International Conventions and Non-State Actors: Selection, Signaling and Reputation Effects *

Kristian Skedre Gleditsch[†]
Department of Government, University of Essex and
Center for the Study of Civil War, PRIO

Simon Hug[‡]
Département de science politique, Université de Genève and
Center for the Study of Civil War, PRIO

Livia Isabella Schubiger[§]
Center for Comparative and International Studies (CIS),
Institute of Political Science, University of Zurich

and Julian Wucherpfennig[¶]
Center for Comparative and International Studies (CIS), ETH Zurich

First version: August 2010, this version: January 9, 2011

*An earlier version of this paper was presented at the CSCW workshop "Political Institutions and Conflict Resolution" (PRIO, Oslo November 4-5, 2010). We gratefully acknowledge helpful comments by the participants at this event. Thanks are also due to Marco Sassoli, chairman of the board of Geneva Call, as well as Pascal Bongard and Chris Rush of Geneva Call, who answered our numerous questions. Livia Isabella Schubiger acknowledges financial support by the Zurich University Alumni Association (fund for the support of aspiring young academics FAN) and the Swiss National Science Foundation (fellowship PBZHP1-133450). Simon Hug and Julian Wucherpfennig acknowledge partial financial support by the Swiss National Science Foundation (Grant No. 105511- 116795).

[†] Department of Government, University of Essex, Wivenhoe Park, Colchester CO4 3SQ, United Kingdom, email: ksg@essex.ac.uk.

[‡] Département de science politique, Faculté des sciences économiques et sociales; Université de Genève; 40 Bd du Pont d'Arve; 1211 Genève 4; Switzerland; phone ++41 22 379 83 78; email: simon.hug@unige.ch

[§] Center for Comparative and International Studies (CIS), Institute of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zurich, Switzerland; email: schubiger@pw.uzh.ch

[¶] Center for Comparative and International Studies (CIS), ETH Zurich, Haldeneggsteig 4,

Abstract

In many violent conflicts non-state actors (NSAs) play a considerable role, but contrary to state actors they cannot (be forced to) sign international conventions tying their hands. The non-governmental organization Geneva Call has stepped into this void and proposes NSAs to sign (and permit monitoring of) conventions banning, for instance, the use of landmines. Based on a game-theoretic model we assess what the motivations for NSAs (and states) are to sign such conventions and how they affect conflict behavior on the ground. We find that selection issues are of crucial importance linked to the incentive to signal resolve (both by states and NSAs). Empirical analyses on conflict behavior in countries where Geneva Call has been active support these theoretical insights.

8092 Zurich, Switzerland; email: wucherpfennig@icr.gess.ethz.ch

1 Introduction

On August 7, 2010 the newly elected president of Colombia Juan Manuel Santos in his inaugural address implored the *Fuerzas Armadas Revolucionarias de Colombia* (FARC) to cease using landmines in their guerrilla fight.[1] Some years before Geneva Call (2006), an international non-governmental organization (NGO) encouraging non-state actors to sign conventions to pledge refraining from using landmines,[2] attempted without success to win the FARC over to its cause, despite the fact that the Columbian government had signed the landmine treaty in 2000.[3] Geneva Call had more success in Sudan. In October 2001 the Sudan People's Liberation Movement/Army (SPLM/A) signed the proposed convention (Geneva Call, 2007), and only two years later the Sudanese government followed suit and signed the treaty.[4] Overall 153 countries have by now signed the landmine treaty, and 34 non-state actors (NSAs) from 7 countries have done the same for Geneva Call's (2007) convention. Few NSAs have signed the convention after their government signed the treaty, many more signed before the government pledged its support.

This raises two questions of importance in the current debate on human rights in international relations in general and civil wars more specifically. First, why would a non-state actor sign a constraining convention? And second, what effects do such conventions have? Both of these questions are intimately related to the current debate on the screening and constraining effects of international agreements (see for instance Simmons, 1998; von Stein, 2005; Simmons and Hopkins, 2005).

In what follows we first discuss the literature on human rights as it relates to our research question. We also discuss the context of Geneva call's intervention. In section three we propose a game-theoretic model focusing on the interaction between governments and NSAs when it comes to signing and complying with conventions related to human rights. Section four presents empirical tests of the implications derived from the theoretical model, while section five concludes.

[1] "Santos assumes Colombia's presidency amid conciliation with Venezuela, Ecuador" *LA Times* August 10, 2010 and "Santos Präsident Kolumbiens" *NZZ* August 9, 2010.

[2] Geneva Call also wishes to cover the areas of child soldiers and abuse of women.

[3] See `http://www.icbl.org/index.php/icbl/Universal/MBT/States-Parties`.

[4] See `http://www.icbl.org/index.php/icbl/Universal/MBT/States-Parties`.

2 Human rights and non-state actors

The field of human rights has seen over the last decades a burgeoning interest among academics (e.g., Finnemore and Sikkink, 1998; Risse, Ropp and Sikkink, 1999; Hathaway, 2002; Hafner-Burton, 2008; Vreeland, 2008; Simmons, 2009; Carey, Gibney and Poe, 2010; Hollyer and Rosendorff, 2010). Below we first succinctly discuss the recent literature on human rights as it is relevant for our research question, before offering a short overview over Geneva Call's actions.

2.1 Human rights

Authors like Finnemore and Sikkink (1998) (see also Risse, Ropp and Sikkink, 1999) saw in the raising importance of human rights norms clear evidence for phenomena expected by sociological institutionalist arguments (e.g., March and Olsen, 1984; Dimaggio and Powell, 1991). More recent work focusing on the tangible effects of human rights conventions highlights the critical issue of enforcement (e.g., Hathaway, 2002; Hafner-Burton, 2008; Vreeland, 2008; Simmons, 2009; Hollyer and Rosendorff, 2010). As several studies noted that authoritarian regimes happily signed on to human rights conventions without enforcing them (e.g., Hathaway, 2002; Hafner-Burton, 2008; Vreeland, 2008; Simmons, 2009; Hollyer and Rosendorff, 2010), the question arose whether the norm diffusion highlighted in earlier studies lacked any tangible consequences.

These interrogations harked back to a more general debate on the effects of international treaties (see for instance Simmons, 1998; von Stein, 2005; Simmons and Hopkins, 2005). In this literature the argument is made that assessing the constraining effects of international treaties (including treaties related to human rights) is hampered by the fact that signing a treaty is often also influenced by the expected compliance and compliance costs. Consequently, seeing signatories of particular treaties behave differently might simply be due to the fact that a particular set of countries chooses to sign the treaty (see for instance von Stein, 2005).[5] In the context of human rights the work by Hafner-Burton and Tsutsui

[5]Simmons and Hopkins (2005) contend in this debate that even when taking this selection process into account, there are constraining effects of treaties to be found. See for similar debates studies of whethr the World Trade Organization leads to trade liberalization or not (see Rose, 2002). In our context the Sudanese SPLA had already started to refrain using landmines when it signed Geneva Call's convention (personal communication by Pascal Bongard, program officer Geneva Call, January 5, 2011).

(2005, 2007) (see also Hafner-Burton, 2008), Vreeland (2008), Simmons (2009), and Hollyer and Rosendorff (2010) deals with these issues more specifically.

When it comes to NSAs, however, their human rights obligations are much less clear (e.g., Clapham, 2006), as they are by definition not signatories to human rights conventions. Only recently have scholars started to become interested in the conditions under which NSAs obey human rights norms (see for instance Jo and Thomson, 2008). The most extensive effort in this area is certainly Geneva call's initiative to propose human rights conventions to NSAs.[6] The first convention it offered NSAs to sign concerns the banning of land mines and is a parallel to the Ottowa convention (see for instance Goose, 1998). As with the early work on human rights, many studies emphasized in a sociological institutionalist perspective the importance of NGOs and civil society to bring about this convention (e.g., Price, 1998; Short, 1999; Anderson, 2000; Rutherford, 2000*b*; Rutherford, 2000*a*; Wexler, 2003; Lins de Albuquerque, 2007). More recently, however, scholars questioned the importance of civil society in this context, as most of the signatories of the Ottawa convention did not stock land mines and the enforcement mechanisms remained particularly weak (see for instance Drezner, 2005).

This makes it all the more interesting to understand first of all why NSAs would sign a convention imitating the Ottowa convention, and second how this affects their human rights record. Studies dealing with theses issues, especially those focusing on the interaction between governments and NSAs, are rather rare. Jo and Thomson (2008), for instance, propose a theoretical model assessing how compliance with human rights norms related to reputation and international organizations.[7]

2.2 Geneva Call

Geneva Call is an NGO that aims at engaging armed non-state actors to respect international humanitarian law and human rights law. It was founded in 1998, the year after the Ottawa convention was adopted, in response to the concern that this convention was only binding on states, allowing armed non-state actors to continue using these weapons. Geneva Call effectively began in 2000 to engage

[6] See `http://www.genevacall.org/` for more details.

