

HiCN Households in Conflict Network

The Institute of Development Studies - at the University of Sussex - Falmer - Brighton - BN1 9RE
www.hicn.org

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

Kristian Skrede Gleditsch[†], Simon Hug[‡], Livia Isabella Schubiger[§] and Julian
Wucherpfennig^{**}

ksg@essex.ac.uk; simon.hug@unige.ch; schubiger@pw.uzh.ch; wucherpfennig@icr.gess.ethz.ch

HiCN Working Paper 108

October 2011

Abstract: Non-state actors (NSAs) play an important role in violent conflicts, but unlike 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 solicits NSAs to sign and allow monitoring of conventions banning particular activities, for example the use of landmines. We propose a game-theoretic model to assess the motivations for NSAs (and states) 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 of conflict behavior in countries where Geneva Call has been active support the implications of the theoretical model.

Keywords: non-state actors, human rights, international convention, violent conflict

* Earlier versions of this paper were presented at the CSCW workshop on “Political Institutions and Conflict Resolution” (Oslo November 4-5, 2010), the “Political Economy of International Organizations” conference (Zurich, January 28-30, 2011), the IR Workshop at Yale University (February 23, 2011) and the ECPR Joint Sessions of Workshops (St Gallen, April 12-17, 2011). We gratefully acknowledge helpful comments by the participants at these event and especially Susanna Campbell, Simone Günther, and Dan Wood. 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. Gleditsch would like to acknowledge support from the Research Council of Norway (180441/V10). Schubiger is grateful for the hospitality of Yale University and acknowledges financial support by the Zurich University Alumni Association (FAN fellowship) and the Swiss National Science Foundation (fellowship PBZHP1-133450). Hug and 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, UK, and Centre for the Study of Civil War, PRIO.

[‡] Département de science politique et relations internationales, Faculté des sciences, économiques et sociales, Université de Genève; 40 Bd du Pont d’Arve; 1211 Genève 4, Switzerland and PRIO.

[§] Center for Comparative and International Studies (CIS), Institute of Political Science, University of Zurich, Affolternstrasse 56, 8050 Zurich, Switzerland.

^{**} Center for Comparative and International Studies (CIS), ETH Zurich, Haldeneggsteig 4, 8092 Zurich, Switzerland.

1 Introduction

In his inaugural address on August 7, 2010, the newly elected Colombian President Juan Manuel Santos implored the *Fuerzas Armadas Revolucionarias de Colombia* (FARC) guerillas to cease using landmines.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 unsuccessfully attempted to win the FARC over to its cause, despite the fact that the Colombian government 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 41 non-state actors (NSAs) from 10 countries have done the same for Geneva Call's convention.⁵ Few non-state actors (NSAs) have signed the convention after their government signed the treaty, many more have signed before the government has pledged its support.

This raises two questions relevant to 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 addition they highlight a more general process between NSAs and governments, namely that both play to at the same time to a domestic and international audience. Obviously, this type of interaction is not necessarily dependent on having available conventions for NSAs, but may also exist in a much looser fashion with informal agreements.

While Geneva Call's approach to engage non-state actors in the area of human rights is quite specific, our study allows for insights going beyond this lim-

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

²Geneva Call also wishes to cover the areas of child soldiers and sexual violence.

³See <http://www.icbl.org/index.php/icbl/Universal/MBT/States-Parties>.

⁴See <http://www.icbl.org/index.php/icbl/Universal/MBT/States-Parties>.

