 *Assume that concessions have future material value.*

- (a) *If future material value can be destroyed, a mediator can be used to eliminate inefficient concessions and improve welfare.*
- (b) *If future material value cannot be destroyed, a mediator can be used to expand the parameter space over which peace can be achieved.*

Proof: (a) Under M , the low type incentive constraint can bind without having high types give a concession to low types. There is no need then to give inefficient concessions: high types will invest in civil society and improve the giver's welfare and low types will never receive the concession. High type utility under M is

$$U_h^M = pX_{TT} + (1-p)X_{FF} = p \left(\frac{T(1+g^{*M})}{1-\delta_h} \right) + (1-p) \left(\frac{W-D}{1-\delta_h} \right) \quad (3)$$

because high types exchange concessions of the same value. Welfare, defined as high-type utility, is improved because no transfer is made to the low types and no future damage is caused in the case of such a transfer.¹⁵

(b) The contraction of the parameter space over which peace could be achieved in Theorem 2 derives from a low type's ability to use concessions to reduce the payoffs of its

¹⁵Low types IR constraint still binds and $U_l = X_{FF}$.

negotiating partner during a ‘fight’ stage. By ensuring that only high types exchange gifts, the mediator removes this expected cost, thus allowing a wider range of high types to separate through concessions. The threshold for high types to prefer this separating-through-mediation equilibrium to the pooling on no concessions equilibrium, δ_h^* , differs from δ_h^1 of Theorem 2 in that it does not have the term $D(1 - p)$ (thus is smaller) and the separating concession $g^* = \frac{g^1}{1-p}$. Since the concession is in the denominator, this also reduces the value of δ_h^* . ■

The low type leader can bluff in the bilateral setting but not when a mediator is present. As per Theorem 4, inefficiency is due to the high types giving an inefficient gift in order to separate while still being able to prevent low types from using concessions against the high types. But because the benefits of giving a concession differ between the types (low types get one period of cheater’s payoff, while high types get future peace), the mediator is able to mandate concessions only between declared high types. This means there is no longer a need for inefficient concessions since concessions will never be used against the giver. Thus, a mediator is able to remove inefficiency in concession giving and bring willing parties to the table to give the gifts necessary to achieve peace.

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.

The value of truth telling by the mediator has been examined in the literature. Kydd (2003) studies 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 be able to improve on unmediated interaction.¹⁶ In our model, a successful mediator relies on the parties trusting her. However, here the mediator had no incentive

¹⁶Horner et al. (2010) take cheap talk communication as their benchmark of unmediated communication, whereas we take the unmediated signaling game as our benchmark.

to lie: lying can only potentially hurt the high types, the only type who the mediator can possibly help.

6 Conclusion

In states of perpetual conflict, concessions have been shown in the literature to ease distrust. Our paper shows 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 is essentially a state of war for all types. 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, the most intuitive or realistic type of equilibrium, a concessions giving separating equilibrium exists where high types offer concessions to eliminate the distrust.

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 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 really do come in sizes, for example the number of prisoners released, an amount of land ceded, or an 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 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.

In practice, in “sons of the soil” conflicts, concessions are indivisible (Walter 2002); if inefficient concessions need to be given to achieve peace, in the absence of a mediator no

concessions will be given as an inefficient concession is an impossibility. The full value is the only concession worth making because it is the only one that will be accepted: although countries may prefer to give inefficient concessions, they obviously do not prefer to receive them. Inefficient concessions can have the same signaling value as efficient concessions, but unless signaling is the only facet of a concession that is valued, this is not necessarily enough to achieve peace.

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 peace when efficient concessions cannot and thus also when a mediator can restore 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 the mechanisms through which a mediator can help. Since the mediator eliminates inefficient concessions, the observation that inefficient concessions may not do the trick only underscores the potential value of a mediator.

This model uses manipulative mediation, which is indeed used in resolving conflicts. However, the usefulness of mediation relies on perfect trust of the mediator. 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 also be noted that mediation is not free, at least to the providing (third) party. Therefore, the inefficiencies borne by the inefficient concessions and/or by the ongoing conflict must not be outweighed by the cost of providing the mediation. 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 when efficient concessions are not sufficient to do so; it also demonstrates how a mediator can achieve peace in these situations without resorting to inefficient concessions. Opportunities for further research clearly 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.