[7] Related is Beber and Blattman's (2010) work dealing with child-soldiers, an area into which Geneva call is also in the process of venturing (see `http://www.genevacall.org/Themes/Children/children.htm`).

non-state actors on the subject of landmines. To this end, Geneva Call offers the "Deed of Commitment for Adherence to a Total Ban on Anti-Personnel Mines and for Cooperation in Mine Action." The convention engages non-state actors to ban the production, use, and transfer of landmines, as well as to participate in mine clearance and mine risk education. Importantly, the convention entails verification missions by Geneva Call. Geneva call is currently engaged in 6 areas, namely Africa, Asia, the Caucasus, Europe, the Middle East and Latin America.[8]

Africa:[9]

> Since 2000, Geneva Call has made significant headway in Africa, where 20 NSAs have signed the Deed of Commitment banning anti-personnel (AP) mines. As a result, AP mine use has decreased, stockpiles are being destroyed and mine action activities have expanded in areas under the control of signatory NSAs. Geneva Call currently works in Niger, Senegal, Somalia and Western Sahara, where it supports and monitors implementation of the Deed of Commitment banning AP mines and continues to engage additional NSAs in an AP mine ban. The successful peace processes in Burundi and Sudan enabled Geneva Call to end its programmes in those countries. Regions of Engagement: Burundi, Niger, Senegal, Somalia, Sudan, Western Sahara.

Asia:[10]

> Geneva Call has been active in Asia since its launch in 2000. The organization currently works in Burma/Myanmar, Northeast India and the Philippines. The successful peace processes in Nepal and Aceh, Indonesia, and the military defeat of the Liberation Tigers of Tamil Eelam (LTTE) in Sri Lanka led to Geneva Call ending its programmes in those countries. Regions of Engagement: Burma-Myanmar, India, Philippines.

Caucasus:[11]

[8] See `http://www.genevacall.org/home.htm` (accessed September 7, 2010)

[9] Source: `http://www.genevacall.org/Africa/africa.htm` (accessed September 7, 2010)

[10] Source: `http://www.genevacall.org/Asia/asia.htm` (accessed September 7, 2010)

[11] Source: `http://www.genevacall.org/Caucasus/caucasus.htm` (accessed Septemer 7, 2010)

> Geneva Call has been working in the South Caucasus since 2006 to engage internationally non - or partially recognized authorities in the region in a ban on AP mines. . . . Regions of Engagement: Armenian and Azerbaijan; Georgia.

Europe:[12]

> Geneva Call has been working in Turkey since 2001 resulting in the Kurdistan People's Congress (KONGRA-GEL)/People's Defence Forces (HPG), also known as the Kurdistan Workers' Party (PKK), signing the Deed of Commitment banning AP mines in 2006.. . . Regions of Engagement: Turkey.

Latin America:[13]

> Since 2003, Geneva Call has been working in Colombia to promote adherence to the AP mine ban by NSAs and to facilitate the implementation of emergency humanitarian mine action in favour of communities affected by AP mines. Regions of Engagement: Colombia.

Middle East:[14]

> Since 2000, Geneva Call has engaged in an ongoing dialogue in an AP mine ban with armed non-State actors (NSAs) operating in Iran, Iraq, Lebanon, and Yemen. As a result, four NSAs in Iran and two in Iraq have signed the Deed of Commitment banning AP mines to date. . . . Regions of Engagement: Iran, Iraq, Lebanon, Yemen

Table 1 contains information on the ratification sequence of both governments and non-state actors of the convention on landmines.

[12]Source: `http://www.genevacall.org/Europe/europe.htm` (accessed September 7, 2010)
[13]Source: `http://www.genevacall.org/Latin-America/latin-america.htm` (accessed September 7, 2010)
[14]Source: `http://www.genevacall.org/Middle-East/middle-east.htm` (accessed September 7, 2010)

Table 1: Ratification of Landmine Ban Convention

countries (number of NSAs that signed in parentheses)
country signed first, NSA after
Burundi (1)
Phillipines (3)
Turkey (1)
NSA signed first, country after
Iraq (2)
Sudan (1)
NSA signed first, country not yet
Burma (6)
India (3)
Iran (6)
Morocco (1)
Somalia (17)
country signed, no NSA signed
148 countries
neither country nor NSA signed
34 countries

Sources:
`http://www.icbl.org/index.php/icbl/Universal/MBT/States-Parties`
Geneva Call (2007)

3 A model

To better understand how Geneva Call's activity (understood as an equivalent to the Ottawa convention) shapes the interaction between the government and non-state actors, we rely on a game-theoretic model. The model is based on the interaction between two actors, namely a government G and a rebel organization R. The sequence of play is as follows:

1. government can sign or not sign treaty;
2. rebel organization can sign or not sign treaty;
3. if only rebel organization has signed, government gets another chance to sign or not sign the treaty.

The payoffs are assumed to be composed of the following elements:

- cw_i, with $i \in \{G, R\}$, corresponds to the costs of the civil war,[15]
- w_i, with $i \in \{G, R\}$, corresponds to the increased costs of warfare if treaty is adhered to (by assumed symmetry, these increased costs generate benefits for the adversary),
- r_G are reputation benefits (if G signs first) or costs (if R signs and G does not.[16]

The extensive form of this model, including payoffs, is depicted in figure 1. This simple structure already generates some insights under what conditions R and G sign the agreement. However, this version of the model does not tell us anything about compliance, and thus whether these agreements have any tangible effects. We therefore extend the model by assuming that w_i depends on compliance (which may be monitored but not directly observed).[17] In particular, we assume that G and R can be of two different types: following Jo and Thomson (2008), they can be either "nice" or "mean." Thus, we assume that complying with the agreement results in the payoffs depicted in figure 1 but that non-compliance by actor i withdraws from both actors' respective payoffs the w_i term while the "offending" actor i pays a cost of c_i related to the lack of compliance detected (possibly stochastically) by monitoring. Thus,

- c_i, with $i \in \{G, R\}$, corresponds to the increased costs related to non-compliance.

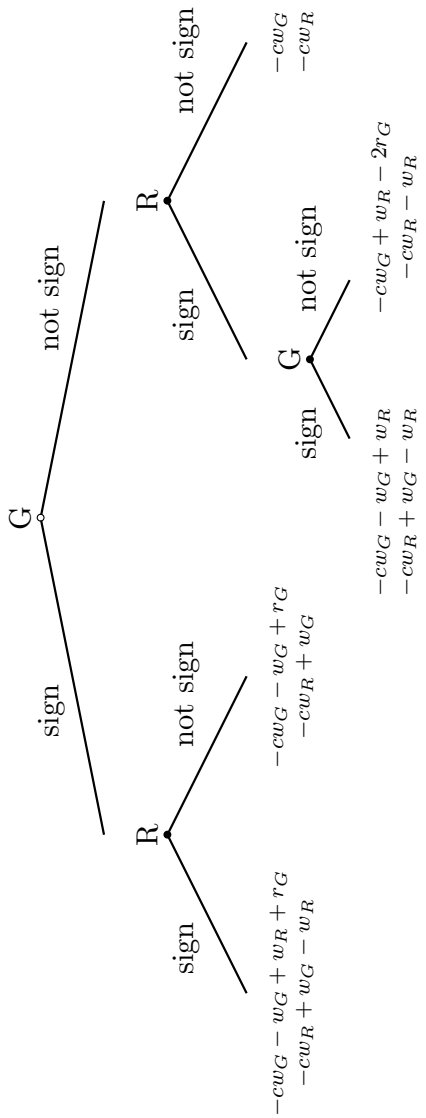
This results in a signaling game with two-sided incomplete information in which the effects of the agreements will become endogenous. The decisions to comply by G and R are then reached simultaneously leading to the game form depicted in figure 2, which for simplicity's sake omits nature's move to select the two types of G and R.

[15] As we will focus on states and NSAs engaged in civil wars, this term will be constant and could be dropped. We nevertheless keep it in what follows to allow for extensions beyond civil war cases.

[16] More precisely, the costs of not signing the treaty after R's signing of the convention is assumed to be twice as large as the benefits of signing first. It is easy to see that if costs and benefits were of the same magnitude, G will always sign at the first decision node if it were also to sign at its second node. We assume that rebels do not face reputation costs or benefits.

[17] Both the Ottawa convention on baning landmines and Geneva call's convention include monitoring provisions (e.g, Geneva Call, 2007).