⁵We will use the term non-state actor as Geneva Call (2007) does.

ited realm. More precisely we find in our theoretical model and in our empirical analyses that decisions by governments and NSAs are interdependent, and that the latter's decisions influence the effect of reputation costs of governments. This is likely to be found also in other areas where no NGO offers formal deeds of commitments to NSAs, but where NGOs operate through "naming and shaming" (e.g., Hafner-Burton, 2008) to influence both governments and NSAs. We also find that evaluating the effect of both the Ottawa convention and Geneva Call's deed of commitment is considerably influenced by selection effects. More precisely, we only find a positive effect for Geneva Call's deed of commitment on NSAs refraining from using landmines, once we consider the decision to sign this deed of commitment. Similarly, the effect of the Ottawa convention only appears once taking into consideration the signing decisions, but this effect is surprisingly negative: signatories of the convention, taking into account the reasons for signing, are more likely to continue using landmines. The rather weak enforcement regime of the Ottawa convention (e.g., Drezner, 2005) is probably not stranger to this result.

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.

2 Human rights and non-state actors

There has been a surge of academic interest in the study of human rights over the last decades (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 review the recent literature on human rights relevant for our research questions, before offering a short overview over Geneva Call's actions.

2.1 Human rights

Authors like Finnemore and Sikkink (1998) and Risse, Ropp and Sikkink (1999) see the growing importance of human rights norms as clear evidence for sociological institutionalist arguments (see, e.g., March and Olsen, 1984; Dimaggio and Powell, 1991). Recent work focusing on the tangible effects of human rights conventions highlights the critical issue of enforcement. Several studies note how authoritarian regimes have agreed to sign human rights conventions without enforcing them (e.g., Hathaway, 2002; Hafner-Burton, 2008; Vreeland, 2008; Simmons, 2009; Hollyer and Rosendorff, 2010). This raises the question of whether the norm diffusion effects highlighted in earlier studies merely entail states paying lip-service when signing treaties and whether signing treaties by itself has any tangible consequences.

The debates over the effects of human rights treaty ratification and subsequent behavior are related to a more general debate on the effects of international treaties on behavior (see for instance Simmons, 1998; von Stein, 2005; Simmons and Hopkins, 2005). This literature highlights the problem of assessing the constraining effects of international treaties (including treaties related to human rights) arising from the fact that signing a treaty often is influenced by the expected compliance and compliance costs. Consequently, observing that signatories of particular treaties behave differently may simply be due to particular types of countries choosing to sign treaties rather than the effects of treaties for behavior *per se* (see for instance von Stein, 2005).⁶ In the context of human rights work by Hafner-Burton and Tsutsui (2005, 2007), Hafner-Burton (2008), Vreeland (2008), Simmons (2009), Hill (2010), and Hollyer and Rosendorff (2010) deals with these issues more specifically.

NSAs differ from state actors in that their human rights obligations are much less clear (e.g., Clapham, 2006), as NSAs by definition are not signatories to standard human rights conventions. Scholars have only recently become interested in the conditions under which NSAs obey human rights norms (see for

⁶Simmons and Hopkins (2005) contend that there are constraining effects of treaties even after taking this selection process into account. Similar debates have emerged in studies of whether the World Trade Organization leads to trade liberalization or not (see Rose, 2002). In our context, the Sudanese SPLA had already started to refrain from using landmines when it signed Geneva Call's convention (personal communication by Pascal Bongard, program officer Geneva Call, January 5, 2011).

instance Jo and Thomson, 2008).⁷ The most extensive effort in this area is certainly Geneva Call’s initiative to propose human rights conventions to NSAs.⁸ The first convention offered by Geneva Call for NSAs to sign concerns banning landmines, and is intended to be a parallel to the Ottawa convention (see for instance Goose, 1998; Moser-Puangsuwan, 2008).

For this convention, as with the early work on human rights, many studies emphasized the importance of NGOs and civil society to bring about this convention from a sociological institutionalist perspective (e.g., Price, 1998; Short, 1999; Anderson, 2000; Rutherford, 2000*b*; Rutherford, 2000*a*; Wexler, 2003; Lins de Albuquerque, 2007).⁹ More recently, however, scholars have questioned the importance of civil society in this context, as most of the signatories to the Ottawa convention did not stock landmines at the time of ratification and the enforcement mechanisms remain particularly weak (see for instance Drezner, 2005).