7 Appendix

Pooling equilibrium

Lemma 3 *From period 1 on, playing fight in all periods regardless of beliefs constitutes an equilibrium strategy of the continuation game for both types.*

Proof: As $D \geq 0$ and $W \geq 0$ the dominant stage game action for all types is to play Fight. Thus the stage game equilibrium is (Fight, Fight) 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—who always play “Fight” by definition—and always play Fight. If both the low and high type of player j choose the strategy that gives no concessions and always plays fight, then either type of player i is made strictly better off by playing Fight in each period given player j ’s behavior. Given that the stage-game equilibrium will be (Fight,Fight), 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 Fight.

No-Concessions Separating Equilibrium

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

Proof of Theorem 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 Fight in the first period and then every period thereafter by the one-shot deviation principle. If both countries play Trust in a period, a Peace equilibrium takes place and both countries continue to play peace in every period with the grim trigger threat of the War equilibrium. Low types play Fight in the first period, so a high type that observed Fight being played would respond by playing Fight in period 2 and in all future periods. A country plays Trust if and only if the following condition is satisfied:

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

Simplifying, we have

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

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

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

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

If the potential advantages of Peace are large enough to outweigh the risks of encountering a low type in expected utility terms, that is $\delta_h \geq \frac{W}{(1 - p)W + p(T + D)} = \delta_h^{nc}$, a country has the incentive to play Trust in the first round rather than Fight. Any player with $\delta < \delta_h^{nc}$ has the incentive—by this same inequality—to play Fight 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 4 is not met, a separating equilibrium cannot be achieved with no concessions in period 0.¹⁷ 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 War 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 that when period 1 is reached, countries know which type the other country is and choose to 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, we focus on concessions separating equilibrium 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 Fight with low types. Low types play Fight (Lemma 3) 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 distinguishes them from the high types *and* this 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 Peace equilibrium, which can

¹⁷Countries need to deliver concessions in period 0, not just promise them. A concession that is promised but not delivered will count as cheap talk.

only be sustained by a pair of high types. Both types benefit from the other country playing Trust so all countries have 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 Theorem 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:

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

This expression represents the low types' incentives by 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 give a gift

$$g_h \geq p(X_{FT}^l - X_{FF}^l). \quad (5)$$

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 and still have separation. This occurs when Expression 5 holds with equality. Substituting in the definitions of X_{FT}^l and X_{FF}^l yields the minimum separating gift:¹⁸

$$g_h^* = p(T + D) \quad (6)$$

¹⁸While 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.

In order to incentivize high types to send nonzero concessions, the high type incentive constraint must not be violated. That is, the payoff from separating must be higher than the payoff from pooling with the low types.

$$pX_{TT}^h + (1-p)X_{FF}^h - g_h + pg_h \geq X_{FF}^h + pg_h \quad (7)$$

Combining Expressions 6 and 7, we have the condition for a concessions separating equilibrium to exist:

$$pX_{TT}^h - pX_{FF}^h - p(X_{FT}^l - X_{FF}^l) \geq 0 \quad (8)$$

In terms of fundamentals, the condition for a concessions separating equilibrium to exist is:

$$\delta_h \geq \frac{W}{T+D} \quad (9)$$

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

Proof of Corollary 1:

To show that the patience threshold for the *no concessions separating equilibrium* is greater than the patience threshold for the *concessions separating equilibrium*, we start with the assumption

$$T + D > W.$$

We can 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

and so we can divide through by both and multiply both sides by W :

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

In the case that both types of equilibria exist, i.e. $\delta_h \geq \delta_h^{nc}$, the high types will prefer the *concessions separating equilibrium* if and only if

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

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

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

Rearranging,

$$\frac{W}{T+D} \geq p \tag{10}$$

i.e. The high types prefer the *concessions separating equilibrium* when p is smaller than $\frac{W}{T+D}$.

■

Concessions Separating Equilibrium with Material Value

Proof of Theorem 3:

In the separating equilibrium constructed here, players will believe that any concession not equal to the equilibrium concession of the high type g^1 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:

$$pg^1 + X_{FF}^l \geq pX_{FT}^l + (1-p)X_{FF}^l - g^1 + pg^1 \tag{11}$$

The α_i 's vary depending on the situation. On the left side, the low type does not give

a concession but receives a concession if facing a high type (pg^1). Then (Fight, Fight) is played, and the payoffs depend on whether this low type is facing a high type or a low type because the implications for future material value are different. With probability p , the low type received a gift and invested in the military. With probability $(1 - p)$ there were no concessions given by either side. The left side of Expression 11 is then $pg^1 + \frac{1}{1-\delta_l} [p(W + Wg^1 - D) + (1 - p)(W - D)]$.