Figure 1: Signing treaty without compliance decision (complete and perfect information)



Using this simple modification leads to four combinations of possible compliance decisions:

1. both G and R comply

$$\begin{aligned} EU_G &= -cw_G - w_G + w_R + r_G \\ EU_R &= -cw_R - w_R + w_G \end{aligned} \tag{1}$$

2. only G complies

$$\begin{aligned} EU_G &= -cw_G - w_G + r_G \\ EU_R &= -cw_R + w_G - c_R \end{aligned} \tag{2}$$

3. only R complies

$$\begin{aligned} EU_G &= -cw_G + w_R + r_G - c_G \\ EU_R &= -cw_R - w_R \end{aligned} \tag{3}$$

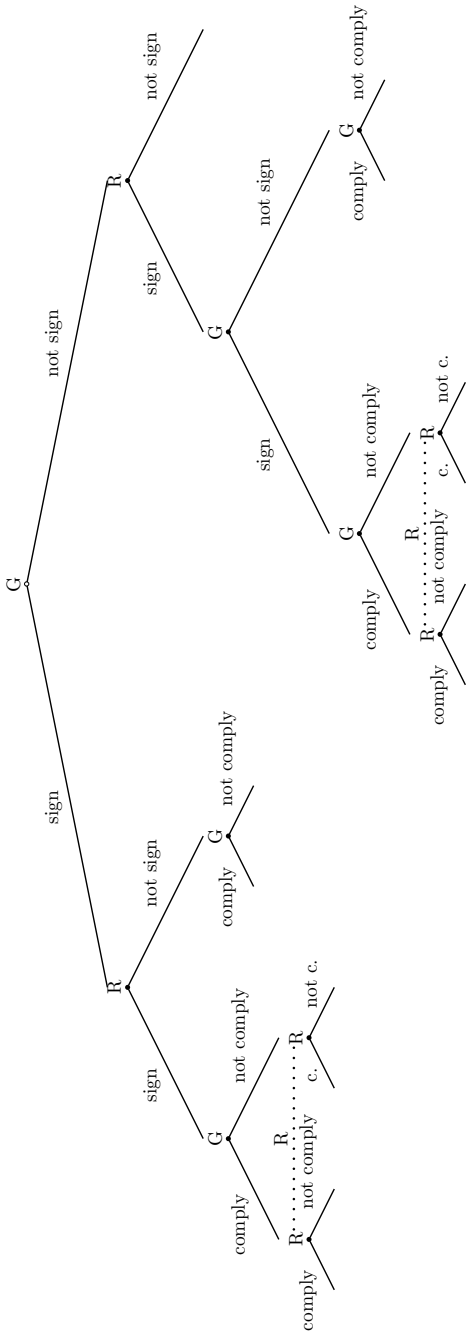
4. neither G nor R complies

$$\begin{aligned} EU_G &= -cw_G + r_G - c_G \\ EU_R &= -cw_R - c_R \end{aligned} \tag{4}$$

From this setup it easily follows that compliance for both actors i depends on the condition that $-w_i > -c_i$. Consequently, we use this condition to define the "mean" and "nice" types of actors. That is for a "nice" G $-w_G > -c_G$ while for a "mean" G $-c_G > -w_G$ and similarly for a "nice" R $-w_R > -c_R$ holds while $-c_R > -w_R$ holds for a "mean" R.[18]

[18]More precisely, we assume that all payoff elements are common knowledge except the c_is which are private information to both is, respectively.

Figure 2: Signing treaty with compliance decision (complete but imperfect information, move by Nature determining types omitted)



Consequently, if both G and R are uncertain about the type of their adversary, compliance will depend on the updated beliefs of these two actors. We denote the prior beliefs as p ($prob(c_G > w_G)$) and q ($prob(c_R > w_R)$). We will use this more general formulation when analyzing the complete and imperfect information version of this model, but replace it with a simplified version for the incomplete information version, where c_i may take two values, namely $2 \times w_i$ for a "nice" type and $\frac{w_i}{2}$ for a "mean" type.[19]

Proposition 1 (Complete and imperfect information) *In any subgame-perfect equilibrium, either G fails to sign at its first decision node but signs after R's signing (if $p = 1$, $q = 1$ and $2 \times r_G > w_G$) or G signs at its first decision node, while R refrains from doing so (in all other cases).*

Proposition 1 suggests that in the complete and imperfect information version of the game R may induce G to sign (or vice versa G by not signing first forces R to sign).

Proposition 2 (Incomplete and imperfect information) *Each of the perfect Bayesian equilibria produce one the following outcomes:*

- "mean" G signs at first decision node and nice G does not, leading both types of R not to sign.

- "mean" G signs at first decision node and "nice" G does not, leading the mean type of R to sign, followed by a signing by R as well.

- both types of G refrain from signing at each of their deicision node, leading both types of R not to sign.

- both types of G sign at their first decision node, leading both types of R not to sign.

- "nice' G does not sign first, while the "mean" R does not sign probabilistically, leading the nice R to sign probabilistically, while the "mean" R always signs, which are followed by both types of G signing.

For the empirical purposes of this paper the proof of proposition 2 offers especially the following interesting implications:

- If w_G is sufficiently high compared to r_G neither G nor R will ever sign

[19] If the more general formulation were to be used, some equilibria would depend on the exact distribution of the two c_is.

- If the prior belief q is low, then both types of G will sign immediately.
- For moderate values of w_G the "nice" G may not sign at first, inducing R to sign on its turn.

Getting back to the more substantive elements one critical element leading NSAs to sign conventions is whether w_R is high, i.e., if the profits for the government and the costs for the NSA of signing the agreement are large. A rather important factor in this is certainly whether or not NSAs occupy territory. Those that occupy territory are likely to face considerable costs by signing a convention and create benefits for the government.

Vice-versa one of the driving elements for governments to sign the treaty is whether the reputation costs of not signing the treaty outweigh the costs generated by the treaty.

4 Implications and empirical tests

4.1 Scope of Data

We begin by describing our data. Since we are interested in evaluating the consequences of Geneva Call's engagement, our analyses are temporally and spatially restricted to cases in which Geneva Call played an active role. Moreover, given the setup of our theoretical model, we require data that allow us to model the (strategic) interaction between non-state actors and their governments. Thus, the format of our data is dyad-year.

The next step is to define the sample. Within the regions (and the respective time periods) of Geneva Call's engagement, the dataset covers all dyads for which the non-state actor has been involved in intra-state armed conflict (as defined by UCDP,footnote"An armed conflict is a contested incompatibility that concerns government and/or territory where the use of armed force between two parties, of which at least one is the government of a state, results in at least 25 battle-related deaths in one calendar year." `http://www.pcr.uu.se/database/definitions_all.htm` (accessed September 14, 2010).) at least once since 1989. More precisely, dyads are included if the non-state actor has been actively involved in armed hostilities with the government, i.e., in intra-state conflict as defined by UCDP,[20]

[20]UCDP Dyadic Dataset v. 1-2010 (Harbom, 2008; Harbom, 2010).

during at least one year during the period from 1989 through 2009.[21] Armed organizations do not enter the dataset prior to their active involvement in an intra-state armed conflict. Once non-state actors have qualified for inclusion, they enter the dataset on a yearly basis during Geneva Call's period of engagement in the respective region, regardless of whether they were actively engaged in armed conflict during a given year. However, we only include non-state actors as long as they qualify as politically active organizations that maintain their own armed wing (our coding effort). To ensure robustness, we run our estimations on both a strict and a more lenient coding of activity (the latter includes dyad-years for which the pattern of activity is unclear).

To illustrate, the conflict between MFDC and the government of Senegal (dyadid 129) was coded active first in 1990 in the UCDP dyadic dataset (v 1-2010). Dyad 129 is therefore included in the dataset during all years since Geneva Call became active in the respective region (2000 onwards), although this dyad did not reach the 25 battle-related threshold every year since 2000. Naturally, we restrict this sample to regions and periods of Geneva Call's engagement. These are listed below.

Regions and time periods of Geneva Call's engagement:[22]]

- Africa (2000 onwards): (Burundi), Niger, Senegal, Somalia, (Sudan), Western Sahara/Morocco.
- Asia (2000 onwards): Burma-Myanmar, India, Philippines, (Nepal), (Sri Lanka), (Indonesia).
- Caucasus (2006 onwards): Azerbaijan [and Armenia][23], Georgia.
- Europe (2001 onwards): Turkey.

[21] Intra-state conflict dyads are composed of the government of a state and an armed opposition organization. UCDP defines armed opposition organizations as "[a]ny non-governmental group of people having announced a name for their group and using armed force to influence the outcome of the stated incompatibility" (Harbom, 2010). The criterion for inclusion of non-state actors into the UCDP dyadic dataset is at least 25 battle-related deaths during the given year in the dyad of the warring party `http://www.pcr.uu.se/database/definitions_all.htm` (accessed September 14, 2010).

[22] `http://www.genevacall.org/home.htm` (accessed September 11, 2010).

[23] We only include Azerbaijan in our dataset.

- Latin America (2003 onwards): Colombia.
- Middle East (2000 onwards): Iran, Iraq, Lebanon[24], Yemen.

Note that Geneva Call provides more accurate start and end dates of engagement for a subset of countries. For this version, the start year as indicated for Geneva Call's regions of engagement was taken. The countries where Geneva Call has already ended its programs are listed in parentheses.

4.2 Variables

Our main variable captures whether (or not) the NSA has signed Geneva Call's deed of commitment banning anti-personnel (AP) mines during a given year. For the government, the corresponding variable denotes ratification of the international mine-ban treaty.

We employ several additional variables in our tentative analyses: *OSV* denotes the extent to which rebels or governments were responsible for one-sided violence according to UCDP.[25] The variables indicate the best estimate of the aggregated estimated fatalities for all incidents of one-sided violence for a given actor and year. Consistent with our sample definition, fatality estimates have been assigned to dyad-years if the perpetrator has been actively involved in a given dyadic conflict in any year since 1989.[26] Accordingly, the fatality estimates

[24]In Lebanon, Geneva Call is mainly in contact with organizations affiliated with Hezbollah. Therefore, the Israel-Hezbollah dyad is included in the dataset.