This makes it all the more interesting to understand why NSAs would sign a convention imitating the Ottawa convention and how this affects their subsequent human rights record. Few studies have examined these issues, in particular the interaction between governments and NSAs.¹⁰ Jo and Thomson (2008), for instance, propose a theoretical model assessing how compliance with human rights norms relates to reputation and international organizations.¹¹

⁷By “non-state actors,” we refer to armed opposition organizations previously or currently engaged in violent intra-state conflict. As outlined above, other types of social actors figure prominently in the study of human rights, primarily in arguments stressing mechanisms of transnational and domestic mobilization for human rights protection (e.g., Keck and Sikkink, 1998; Simmons, 2009). In general, however, the literature on compliance with international law suffers from a state-centric focus (see Simmons, 2010). An important exception is Morrow (2007), who in his study of state compliance with the laws of war discusses the agency problem introduced by non-compliance by individual soldiers.

⁸See <http://www.genevacall.org/> for more details.

⁹In a recent contribution on global norm creation with regards to specific weapons types, Carpenter (2011) argues that it is not the existence of advocacy networks around an issue per se that explains variation in the emergence and success of campaigns on particular weapons norms. Instead, she argues that processes of issue selection within nongovernmental and international organizations that are most central to advocacy networks (“network hubs”) are crucial in bringing particular issues to global prominence.

¹⁰Variations in treaty compliance during wartime have been studied with regards to the conduct of states and their adherence to the laws of war (Valentino, Huth and Croco, 2006; Morrow, 2007). Another strain of literature deals with the determinants of violence against civilians by states and/or NSAs during wartime more generally (e.g., Azam and Hoefler, 2002; Valentino, Huth and Balch-Lindsay, 2004; Azam, 2006; Downes, 2006; Humphreys and Weinstein, 2006; Kalyvas, 2006; Eck and Hultman, 2007; Balcells, 2010; Stanton, 2010).

¹¹Related is Beber and Blattman’s (2010) work dealing with child-soldiers, an area into which

2.2 Geneva Call

Geneva Call is an NGO that aims at engaging armed NSAs 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, and did not prevent armed NSAs to continue to use these weapons. Geneva Call effectively began in 2000 to engage NSAs 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 NSAs 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 (since 2000), Asia (since 2000), the Caucasus (since 2006), Europe (since 2001), the Middle East (since 2000), and Latin America (since 2003).¹²

Table 1 gives an overview of the numbers of countries and NSAs which have signed the Ottawa treaty and the Geneva Call convention to date.

Table 1: Ratification of Landmine Ban Convention and Sequence

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)	

Geneva Call is also in the process of venturing (see <http://www.genevacall.org/Themes/Children/children.htm>). Another recent contribution is the study of Bussmann and Schneider (2011) analyzing the impact of the ratification of international humanitarian law and the presence of the ICRC in conflict zones on violence against civilians in intra-state armed conflicts.

¹²See <http://www.genevacall.org/home.htm> (accessed September 7, 2010).

3 A model

We propose a game-theoretic model to better understand how Geneva Call's activity on landmines, understood as an equivalent to the Ottawa convention, shapes the interaction between the government and NSAs. The model is based on the interaction between two actors, namely a government G and an NSA N . The sequence of play is as follows:

1. The government G can sign or not sign treaty;¹³
2. The NSA N can sign or not sign treaty;
3. If only the NSA has signed, the government gets another chance to sign or not sign the treaty.

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

- The costs of the civil war cw_i , with $i \in \{G, N\}$;¹⁴
- The increased costs of warfare if a treaty is adhered to w_i , with $i \in \{G, N\}$ (by assumed symmetry, these increased costs generate benefits for the adversary);
- The reputation benefits r_G (if G signs first) or costs (if N signs and G does not).¹⁵

¹³In principle one can envision a case where a state for practical purposes does not exist as an actor, either because an NSA operates in a failed state, or a secession is fought for a state to be created. As NSAs can still be the first actor to sign in our model, we believe this assumption will not affect unduly our results. We thank Simone Günther for alerting us to this point.