On the left side, the low type is giving a gift and also receiving 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 low type does not receive a gift. So the right hand side of Expression 11 becomes $-g^1 + pg^1 + p \left(T + Tg^1 + W + Wg^1 + \frac{\delta_l}{1-\delta_l} (W + Wg^1 - D) \right) + (1 - p) \left(\frac{1}{1-\delta_l} (W - D - Dg^1) \right)$.

Substituting into Expression 11 and simplifying a bit, we have

$$\begin{aligned}
g^1 + p(W + Wg^1 - D) &\geq p(T + Tg^1 + W + Wg^1) + (1 - p) \left(\frac{-Dg^1}{1 - \delta_l} \right) \\
(1 - \delta_l) g^1 (1 - pT) &\geq (1 - \delta_l) p(D + T) - (1 - p)Dg^1 \\
g^1 &\geq \frac{(1 - \delta_l) p(D + T)}{(1 - \delta_l)(1 - pT) + (1 - p)D} = g_h^1
\end{aligned} \tag{12}$$

This implies that the high types need give a gift at least as large as g_h^1 in order to separate from the low type; that is, if the high types give a gift smaller than g_h^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 12 holds with equality. That is, the optimal high type gift in this equilibrium is $g^1 \geq \frac{(1-\delta_l)p(D+T)}{(1-\delta_l)(1-pT)+(1-p)D}$.¹⁹

¹⁹While 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

In order to incentivize high types to send nonzero concessions, the high type incentive constraint must not be violated. 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 be playing Fight and so invests any concession received in military buildup.

$$\frac{1}{1 - \delta_h} \{pT(1 + g^1) + (1 - p)(W - D(1 + g^1))\} - (1 - p)g^1 \geq \frac{1}{1 - \delta_h} [p(W + Wg^1 - D) + (1 - p)(W - D)] + pg^1$$

Simplifying, we have

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

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

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

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

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

Substituting the minimum efficient concession from Expression 12, we have the condition for a concessions separating equilibrium to exist in terms of fundamentals:

$$\delta_h \geq \frac{p(W - D - T)}{\frac{(1 - \delta_l)p(D + T)}{(1 - \delta_l)(1 - pT) + (1 - p)D}} + p(W - T) + (1 - p)D + 1 \quad (13)$$

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

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.

Mediation

Proof of Theorem 4:

High types choose $\alpha = 1$ (because a concession would only come from another high type in equilibrium) and low types choose $\alpha = 0$. Because the low type's IC constraint binds, with $e = 1$ the optimal concession can be calculated to be $g^1 = \frac{p(1-d_l)(D+T)}{(1-d_l)(1-pT)+(1-p)D}$. The high type's utility when $e = 1$ is then $\frac{1}{1-d_h} \{pT(1+g^*) + (1-p)(W - D(1+g^*))\} - g^1$.

When $e = 0$ —that is, the future material value of the concessions is destroyed—we are essentially back in the case of Section 3 and the minimum separating concession is $g^0 = p(T + D)$. The threshold for high types to cooperate is again $\frac{W}{T+D}$. High type utility when $e = 0$ is $\frac{1}{1-d_h} \{pT + (1-p)(W - D)\} - g^0$.

Whether it is beneficial for high types to settle on the efficient concession or the inefficient concession is determined by comparing the utility under each, and either can be larger. Inefficient concessions are made optimal when p is low (there's a good chance of facing a low type who will use the concession against you) as well as when δ_l is low (g^1 is decreasing in the patience of the low type, as a larger concession is required to deter an impatient low-type from being untruthful). ■

8 Bibliography

- Allee, Todd L. and Paul K. Huth. (2006). "Legitimizing Dispute Settlement: International Legal Rulings as Domestic Political Cover." *American Political Science Review* 100:2, 219-34.
- Arena, Philip. (2013). "Costly Signaling, Resolve, and Martial Effectiveness." Unpublished manuscript.
- Beardsley, Kyle. (2010). "Pain, Pressure, and Political Cover: Explaining Mediation Incidence." *Journal of Peace Research*. 47:2, 395-406.
- 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, ed. (1996). *Resolving International Conflicts: The Theory and Practice of Mediation*. Boulder, CO: Lynne Rienner.
- Bercovitch, Jacob. (1997). "Mediation in International Conflict," in I.W. Zartman and J.L. Rasmussen (eds.), *Peacemaking in International Conflict*, Washington DC: US Institute of Peace.
- 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), 231-49.