[25]To construct these variables, the UCDP One-sided Violence Dataset (Eck and Hultman, 2007) was used, an actor-year dataset on deadly attacks on civilians by governments and armed groups. It is based on media reports and provides information on the unilateral use of armed force by governments and formally organized groups against unarmed persons resulting in at least 25 deaths per calendar year (Kreutz, 2004; Kreutz, Eck, Wallensteen, Harbom, Hgbladh and Sollenberg, 2005). The most recent version, 1.3-2010 (as updated on August 30, 2010), covers the period 1989-2008. Information on one-sided violence during 2009 was adopted from the UCDP database (accessed October 30, 2010).

[26]Instances of one-sided violence were not assigned to a conflict-year if the perpetrating actor did not constitute one of the conflict parties according to UCDP/PRIO-criteria. Exceptions are militias that allegedly acted on behalf of - or supported by - the state (Janjaweed in Sudan, AUC in Colombia). The following perpetrators of one-sided violence have not been assigned to conflict dyads: Abu-Hafs al-Masri Brigades (Turkey), ACCU (Colombia), BLTF (India), DHD (Dima Halam Daogah) (only DHD-BW included in Dyadic Dataset) (India), Gov of USA (Iraq and Lebanon), Gov. of Lybia (Niger), HPC (India), Indian Mujahideen (India), Jamaat Jund al-Sahaba/Soldiers of the Prophet's Companions (Iraq), Jemaah Islamiya/Islamic Association (Indonesia), KRA (India), Lashkar-e-Taiba/Army of Taiba (India), Lord's Resistance Army (Sudan and Uganda), Medellin cartel (Colombia), MFDC-FN (Mouvement des forces dmocra-

attributable to one particular actor and year appear multiple times in the dataset where the respective actor has been involved in more than one dyad since 1989.[27] We employ OSV as a proxy for compliance.

territorial control is a dummy variable denoting whether the NSA exerts at least a moderate level of control over its main territory. As outlined above, we argue that this variable is related to w_R, the costs induced by treaty adherence. The logic is simple. Land mines are an effective way of securing territory from governmental intrusion, hence relinquishing their usage is likely to make the non-state actor more vulnerable since it removes an effective military strategy from her portfolio.

To capture size-related effects, such as military capacity, we also use an estimate of the estimated *troop size* of the non-state actor.

4.3 Ratifying Mine-Ban Treaties

For the time being we consider the decision by Geneva call to propose conventions in particular areas as exogenous. Our first set of analyses addresses some of the formal model's empirical implications for the ratification of mine-ban treaties by both governments and non-state actors.

We begin with some descriptive statistics given in table 2. The table lists the number of dyad-years in which governments or non-state actors by signatory status (dyad years following signature are dropped), and in parentheses the respective numbers when the 'other' actor has previously ratified. Substantially, the table suggests that the rate of signatory is roughly the same for governments and non-state actors, but once non-state actors have signed the Geneva Call

tiques de Casamance) (Senegal), Ranvir Sena (Army of Ranvir) (India), RRA (Rahanweyn Resistance Army) (Somalia), Salafia Jihadia (Morocco), SIMI/Students' Islamic Movement (India), SSDF (South Sudan Defence Force) (Sudan), SPM/SNA (Splinter of SPM, Somalia), UPDS (United People's Democratic Solidarity) (India), VHP/Vishwa Hindu Parishad (India), Ampatuan Militia (Philippines), Asa'ib Ahl al-Haqq (Iraq).

[27] To give an example, the government of Burundi was involved in several dyadic conflicts during the period 1989-2009. Therefore, the osv fatality estimates attributed to the government of Burundi in a given year have been assigned to all dyads that qualify for inclusion in our sample during this year (cf. sample definition, section 4.1.). Similarly, the actor 'Hutu rebels' encompasses more than one non-state actor involved in intra-state conflict (e.g., Palipehutu and Palipehutu-FNL) (cf. Harbom and Wallensteen, 2009); osv fatality estimates attributable to this actor are therefore assigned to several dyads. One exception to this general coding rule is Israel, which as a special case was coded only with respect to the conflict with Hezbollah (cf. section 4.1.).

Table 2: Signatories and "Follow-Suit" Signatories

	all signatories (follow suit signatories)	
	government	nsa
no	329 (40)	289 (100)
yes	17 (5)	13 (4)

convention, governments are much more likely to follow suit than in the reverse scenario (i.e., when governments sign first).

Next, we estimate a series of (corresponding) logit models with signatory by both governments and non-state actors as the dependent variable. Naturally, where applicable, we exclude subsequent years following the year of ratification. The results are given in Table 3.

A key result from the formal model was that governments are likely to face high reputation costs if they fail to sign the convention if the non-state actors has already done so. This suggest a positive effect of prior ratifications of Geneva Call's deed of commitment by non-state actors. Models 1 and 2 in Table 3 provide such an estimate for both the strict and the lenient sample definition: In the event that non-state actors signed a mine-ban treaty before the government, governments are significantly more likely to follow suit. In accordance with the implications of the incomplete information model the results also suggest that governments are more likely to ratify the mine-ban treaty in the presence of more "meaner" non-state actors (i.e., prior belief of facing a "nice" rebel group (q) as measured by their activity in one-sided violence.

Turning our attention to non-state actors (models 3 to 6), we find little evidence for a similar emulation effect. This non-result is in line with the formal model. In addition, we also find that rebel organizations which exert direct control over their core territory are less likely to accept the costs induced by a mine-ban, and thus refrain from signing the convention. Finally, there is some indication that more sizable non-state actors are more likely to sign the deed of commitment.

Thus, by and large these preliminary results yield considerable support when it comes to the selection effects determining when non-state actors and governments sign mine-ban treaties.

Table 3: Logit Estimates of Signatory Status

	(1)	(2)	(3)	(4)	(5)	(6)
	mb_gov	mb_gov	mb_nsa	mb_nsa	mb_nsa	mb_nsa
osv_nsa (log)	0.0430***	0.0225*	-0.0138	0.00336	-0.0112	-0.00634
	(0.016)	(0.012)	(0.030)	(0.032)	(0.050)	(0.054)
osv_gov (log)	0.0356	0.0356	0.00611	0.00761	0.0394	0.0472
	(0.041)	(0.036)	(0.026)	(0.026)	(0.050)	(0.056)
mbtreaty_nsa	1.330**	0.646**				
	(0.52)	(0.31)				
mbtreaty_gov			0.287	-0.0960	1.417	1.596*
			(0.63)	(0.65)	(0.97)	(0.96)
territorial control $\geq$ moderate					-2.804***	-2.561***
					(0.92)	(0.93)
rebel size (log)					1.788*	1.668*
					(0.95)	(0.97)
Constant	-2.362***	-2.291***	-3.636***	-3.630***	-18.72**	-17.75**
	(0.56)	(0.56)	(0.56)	(0.55)	(8.78)	(8.97)
Observations	346	469	402	599	192	200
ll	-62.30	-96.50	-57.19	-62.56	-15.87	-16.53

Robust standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

4.4 Evaluating the Effectiveness

Having established these patterns, we now turn to evaluating the effect of such conventions. For the time being we will neglect the fact that signatories are potentially endogenous to their effect.

One way of evaluating the effectiveness of international conventions baning the usage of land-mines is to assess civilian casualties. Since the main problem with land-mines is that they tend to produce victims among innocent civilians, we argue that all else being equal, renouncing land mines will lead to lower numbers in one-sided violence. More precisely, we view one-sided violence as a proxy variable for the extent to which states and non-state actors are willing to value the life of civilians, which in turn should be directly linked to the willingness to employ or ban the use of landmines.[28]

We therefore estimate negative binomial regressions with one-sided violence perpetrated by states or rebels as the dependent variable (table 4). To account for the panel structure of the dataset, we allow for an AR(1) serial correlation in the error term. The models are estimated using generalized estimating equations (GEE).

[28]While Geneva Call also monitors compliance with the convention, the information is not sufficiently detailed to allow for a direct test of the theoretical implications related to compliance,

Table 4: GEE Negative Binomial Estimates (AR(1) errors) of One-Sided Violence

	(1)	(2)	(3)	(4)	(5)	(6)
	osv_gov	osv_gov	osv_nsa	osv_nsa	osv_nsa	osv_nsa
mineban government	1.200	1.005	1.179	0.998	-0.395	-0.629
	(0.81)	(0.62)	(0.96)	(0.77)	(0.38)	(0.50)
mineban nsa	-0.220	-0.242	-1.621***	-1.513***	-0.372	-1.054
	(0.37)	(0.35)	(0.38)	(0.41)	(0.33)	(0.69)
territorial control $\geq$ moderate					0.222	-0.223
					(0.63)	(0.81)
mineban $\times$ territorial control						-1.149*
						(0.66)
rebel size (log)						0.598**
						(0.26)
Constant	3.749***	3.711***	2.834***	2.542***	3.222***	-1.337
	(0.19)	(0.19)	(0.35)	(0.36)	(0.54)	(1.78)
Observations	434	645	442	645	254	206
Number of dyadid	64	85	65	85	32	27
deviance	2701	4236	2805	4064	1490	1115

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

As before, we begin with the government (models 1 and 2). Surprisingly, we find that signatories are associated with higher numbers of one-sided violence, although this finding is only marginally significant under the lenient sample definition. Whether NSAs sign the convention does not seem to condition the governments' behavior. For non-state actors, however, we find a strong negative and statistically significant effect of convention ratification (models 3 and 4). However, this effect disappears once we control for territorial control (model 5), but comes to light again when including the multiplicative interaction term between territorial control and convention ratification; it is strongest in the (rare) event that non-state actors exert control over their core territory, yet ratify the convention.