¹⁴As 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.

¹⁵More precisely, the costs of not signing the treaty after N '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 NSAs do not face reputation costs or benefits related to human rights conventions. It is likely that NSAs do face such costs, but this will depend strongly on the position of the NSA. NSAs close to winning a war might be concerned much more about civilian victims than other NSAs. We thank Susanna Campbell for raising this point. Moreover, empirically the cost of civil war will also vary across time depending on the level of intensity.

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

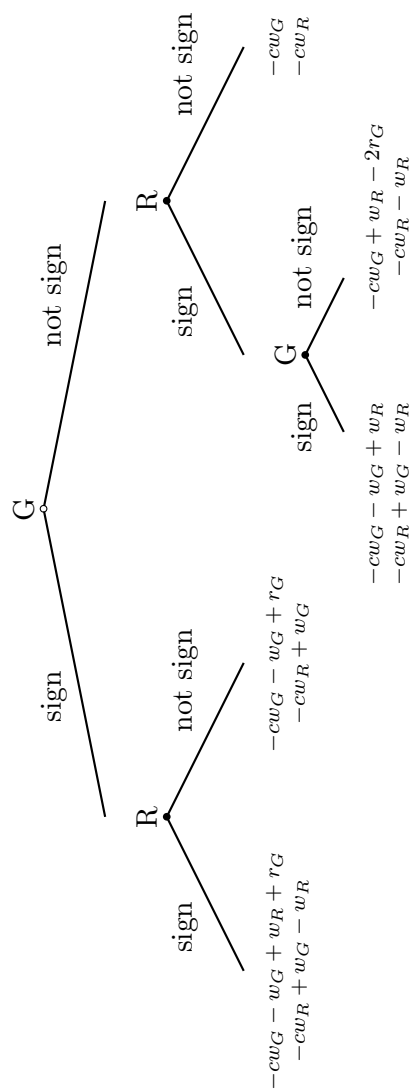


Figure 1 displays the extensive form of this model, including payoffs. This simple structure already generates some insights about under what conditions the NSA N and the government G will chose to sign the agreement. However, this version of the model does not tell us anything about compliance, and whether these agreements have any tangible effects. We therefore extend the model by assuming that the costs of warfare under a treaty w_i depend on compliance (which may be monitored but not directly observed).¹⁶ In particular, we assume that G and N 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, we add:

- The increased costs related to non-compliance c_i , with $i \in \{G, N\}$.

This results in a signaling game with two-sided incomplete information where the effects of the agreements will become endogenous. The decisions to comply by G and N are then reached simultaneously, leading to the game form depicted in figure 2. For simplicity we omit the initial move by Nature to select the two types of G and N .

¹⁶Both the Ottawa convention on baning landmines and Geneva Call’s convention include monitoring provisions (e.g, Geneva Call, 2007).

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

1. Both G and N comply

$$\begin{aligned} EU_G &= -cw_G - w_G + w_N + r_G \\ EU_N &= -cw_N - w_N + w_G \end{aligned} \tag{1}$$

2. Only G complies

$$\begin{aligned} EU_G &= -cw_G - w_G + r_G \\ EU_N &= -cw_N + w_G - c_N \end{aligned} \tag{2}$$

3. Only N complies

$$\begin{aligned} EU_G &= -cw_G + w_N + r_G - c_G \\ EU_N &= -cw_N - w_N \end{aligned} \tag{3}$$

4. Neither G nor N comply

$$\begin{aligned} EU_G &= -cw_G + r_G - c_G \\ EU_N &= -cw_N - c_N \end{aligned} \tag{4}$$