- Camerer, Colin. (1988). "Gifts as Economic Signals and Social Symbols." *The American Journal of Sociology*, Vol. 94, Supplement: Organizations and Institutions: Sociological and Economic Approaches to the Analysis of Social Structure: S180-S214.
- Carmichael, H. Lorne and W. Bentley MacLeod. (1997). "Gift Giving and the Evolution of Cooperation." *International Economic Review* 38:3, 485-509.
- 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.
- Fearon, James. (1996). "Bargaining Over Objects that Influence Future Bargaining Power." *Chicago Center on Democracy Working Paper*, University of Chicago.
- 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, 529-60.
- 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. (2005a). "Moments of Opportunity: Recognizing Conditions of Ripeness for International Mediation Between Enduring Rivals," *Journal of Conflict Resolution* 45, 691-718.

- Greig, J. Michael, (2005b). "Stepping into the Fray: When Do Mediators Mediate?" *American Journal of Political Science* 49:2, 249-66.
- Greig, J. Michael and Paul F. Diehl. (2006). "Softening Up: Making Conflicts More Amenable to Diplomacy," *International Interactions* 32, 355-84.
- Horner, Johannes, Massimo Morelli and Francesco Squintani. (2010). "Mediation and Peace." *The Review of Economic Studies* 84, 1483-1501.
- Iannaccone, Laurence. (1992). "Sacrifice and Stigma: reducing free-riding in cults, communes, and other collectives." *Journal of Political Economy* 100:2, 271-291.
- Kranton, Rachel (1996). "The formation of cooperative relationships." *Journal of Law, Economics, and Organization*, 12:1, 214-233.
- Kydd, Andrew. (2003). "Which Side Are You On? Bias, Credibility, and Mediation." *American Journal of Political Science* 47:4, 597-611.
- Larson, Deborah Welch. (1987). "The psychology of reciprocity in international relations." *Negotiation Journal*. 4:3, 281-301.
- 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. (2015). "Dispute Resolution Institutions and Strategic Militarization." Unpublished manuscript.
- Myerson, Roger B. (1979). "Incentive Compatibility and the Bargaining Problem." *Econometrica* 47, 61-73.
- Prendergast, Canice and Lars Stole. (2001). "The Non-Monetary Nature of Gifts." *European Economic Review* 45:10, 1793-810.
- Rabin, Matthew. (1993). "Incorporating fairness into game theory and economics." *American Economic Review* 83:5, 1281-1301.

- Rauchhaus, Robert. (2006). "Asymmetric Information, Mediation, and Conflict Management." *World Politics* 58:2, 207-401.
- Slantchev, Branislav. (2011). "Military Threats: The Costs of Coercion and the Price of Peace." Cambridge: Cambridge University Press.
- Smith, Alastair, and Alan Stam. (2003). "Mediation and Peacekeeping in a Random Walk Model of Civil and Interstate War." *International Studies Review* 5:4, 115-35.
- Spence, Michael. (1973). "Job Market Signaling." *The Quarterly Journal of Economics*, 87:3, 355-74.
- Svensson, Isak. (2008). "Do Mediators Go Where they Are Needed Most? Mediation Selection in Civil Wars," Unpublished manuscript.
- Terris, Lesley G. and Zeev Maoz. (2005). "Rational Mediation: A Theory and a Test." *Journal of Peace Research* 44:5, 563-583.
- Van Damme, E. (1989). "Stable equilibria and forward induction. *Journal of Economic Theory* 48:2, 476-496.
- Van de Ven, Jeroen. (2000). "The economics of the gift." Unpublished manuscript.
- Walter, Barbara F. (2002). *Committing to Peace: The Successful Settlement of Civil Wars*. Princeton: Princeton University Press.
- Waltz, Kenneth (2001). *Man, the State and War: A Theoretical Analysis*. New York : Columbia University Press.
- Watson, Joel. (1999). "Starting Small and Renegotiation," *Journal of Economic Theory* 85:1, 52-90.