5 Discussion and Conclusion

The present paper offered a first look and assessment of how NSAs decide on signing Geneva Call's convention on the ban of landmines and its effectivness. To do so we proposed a simple game-theoretical model on the interactions between governments and NSAs. The equilibrium analysis of this model allows for a rich set of implications, some of which we were able to test in a preliminary fashion in the present paper.

We found rather clear evidence that the decisions by governments and NSAs to sign a landmine ban convention are not independent. Especially governments appear to be affected quite considerably by the fact whether NSAs had signed before them. Similarly, for the governments decision to sign first its assessment of possible compliance by NSAs also appears to be an important factor. For NSAs not surprisingly territorial control is an important factor influencing the costs of implementation and thus also the signing decision.

Regarding the consequences of signing such conventions our results are less rich. We find that NSAs having signed Geneva Call's convention are less related to one-sided violence (our proxy for compliance), but given possible selection biases, this estimate has to be taken with a grain of salt.

Both the theoretical model and first empirical results suggest, however, that our research endeavour is a fruitful avenue. We plan on linking more closely our empirical analysis to the theoretical implications, first by deriving additional insights and second by addressing the strategic nature in the empirical testing (see for instance Signorino, 1999; Signorino, 2002; Signorino, 2003; Signorino and Yilmaz, 2003; Signorino and Tarar, 2006).

Appendix

In this appendix we first present a few observations helpful in proving the main propositions presented in the main text. We then prove the two propositions characterizing equilibrium behavior under complete and imperfect and incomplete and imperfect information.

1. Observation

   If G signs at its first decision node, R will never sign, since it obtains the benefit of compliance by G for free, or can not improve on its own its situation if G should sign but not comply.

   Proof: Simply comparing expected utilities with p' the possibly updated prior belief yields:

$$\begin{aligned} EU_R(sign) &= p' \times (-cw_R - q \times w_R + w_G - (1-q) \times c_R) \\ &= +(1-p') \times (-cw_R - qw_R + (1-q) \times c_R) \\ &= -cw_R - q \times w_R - (1-q)c_R + p' \times w_G \qquad (5) \\ EU_R(not\ sign) &= p' \times (-cw_R + w_G) + (1-p) \times (-cw_R) \\ &= -cw_R + p' \times w_G \qquad (6) \end{aligned}$$

   As w_R and c_R are both positive, independent of q R will never sign. QED.

2. Observation

   If R signs the agreement (when G has not in the first round), G's decision to sign after R is independent of its possibly updated belief of R's type q'.

   Proof: To see this assume first that G is "nice" (i.e., $p = 1$)

$$\begin{aligned} EU_G(sign) &= q' \times (-cw_G - w_G + w_R) + (1-q') \times (-cw_G - w_G) \\ &= -cw_G - w_G + q' \times w_R \qquad (7) \\ EU_G(not\ sign) &= q' \times (-cw_G - 2 \times r_G + w_R) + (1-q') \times (-cw_G - 2 \times r_G) \\ &= -cw_G - 2 \times r_G + q' \times w_R \qquad (8) \end{aligned}$$

Consequently, G signs if $-cw_G - w_G + q' \times w_R > -cw_G - 2 \times r_G + q' \times w_R$, hence only if $2 \times r_G > w_G$.

If G is "mean" (i.e., $p = 0$)

$$\begin{aligned} EU_G(sign) &= q' \times (-cw_G + w_R - c_G) + (1 - q') \times (-cw_G - c_G) \\ &= -cw_G - c_G + q' \times w_R \qquad (9) \\ EU_G(not\ sign) &= q' \times (-cw_G - 2 \times r_G + w_R) + (1 - q') \times (-cw_G - 2 \times r_G) \\ &= -cw_G - 2 \times r_G + q' \times w_R \qquad (10) \end{aligned}$$

In that case G will sign if $-c_G > -2 \times r_G$ or $2 \times r_G > c_G$.

In both cases, i.e. independent of p, the decision of G to sign or not is independent of q'. $QED.$

3. Observation

From observation 2 follows that if $2 \times r_G > w_G$ then independent of its type G will always sign at its second decision node.[29] If, however, $w_G > 2 \times r_G$ then the nice type does not sign, but the mean type signs as long as $2 \times r_G > c_G$, but will not comply or does not sign if $c_G > 2 \times r_G$. As in this case the payoff for R is identical, it can anticipate its payoff, namely if $2 \times r_G > w_G$ and $q = 1$

$$\begin{aligned} EU_R(sign) &= p' \times (-cw_R + w_G - w_R) + (1 - p') \times (-cw_R - w_R) \\ &= -cw_R - w_R + p' \times w_G \qquad (11) \\ EU_R(not\ sign) &= -cw_R \qquad (12) \end{aligned}$$

Consequently a nice R signs in that case if $p' > \frac{w_R}{w_G}$

For a mean R the payoffs are as follows:

$$\begin{aligned} EU_R(sign) &= p' \times (-cw_R + w_G - c_R) + (1 - p') \times (-cw_R - c_R) \\ &= -cw_R - c_R + p \times w_G \qquad (13) \\ EU_R(not\ sign) &= -cw_R \qquad (14) \end{aligned}$$

[29] This follows from the fact that $p = 0$ implies $w_G > c_G$.

Consequently a mean R signs in that case if $p' > \frac{c_R}{w_G}$

If on the other hand $w_G > 2 \times r_G$ then R knows that G either won't sign or won't comply. Consequently, its payoffs for a "nice" ($q = 0$) type are

$$EU_R(sign) = -cw_R - w_R \quad (15)$$
$$EU_R(not\ sign) = -cw_R \quad (16)$$

As $w_R > 0$ R will never sign.

For a "mean" type

$$EU_R(sign) = -cw_R - c_R \quad (17)$$
$$EU_R(not\ sign) = -cw_R \quad (18)$$

As $c_R > 0$ R will never sign.

Complete and imperfect information

Proof of Proposition 1

Based on the observations above the following subgame-perfect equilibrium can be established:

1. If $p = 1$, $q = 1$ and $2 \times r_G > w_G$ $G : \{not\ sign, sign, comply\}$, $R : \{not\ sign, sign, comply\}$[30]
2. If $p = 1, q = 1$ and $w_G > 2 \times r_G$ $G : \{sign, ., comply\}$, $R : \{not\ sign, not\ sign, .\}$
3. If $p = 0, q = 1$ and $2 \times r_G > c_G$ $G : \{sign, ., not\ comply\}$, $R : \{not\ sign, not\ sign, .\}$
4. If $p = 0, q = 1$ and $c_G > 2 \times r_G$ $G : \{not\ sign, ., .\}$, $R : \{not\ sign, not\ sign, .\}$
5. If $p = 1, q = 0$ and $r_G > w_G$ $G : \{sign, ., comply\}$, $R : \{not\ sign, not\ sign, .\}$
6. If $p = 1$, $q = 0$ and $w_G > r_G$ $G : \{not\ sign, ., .\}$, $R : \{not\ sign, not\ sign, .\}$

[30]For simplicity's sake we shorten the strategies for both actors by only stating their actions at their first two decisions nodes and indicating with the third element the action taken at their remaining decision nodes, as these do not vary.

7. If $p = 0$, $q = 0$ and $r_G > c_G$ $G : \{sign, ., not\ comply\}$, $R : \{not\ sign, not\ sign, .\}$

8. If $p = 0$, $q = 0$ and $c_G > r_G$ $G : \{not\ sign, ., .\}$, $R : \{not\ sign, not\ sign, .\}$

As these equilibria exhaust all possible conditions proposition 1 simply summarizes the insights from these equilibrium characterizations. $QED.$

Incomplete information

As mentioned in the main text we simplify the model for the incomplete information version by letting $c_i \in \{2 \times w_i, \frac{w_i}{2}\}$. c_i takes the higher value if i is a "nice" type, and the lower one when i is a "mean" type. This allows us as an extension of the discussion above already to establish the following observation

1. Observation

   If G does not sign at the first decision node a R does at its second decision node, then if $w_G > 4 \times r_G$ then neither type of G will sign at its second decision node, while if $4 \times r_G > w_G > 2 \times r_G$ then only the "mean" type of G will sign, while if $2 \times r_G > w_G$ both types will sign.