From this setup it easily follows that compliance for both actors i depends on the condition $-w_i > -c_i$ (i.e., the costs of warfare under the treaty must exceed the increased costs of non-compliance). Consequently, we use this condition to define the “mean” and “nice” types of actors. For a “nice” G $-w_G > -c_G$, while for a “mean” G $-c_G > -w_G$. Similarly, for a “nice” N $-w_N > -c_N$ holds, while $-c_N > -w_N$ holds for a “mean” N .¹⁷

¹⁷More precisely, we assume that all payoff elements are common knowledge except the c_i s, which are private information to both i s, respectively.

Consequently, if both G and N 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 ($\text{prob}(c_G > w_G)$) and q ($\text{prob}(c_N > w_N)$). 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.¹⁸

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

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

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

- *Both types of G refrain from signing at each of their decision node, leading both types of N not to sign.*
- *Both types of G sign at their first decision node, leading both types of N not to sign.*
- *Both types of G refrain from signing at their first decision, leading both types of N to sign, followed by both types of G signing as well.*
- *A “mean” G signs at first decision node and a “nice” G does not, leading both types of N not to sign.*
- *A “mean” G signs at first decision node and a “nice” G does not, leading the “mean” type of N to sign, followed by a signing by both types of G as well.*

¹⁸If the more general formulation were to be used, some equilibria would depend on the exact distribution of the two c_i s.

- A “nice” G does not sign first, while the “mean” N does not sign probabilistically, leading the “nice” N to sign probabilistically, while the “mean” N 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 N will ever sign.
- If the prior belief q is high, then both types of G will first refrain from signing, but after N 's signing will sign as well.
- 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 the “mean” N to sign on its turn, followed by G signing.

The first three implications basically focus on the decision to sign a convention, while the last one also has some implications regarding the compliance to be observed in equilibrium. The first implication draws on a comparison of two costs for government namely its reputation costs and the cost of refraining from using landmines. Below we will presume that reputation costs for democracies are higher than for non-democratic countries. The costs of refraining from using landmines may be proxied by previous use or stockpiles of landmines (see <http://www.the-monitor.org/>). For the second and third implication one-sided violence committed by an NSA may provide a proxy for the government's prior belief of facing a “nice” or “mean” NSA. the costs of being caught cheating exceed the compliance costs. As we assume the former to be private information, we start from the presumption that observing one-sided violence by one actor gives the other actor information on how costly it is for the former to engage in non-tolerated activities. Finally, for the last implication the costs of refraining from using landmines might again be proxied by previous landmine use (with the same caveat as above), while the cost of treaty adherence for N can be proxied by the degree of territorial control that an NSA is able to exert.

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 countries in which Geneva Call played an active role and to NSAs in contact with the latter NGO. Moreover, given the setup of our theoretical model, we require data that allow us to model the (strategic) interaction between NSAs and their governments. To retain time varying information, our unit of analysis is the 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 NSA has been involved in intra-state armed conflict (as defined by UCDP¹⁹) at least once since 1989. More precisely, dyads are included if the NSA has been actively involved in armed hostilities with the government, i.e., in intra-state conflict as defined by UCDP,²⁰ 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 NSAs 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.²² However, we only include NSAs as long as they qualify as politically active organizations that maintain their own

¹⁹“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).

²⁰UCDP Dyadic Dataset v.1-2011 (Harbom, Melander and Wallensteen, 2008; Harbom, 2010; Themnr and Wallensteen, 2011 (forthcoming)).

²¹Intra-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 NSAs 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).

²²To illustrate, the conflict between the *Mouvement des forces démocratiques de Casamance* (MFDC) and the government of Senegal was coded active first in 1990 in the UCDP dyadic dataset (v. 1-2010). This dyad 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.

armed wing (our coding effort) and have been engaged by Geneva Call. 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; see table 7 in the appendix).

Naturally, we restrict this sample to regions and periods of Geneva Call's engagement. These are listed below. The countries where Geneva Call has already ended its programs are listed in parentheses.²³

Regions and time periods of Geneva Call's engagement:²⁴

- 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],²⁵ Georgia.
- Europe (2001 onwards): Turkey.
- Latin America (2003 onwards): Colombia.
- Middle East (2000 onwards): Iran, Iraq, Lebanon,²⁶ Yemen.

4.2 Variables

In our analyses we rely on two sets of dependent variables. The first set captures whether (or not) the NSA has signed Geneva Call's deed of commitment

²³Note 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.

²⁴Source: Geneva Call. <http://www.genevacall.org/home.htm> (accessed September 11, 2010).

²⁵Armenia does not qualify as a primary conflict party during the period of investigation. Rather than having been directly challenged by a non-state actor itself, Armenia supported the pro-independence movement in Nagorno-Karabakh against Azerbaijan (UCDP Database, Uppsala Conflict Data Program, accessed October 24, 2010: <http://www.ucdp.uu.se/gpdata/gpcountry.php?id=6®ionSelect=9-Eastern.Europe>, Uppsala University.)

²⁶In Lebanon, Geneva Call is mainly in contact with organizations affiliated with Hezbollah. Therefore, the Israel-Hezbollah dyad is included in the dataset.

banning anti-personnel (AP) mines during a given year ($mbtreaty_{nsa}$).²⁷ For the government, the corresponding variable denotes ratification of the international mine-ban treaty ($mbtreaty_{gov}$). The second set deals with compliance and relies on information on mine use as provided by IISS.²⁸ This measure is, however, fraught with some problems, as the indicator covers all types of mines and other “improvised explosive devices.” Nevertheless, we consider it the best measure available.²⁹

We employ several additional variables in our analyses:

Territorial control is a dummy variable denoting whether the NSA exerts at least a moderate level of control over its main territory.³⁰ As outlined above, we argue that this variable is related to w_N , the costs induced by treaty adherence. The logic is simple. Landmines are an effective way of securing territory from governmental intrusion, hence relinquishing their usage is likely to make the NSA more vulnerable since it removes an effective military strategy from its portfolio.

Use of mines by government indicates whether landmines and improvised explosive devices were among the weapons used by the state actor of this dyad during any year between 1997 and the start year of Geneva Call’s engagement in the respective region.³¹

OSV denotes the extent to which NSAs or governments were responsible for one-sided violence according to UCDP.³² The variables indicate the best estimate of the aggregated estimated fatalities for all incidents of one-sided violence for a

²⁷Source: Geneva Call, <http://www.genevacall.org/resources/list-of-signatories/list-of-signatories.htm> (accessed September 16, 2010).

²⁸Source: IISS Armed Conflict Database, http://acd.iiss.org/armedconflict/report/dsp_MainForm.asp (accessed February 11, 2011).

²⁹Geneva Call also assesses compliance and deploys monitoring missions. So far these have, however, not been carried out in a systematic and recurrent fashion (personal communication by Pascal Bongard, program officer Geneva Call, January 5, 2011).

³⁰Information on territorial control is largely adopted from Cunningham, Gleditsch and Salehyan (2009).

³¹Source: IISS Armed Conflict Database, http://acd.iiss.org/armedconflict/report/dsp_MainForm.asp (accessed February 11, 2011).

³²To 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).

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. Instances of one-sided violence were not assigned to a conflict-year if the perpetrating actor did not constitute one of the primary conflict parties in the respective countries according to UCDP/PRIO-criteria.³³ Accordingly, the fatality estimates 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 1989.³⁴ We employ *OSV* as a proxy for p and q , i.e., the prior beliefs. With regard to one-sided violence serving as proxy for prior beliefs this assumes that higher such violence suggests being caught in violation of a treaty is less costly (i.e., smaller c_i s).