Using this observation we start by deriving the conditions under which completely pooling and separating equilibria may occur before moving to semi-pooling equilibria

Pooling equilibria

We start by looking at a candidate equilibrium where both types of G refrain from signing at the first decision node. We first assume that $4 \times r_G > w_G$ implying that no type of G would sign at its second decision. Consequently R must evaluate the following expected utilities:

$$EU_R(sign|q = 1) = -cw_R - w_R \tag{19}$$

$$EU_R(sign|q = 0) = -cw_R - \frac{w_R}{2} \tag{20}$$

$$EU_R(not\ sign|q = .) = -cw_R \tag{21}$$

Consequently, both types of R will never sign in this case. For this to be part of a pooling equilibrium the following has to be evaluated:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G && (22)\\
EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G && (23)\\
EU_G(not\ sign|p=.) &= -cw_G && (24)
\end{aligned}
$$

For both types not to sign at the first decision node is optimal provided $r_g < w_G$ and $2 \times r_g < w_G$ hold, which is the case given our assumption from above. This establishes a first pooling equilibrium with the following conditions to hold:

- $w_G > 4 \times r_G$

Assuming now that $4 \times r_G > w_G > 2 \times r_G$ we know that a "mean" G will sign after R's decision to sign, while a "nice" type will not. As not signing by G has the same consquences for R as not complying, the expected utility calculations both for R and G are as above, establishing a second pooling equilibrium under the following conditions:

- $4 \times r_G > w_G > 2 \times r_G$

Finally if $2 \times r_G > w_G$ both types of G will sign after R's decision to sign. Consequently R evaluates the following expected utilities where the updated belief p' is identical to the prior belief, giving the assumption of a pooling equilibrium:

$$
\begin{aligned}
EU_R(sign|q=1) &= p'(-cw_R - w_R + w_G) + (1-p')(-cw_R - w_R)\\
&= -cw_R - w_R + p' \times w_G && (25)\\
EU_R(sign|q=0) &= p'(-cw_R - c_R + w_G) + (1-p')(-cw_R - \frac{w_R}{2})\\
&= -cw_R - \frac{w_R}{2} + p' \times w_G && (26)\\
EU_R(not\ sign|q=.) &= -cw_R && (27)
\end{aligned}
$$

This implies that a "nice" R will sign if $p' > \frac{w_R}{w_G}$, while a "mean" R will do so if $p > \frac{w_R}{2 \times w_G}$. From this it follows that we need to evaluate a series of possible configurations.

First consider $w_R > 2 \times w_G$ implying that both Rs will refrain from signing. For G the following expected utilities are relevant

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G && (28)\\
EU_G(not\ sign|p=.) &= -cw_G && (29)\\
EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G && (30)
\end{aligned}
$$

For both types of G not to sign requires that $w_G > r_G$ and $w_G > 2 \times r_G$. As the latter is in contradiction with the initial assumption no pooling equilibrium exists.

Second, assume that $2 \times w_G > w_R > w_G$ and $p > \frac{w_R}{2 \times w_G}$. As in this case again both Rs refrain from signing the above expected utilities for G apply, establishing that no pooling equilibrium exists.

Third, assume that $2 \times w_G > w_R > w_G$ and $p > \frac{w_R}{2 \times w_G}$ which implies that $p < \frac{w_R}{w_G}$. Consequently a "nice" G will not sign while a "mean" one will. Consequently, for G the following expected utilities become relevant:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G && (31)\\
EU_G(not\ sign|p=.) &= -cw_G && (32)\\
EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G && (33)
\end{aligned}
$$

from which it follows that both types of G would not sign if $w_G > r_G$ and $w_G > 2 \times r_G$ hold. As the latter conditions is in contradiction with the assumption that $2 \times r_G > w_G$ no pooling equilibrium exists.

Fourth, if $w_G > w_R$ and $p < \frac{w_R}{2 \times w_G}$ then neither of the two types of R will sign. Hence we are in the same situation as above and no pooling equilibrium exists.

Fifth, if $w_G > w_R$ and $\frac{w_R}{2 \times w_G} < p < \frac{w_R}{w_G}$ then only the mean R signs which is equivalent to the third situation implying again the absence of a pooling equilibrium.

Finally, if $w_G > w_R$ and $\frac{w_R}{w_G} < p$ then both Rs will sign. Consequently for G the following is relevant:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G && (34)\\
EU_G(not\ sign|p=1) &= q(-cw_G - w_G + w_R) + (1-q)(-cw_G - w_G)\\
&= -cw_G - w_G + q \times (w_R) && (35)
\end{aligned}
$$

Thus a nice G will not sign if $-cw_G - w_G + q \times w_R > -cw_G - w_G + r_G$ which is equivalent to $q > \frac{r_G}{w_R}$. For the mean G the following expected utilities are relevant:

$$\begin{aligned} EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G & (36) \\ EU_G(not\ sign|p=0) &= q(-cw_G - \frac{w_G}{2} + w_R) + (1-q)(-cw_G - \frac{w_G}{2}) \\ &= -cw_G - \frac{w_G}{2} + q \times (w_R) & (37) \end{aligned}$$

Thus a mean G will not sign if $-cw_G - \frac{w_G}{2} + q \times w_R > -cw_G - \frac{w_G}{2} + r_G$ which is equivalent to $q > \frac{r_G}{w_R}$. Consequently, a pooling equilibrium exists if

- $w_G > w_R$
- $p > \frac{w_R}{w_G}$
- $q > \frac{r_G}{w_R}$
- $2 \times r_G > w_G$

As this exhausts all possible conditions for the first type of pooling equilibrium, we now consider a pooling equilibrium where both types of G sign at the first decision node. Given the derivations of the first set of pooling equilibria, this can only occur if $2 \times r_G > w_G$.

Assume first that $w_G > w_R$ and that the out-of equilibrium belief is $p' = 1$, which leads both types of R to sign. The relevant expected utilities for G are as follows:

$$\begin{aligned} EU_G(sign|p=1) &= -cw_G - w_G + r_G & (38) \\ EU_G(not\ sign|p=1) &= q(-cw_G - w_G + w_R) + (1-q)(-cw_G - w_G) \\ &= -cw_G - w_G + qw_R & (39) \end{aligned}$$

Consequently, a "nice" G will prefer signing if $r_G > qw_R$ or if $q < \frac{r_G}{w_R}$ For a mean G the following is relevant:

$$
\begin{aligned}
EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G && (40)\\
EU_G(not\ sign|p=1) &= q(-cw_G - \frac{w_G}{2} + w_R) + (1-q)(-cw_G - \frac{w_G}{2}) \\
&= -cw_G - \frac{w_G}{2} + qw_R && (41)
\end{aligned}
$$

which again requires $q < \frac{r_G}{w_R}$ for a "mean" G to sign at the first decision node, establishing thus a pooling equilibrium under the following conditions:

- $2 \times r_G > w_G$
- $q < \frac{r_G}{w_R}$
- $w_G > w_R$
- $p' = 1$

Let's next assume that $2 \times w_G > w_R > w_G$ and $p' = 1$, implying that only the "mean" R will sign, which implies the following:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G && (42)\\
EU_G(not\ sign|p=1) &= q(-cw_G) + (1-q)(-cw_G - w_G) \\
&= -cw_G - w_G + qw_G && (43)
\end{aligned}
$$

Consequently, a "nice" G will prefer signing if $r_G > qw_G$ or if $q < \frac{r_G}{w_G}$ For a mean G the following is relevant:

$$
\begin{aligned}
EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G && (44)\\
EU_G(not\ sign|p=1) &= q(-cw_G) + (1-q)(-cw_G - \frac{w_G}{2}) \\
&= -cw_G - \frac{w_G}{2} + q\frac{w_G}{2} && (45)
\end{aligned}
$$

which implies that a "mean" G will only sign if $r_G > q\frac{w_G}{2}$ or $q < \frac{2 \times r_G}{w_G}$. Hence a pooling equilibrium exists if

- $2 \times r_G > w_G$

- $2 \times w_G > w_R > w_G$
- $q < \frac{r_G}{w_G}$
- $p' = 1$

Next, let's assume that $w_R > 2 \times w_G$ which with $p' = 1$ will lead both types of R not to sign. Consquently the relevant expected utilities for G are the following:

$$\begin{aligned} EU_G(sign|p=1) &= -cw_G - w_G + r_G \quad (46) \\ EU_G(not\ sign|p=.) &= -cw_G \\ EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G \quad (47) \end{aligned}$$

Consequently, a both types of G will prefer signing if $r_G > w_G$ and $2 \times r_G > w_G$ hold, establishing a last pooling equilibrium:

- $r_G > w_G$
- $w_R > 2 \times w_G$
- $p' = 1$

Let's next assume that the out-of-equilibrium belief is $p' = 0$. In that case neither types of R will sign. Consequently, for G the following expected utilities are relevant

$$\begin{aligned} EU_G(sign|p=1) &= -cw_G - w_G + r_G \quad (48) \\ EU_G(not\ sign|p=.) &= -cw_G \quad (49) \\ EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G \quad (50) \end{aligned}$$

Consequently, a "nice" G will prefer signing if $r_G > w_G$ and a "mean" one will prefer the same if $r_G > \frac{w_G}{2}$ or $2 \times r_G > w_G$. This establishes again a pooling equilibrium under the following conditions

- $2 \times r_G > w_G$
- $r_G > w_G$

This equilibrium is based, however, on a counter-intuitive out-of-equilibrium belief, as it is mostly the "nice" G that could profit from not signing.

Separating equilibria

Given the complete and imperfect information equilibria, a first candidate for a separating equilibrium is that the nice G does not sign the treaty and the mean G signs it. Whether this can be a separating equilibrium depends, as above, on relationship between r_G and w_G.

Assuming $w_G > 4 \times r_G$ we know from above that neither type of G will sign if R signs. But then R will neither sign. Hence, for this condition to allow for a separating equilibrium to exist the following has to hold:

$$EU_G(sign|p=1) = -cw_G - w_G + r_G \quad (51)$$

$$EU_G(not\ sign|p=.) = -cw_G \quad (52)$$

$$EU_G(sign|p=0) = -cw_G - \frac{w_G}{2} + r_G \quad (53)$$

Consequently a "nice" G will not sign if $w_G > r_G$, while for the "mean"' G $r_G > \frac{w_G}{2}$ has to hold. But the latter is in contradiction with the assumption that $w_G > 4 \times r_G$ so that no separating equilibrium can exist.

Assuming next that $4 \times r_G > w_G > 2 \times r_G$ we know that a "mean" G will sign after R's signing, while the "nice" G will not. This induces R not to sign either. But then the same conditions as discussed above have to hold for a separating equilibrium to exist, which are again in contradiction with the assumption that $4 \times r_G > w_G > 2 \times r_G$. Hence no separating equilibrium exists.

Next assume that $2 \times r_G > w_G$ implying that both types of G will sign after R's decision to sign. As shown above in this situation a "nice" R will sign if $p' > \frac{w_R}{w_G}$ while a "mean" R will do the same if $p' > \frac{w_R}{w_G}$ holds. Consequently, a series of configurations have to be evaluated.

First, assume that $w_R > 2 \times w_G$ which implies that the threshold values for the updated beliefs of R are both higher than 1 implying that both Rs will refrain from signing. Consequently, the question becomes how this situations looks from G's perspective:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G && (54)\\
EU_G(not\ sign|p=.) &= -cw_G && (55)\\
EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G && (56)\\
& && (57)
\end{aligned}
$$

Consequently, a nice G will not sign if $w_G > r_G$ while a "mean" G will sign if $r_G > \frac{w_G}{2}$. This establishes a separating equilibrium under the following conditions:

- $2 \times r_G > w_G$
- $w_G > r_G$
- $w_R > 2 \times w_G$

Second, let's assume that $2 \times w_G > w_R > w_G$. As in the proposed separating equilibrium $p' = 1$ and this value is smaller than $\frac{w_R}{W_G}$ but larger than $\frac{w_R}{2W_G}$ the "mean" R will sign, while the "nice" R will refrain from doing so. Thus from G's perspective the following expected utilities are relevant:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G && (58)\\
EU_G(not\ sign|p=1) &= q \times (-cw_G) + (1-q) \times (-cw_G - w_G) \\
&= -cw_G - w_G + q \times w_G && (59)
\end{aligned}
$$

From this it follows that a "nice" G will not sign if $\frac{r_G}{w_G} < q$. This can only happen if $w_G > r_G$. For the "mean" G the following expected utilities apply:

$$
\begin{aligned}
EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G && (60)\\
EU_G(not\ sign|p=0) &= q \times (-cw_G) + (1-q) \times (-cw_G - \frac{w_G}{2}) \\
&= -cw_G - \frac{w_G}{2} + q \times \frac{w_G}{2} && (61)
\end{aligned}
$$

so that a "mean" G will sign if $r_G > q \times \frac{w_G}{2}$ or that $q < \frac{2 \times r_G}{w_G}$. As we assume that $2 \times r_G > w_G$ this latter condition will always hold, establishing thus a separating equilibrium under the following conditions:

- $2 \times r_G > w_G$
- $2 \times w_G > w_R > w_G$
- $w_G > r_G$
- $q > \frac{r_G}{w_G}$

Finally, let's assume that $w_G > w_R$ which implies that both thresholds for the updated belief p' are smaller than one leading R in the proposed separating equilibrium to sign under all circumstanes. Hence, from G's perspective the following expected utilities are of importance:

$$\begin{aligned} EU_G(sign|p=1) &= -cw_G - c_G + r_G \quad (62) \\ EU_G(not\ sign|p=1) &= q(-cw_G - w_G + w_R) + (1-q)(-cw_G - w_G) \\ &= -cw_G - w_G + q \times w_R \quad (63) \end{aligned}$$

Consequently, the "nice" G will not sign if $q > \frac{r_G}{w_R}$ which requires $w_R > r_G$. For the "mean" G the following expected utilities are of interest:

$$\begin{aligned} EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G \quad (64) \\ EU_G(not\ sign|p=0) &= q(-cw_G - \frac{w_G}{2} + w_R) + (1-q)(-cw_G - \frac{w_G}{2}) \\ &= -cw_G - \frac{w_G}{2} + q \times w_R \quad (65) \end{aligned}$$

which implies that signing requires $r_G > q \times \frac{w_G}{2}$. This, however, is only possible of $q < \frac{r_G}{w_R}$, which is in contradiction with the condition for the "nice" G to not sign. Consquently, no separating equilibrium can exist.

A completely separating equilibrium may also exist where the nice G signs the treaty and the mean G does not. From above we know that the relationship between w_G and r_G are relevant.

Assuming that $w_G > 4 \times r_G$ we know that no type of G will sign at its second decision node so that R will also refrain from signing. Thus the following expected utilities becomes relevant:

$$EU_G(sign|p=1) = -cw_G - w_G + r_G \quad (66)$$
$$EU_G(not\ sign|p=1) = -cw_G \quad (67)$$

Thus a "nice" G will sign if $r_G > w_G$ which is in contradiction with the assumption $w_G > 4 \times r_G$. Consquently no separating equilibrium of this type exists.

Second assuming that $4 \times r_G > w_G > 2 \times r_G$ we know that the "mean" type of G will sign at its second decision while the "nice" type will not. But then again both types of R will also refrain from signing, so that the same conditions should hold for a separating equilibrium, which are again in contradiction with $4 \times r_G > w_G > 2 \times r_G$. Consquently, no separating equilibrium can exist.

Finally, if $2 \times r_G > w_G$ we know that both types of G will sign after R's signing. Given the proposed separating equilibrium we know that the updated belief p' is 0, leading both types of R to refrain from signing. Consequently, the following expected utilities become relevant:

$$EU_G(sign|p=1) = -cw_G - w_G + r_G \quad (68)$$
$$EU_G(not\ sign|p=.) = -cw_G \quad (69)$$
$$EU_G(sign|p=0) = -cw_G - \frac{w_G}{2} + r_G \quad (70)$$

Consequently, the "nice" G will sign at its first decision node if $r_G > w_G$, while the "mean" type will not sign the treaty if $\frac{w_G}{2} > r_G$. But the latter condition is in contradiction with $2 \times r_G > w_G$ so that no separating equilibrium of this type exists.

Semi-pooling equilibria

From above it follows that semi-pooling equilibria can only exist under the condition of $2 \times r_G > w_G$

The first candidate equilibrium is based on the following (partial) strategy for the two types of G at their first decision node:

$$p(not\ sign|p=1) = 1$$
$$p(not\ sign|p=.) = s \quad (71)$$

From this it follows that the updated belief for R is the following: $\frac{p}{p+s(1-p)}$.

We first assume that G chooses s in such a way that $p' = \frac{w_R}{w_G}$ implying that the "nice" R is indifferent between signing and not signing, while the "mean" R will sign with certainty. Consequently, s can be determined as follows:

$$\begin{aligned} \frac{p}{p+s(1-p)} &= \frac{w_R}{w_G} \\ p &= \frac{w_R}{w_G}s(1-p) \\ p(\frac{w_G}{w_R}-1) &= s(1-p) \\ s &= \frac{p(w_G-w_R)}{w_R(1-p)} \end{aligned} \tag{72}$$

For s to be larger than 0 $w_G > w_R$ has to hold, while $p < \frac{w_R}{w_G}$ assures that $s < 1$. As a "nice" R is in this case indifferent between signing or not signing its (partial) strategy will be $p(sign|q=1) = t$. Hence from G's perspective the following expected utilities are relevant:

$$EU_G(sign|p=1) = -cw_G - w_G + r_G \tag{73}$$

$$\begin{aligned} EU_G(not\ sign|p=1) &= q(t(-cw_G - w_G + w_R) + (1-t)(-cw_G)) \\ &\quad +(1-q)(-cw_G - w_G) \\ &= -cw_G - w_G + qw_G + qtw_R - qtw_G \end{aligned} \tag{74}$$

Consquently, the "nice" G will not sign if $q > \frac{r_G}{w_G + t(w_R - w_G)}$. For the "mean" G the following has to hold:

$$EU_G(sign|p=0) = -cw_G - \frac{w_G}{2} + r_G \tag{75}$$

$$\begin{aligned} EU_G(not\ sign|p=0) &= q(t(-cw_G - \frac{w_G}{2} + w_R) + (1-t)(-cw_G)) \\ &\quad +(1-q)(-cw_G - \frac{w_G}{2}) \\ &= -cw_G - \frac{w_G}{2} + q\frac{w_G}{2} + qtw_R - qt\frac{w_G}{2} \end{aligned} \tag{76}$$

As the "mean" type has to be indifferent the following has to hold:

$$r_G = -cw_G - \frac{w_G}{2} + q\frac{w_G}{2} + qtw_R - qt\frac{w_G}{2} \tag{77}$$

which implies

$$t = \frac{r_G - q\frac{w_G}{2}}{2q(w_R - \frac{w_G}{2})} \tag{78}$$

t will be positive if $q < \frac{2 \times r_G}{w_G}$ which will always be the case given that we assime that $2 \times r_G > w_G$. And t will be smaller than 1 if $r_G < w_R$. It can also easily be checked that the t determined here satisfies the condition for t for the "nice" G to sign. Consequently a semi-pooling equilibrium exists under the following conditions:

- $2 \times r_G > w_G$
- $p < \frac{w_R}{w_G}$
- $q > \frac{r_G}{w_R}$
- $w_R > r_G$

Assuming next that G will choose s in such a way that the "mean" R will be indifferent between signing and not signing, implying that the "nice" R will not sign, the following has to hold:

$$\begin{aligned} \frac{p}{p + s(1-p)} &= \frac{w_R}{2 \times w_G} \\ 2 \times w_G p &= w_R(p + s(1-p)) \\ s &= \frac{p(2 \times w_G - w_R)}{w_R(1-p)} \end{aligned} \tag{79}$$

s will be positive if $2 \times w_G > w_R$ and smaller than 1 if $p < \frac{w_R}{2 \times w_G}$. As a "mean" R is in this case indifferent between signing or not signing its (partial) strategy will be $p(sign|q = 0) = t$. Hence from G's perspective the following expected utilities are relevant:

$$EU_G(sign|p = 1) = -cw_G - w_G + r_G \tag{80}$$

$$\begin{aligned} EU_G(not\ sign|p = 1) &= q(-cw_G) + (1-q)(t(-cw_G - w_G) + (1-t)(-cw_G)) \\ &= -cw_G - (1-q)tw_G \end{aligned} \tag{81}$$

Consquently, the "nice" G will not sign if $-(1-q)tw_G > -w_G + r_G$ which implies that $t < \frac{w_G - r_G}{w_G(1-q)}$. For the "mean" G the following has to hold:

$$\begin{aligned} EU_G(sign|p=0) &= -cw_G - \frac{w_G}{2} + r_G \quad (82) \\ EU_G(not\ sign|p=0) &= q(-cw_G) + (1-q)(t(-cw_G - \frac{w_G}{2}) + (1-t)(-cw_G)) \\ &= -cw_G - \frac{w_G}{2}t(1-q) \quad (83) \end{aligned}$$

As the "mean" type has to be indifferent the following has to hold:

$$\begin{aligned} r_G - \frac{w_G}{2} &= -\frac{w_G}{2}t(1-q) \\ t &= \frac{\frac{w_G}{2} - r_G}{(1-q)\frac{w_G}{2}} \quad (84) \end{aligned}$$

For t to be positive $w_G > 2 \times r_G$ has to hold which is in contradiction with the assumption that $2 \times r_G > w_G$. Hence, no semi-pooling equilibrium of this type can exist.

References

Anderson, K. 2000. "The Ottawa Convention banning landmines, the role of international non-governmental organizations and the idea of international civil society." *European Journal of International Law* 11(1):91–120.

Beber, Bernd and Christopher Blattman. 2010. "The Industrial Organization of Rebellion: The Logic of Forced Labor and Child Soldiering." NYU and Yale University.

Carey, Sabine C., Mark Gibney and Steven C. Poe. 2010. *The Politics of Human Rights*. New York: Cambridge University Press.

Clapham, Andrew. 2006. "Human rights obligations of non-state actors in conflict situations." *International Review of the Red Cross* 88:491–523.

Dimaggio, Paul J. and Walter W. Powell. 1991. *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.

Drezner, Daniel. 2005. "Gauging the Power of Global Civil Society: Intellectual property and public health." Paper presented at the annual meeting of the American Political Science Association, Marriott Wardman Park, Omni Shoreham, Washington Hilton, Washington, DC, Sep 01, 2005.

Eck, Kristine and Lisa Hultman. 2007. "Violence Against Civilians in War." *Journal of Peace Research* 44(2):233–246.

Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 54(2):891.

Geneva Call. 2006. "The Impact of Armed Non-State Actors on the Mine Ban Treaty: Research and Analysis." Working paper presented at the 7th Meeting of States Parties to the Mine Ban Treaty in Geneva, Switzerland (1822 September 2006).

Geneva Call. 2007. "Engaging armed non-state actors in a landmine ban. The Geneva call progess report (2000-2007)." Geneva.

Goose, Stephen D. 1998. "The Ottawa Process and the 1997 Mine Ban Treaty." *Yearbook of International Humanitarian Law* 1:269–291.

Hafner-Burton, Emilie M. 2008. "Sticks and Stones: Naming and Shaming and the Human Rights Enforcement Problem." *International Organization* 62:689–716.

Hafner-Burton, Emilie Marie and Kiyoteru Tsutsui. 2005. "Human Rights in a Globalizing World: The Paradox of Empty Promises." *American Journal of Sociology* 110(5):1373–411.

Hafner-Burton, Emilie Marie and Kiyoteru Tsutsui. 2007. "Justice Lost! The Failure of International Human Rights Law to Matter Where Needed Most." *Journal of Peace Research* 44(4):407–425.

Harbom, Lotta. 2008. *UCDP Dyadic Dataset Codebook Version 1.0.* Uppsala Conflict Data Program, Uppsala University, Department of Peace and Conflict Research.

Harbom, Lotta. 2010. "UCDP Dyadic Dataset Codebook Version 1-2010." Uppsala Conflict Data Program (UCDP), Department of Peace and Conflict Research, Uppsala University.

Harbom, Lotta and Peter Wallensteen. 2009. "Armed Conflicts 19892008." *Journal of Peace Research* 46(4):577–587.

Hathaway, Oona A. 2002. "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111:1935–2042.

Hollyer, James R. and Peter Rosendorff. 2010. "Domestic Politics and the Accession of Authoritarian Regimes to Human Rights Treaties." Paper prepared for presentation at the 3rd Annual Conference on The Political Economy of the International Organizations January 28-30, 2010, Washington D.C.

Jo, Hyeran and Catarina P. Thomson. 2008. "Why Rebels Comply with Humanitarian Rules When Governments Don't: The Effects of Reputation and International Organizations in Armed Conflicts 1991-2006." Paper to be presented in the American Political Science Association (APSA) Annual Meeting in Boston MA, August 27-31, 2008.

Kreutz, Joakim. 2004. Collective Violence in 2004: Focus on Human Insecurity on the Horn of Africa. In *States in Armed Conflict 2004.*, ed. Lotta Harbom. Uppsala University: Department of Peace and Research Report 71 pp. 147–158.

Kreutz, Joakim, Kristine Eck, Peter Wallensteen, Lotta Harbom, Stina Hgbladh and Margareta Sollenberg. 2005. "UCDP One-sided Violence Codebook. Version 1.0 - September 28, 2005.".

Lins de Albuquerque, Adriana. 2007. "The Power of Norms?: Reassessing the Significance of International Support for the Mine Ban Treaty." Columbia University.

March, James G. and Johan P. Olsen. 1984. "The New Institutionalism: Organized Factors of Political Life." *American Political Science Review* 78:734–749.

Price, Richard. 1998. "Reversing the Gun Sights: Transnational Civil Society Targets Land Mines." *International Organization* 52(3):613–644.

Risse, Thomas, Stephen C. Ropp and Kathryn Sikkink, eds. 1999. *The Power of Human Rights. International Norms and Domestic Change.* Cambridge: Cambridge University Press.

Rose, Andrew K. 2002. "Do WTO Members Have a More Liberal Trade Policy?" NBER Working Paper Series Working Paper 9347.

Rutherford, Kenneth R. 2000*a*. "The Evolving Arms Control Agenda: Implications of the Role of NGOs in Banning Antipersonnel Landmines." *World Politics* 53:74–114.

Rutherford, Kenneth R. 2000*b*. "Internet activism: NGOs and the Mine Ban Treaty." *International Journal on Grey Literature* 1(3):99 – 106.

Short, N. 1999. "The Role of NGOs in the Ottawa Process to Ban Landmines." *International Negotiation* 4(3):483–502.

Signorino, Curtis S. 1999. "Strategic Interaction and the Statistical Analysis of International Conflict." *American Political Science Review* 93(2):279–297.

Signorino, Curtis S. 2002. "Strategy and Selection in International Relations." *International Interactions* 28(1):1–3.

Signorino, Curtis S. 2003. "Structure and Uncertainty in Discrete Choice Models." *Political Analysis* 11:316–344.

Signorino, Curtis S. and Ahmer R. Tarar. 2006. "A Unified Theory and Test of Extended Immediate Deterrence." *American Journal of Political Science* .

Signorino, Curtis S. and Kuzey Yilmaz. 2003. "Strategic Misspecification in Regression Models." *American Journal of Political Science* 47(3):551 – 566.

Simmons, Beth A. 1998. "Compliance with International Agreements." *The Annual Review of Political Science* 1:75–93.

Simmons, Beth A. 2009. *Mobilizing for rights : international law in domestic politics*. New York: Cambridge University Press.

Simmons, Beth A. and Daniel J. Hopkins. 2005. "The Constraining Power of International Treaties: Theoryand Methods." *American Political Science Review* 99(4):623–631.

von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99:611–622.

Vreeland, James Raymond. 2008. "Political Institutions and Human Rights: Why Dictatorships enter into the United Nations Convention Against Torture." *International Organization* 62(1):65–101.

Wexler, Lesley. 2003. "International Deployment of Shame, Second-Best Responses, and Norm Entrepreneurship: The Campaign to Ban Landmines and the Landmine Ban Treaty." *Arizona Journal of International and Comparative Law* 20(3):561–606.