Democracy is the Cheibub, Gandhi and Vreeland’s (2010) democracy indicator and is meant to proxy reputation effects.³⁵

To capture size-related effects, such as military capacity, we also use an estimate of the *troop size* of the NSA.³⁶

³³Exceptions are militias that allegedly acted on behalf of - or supported by - the state (Janjaweed in Sudan, Autodefensas Unidas in Colombia) and were therefore attributed to the state actor of a given dyad-year. The following perpetrators of osv active in Geneva Call’s countries of engagement do not enter our sample as primary conflict parties: Colombia: Autodefensas Campesinas de Córdoba y Urabá, Medellín Cartel; India: Bodo Liberation Tiger Force, Dima Halam Daogah, Hmar People’s Convention, Indian Mujahideen, Kuki Revolutionary Army, Lashkar-e-Taiba, Ranvir Sena, Students’ Islamic Movement, United People’s Democratic Solidarity, Vishwa Hindu Parishad; Indonesia: Jemaah Islamiya; Iraq: Gov of USA, Asa’ib Ahl al-Haq, Jamaat Jund al-Sahaba; Lebanon: Gov. of USA; Morocco: Salafia Jihadia; Niger: Gov. of Lybia; Philippines: Ampatuan Militia; Senegal: MFDC - Front Nord, Somalia: Rahanweyn Resistance Army, SPM - Somali National Alliance; Sudan: South Sudan Defence Force, Lord’s Resistance Army; Turkey: Abu-Hafs al-Masri Brigades.

³⁴To 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 (see sample definition, section 4.1.). Similarly, the actor “Hutu rebels” encompasses more than one NSA involved in intra-state conflict (e.g., Palipehutu and Palipehutu-FNL) (see Harbom and Sundberg, 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 (see section 4.1.).

³⁵As this data series ends in 2008, we extrapolated for 2009.

³⁶Information on troop size is adopted from Cunningham, Gleditsch and Salehyan (2009) as well as the IISS Armed Conflict Database.

Table 2: Signatories and “Follow-Suit” Signatories: number of countries, number of dyads (% of countries, % of dyads)

	NSA has not signed	
	Government has signed	Government has not signed
NSA does not sign	6, 54 (75.0, 96.4)	13, 115 (61.9, 93.2)
NSA signs	2, 2 (25.0, 3.6)	8, 11 (38.1, 6.8)
Total	8, 56 (100.0, 100.0)	21, 126 (100.0, 100.0)

	Government has not signed	
	NSA has signed	NSA has not signed
Government does not sign	34 (91.2)	146 (96.6)
Government signs	3 (8.8)	5 (3.4)
Total	37 (100.0)	146 (100.0)

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.³⁷ 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 NSAs.

We begin with some descriptive statistics given in table 2. The table contains information on the ratification sequence of both governments and NSAs of the convention on landmines. It indicates the numbers of dyad-years that correspond to the respective signatory status of governments or NSAs (dyad-years following signature are dropped). Depicted in parentheses are the respective numbers of cases where the “other” actor has previously ratified. Substantially, the table suggests that an NSA’s probability of signing the Geneva Call convention is larger if the government has failed to sign than if the government has signed the Ottawa convention. The differences are, however, rather small. Conversely, if an NSA signs first it increases the probability of a government signing from 0.03 to 0.09 in our strict sample. Consequently, once NSAs have signed the Geneva Call convention, governments are much more likely to follow suit than in the reverse scenario (i.e., when governments sign first).

Next we assess with a series of (corresponding) logit models whether the implications concerning the signing find empirical support in our data.³⁸ As many

³⁷Geneva Call has provided us with a list of NSAs contacted by them. As reported by Pascal Bongard (program officer Geneva Call, personal communication of January 5, 2011) the selection focuses essentially on NSAs that have employed landmines. Not surprisingly in exploratory analyses we find that in our dataset the engagement decision by Geneva Call is heavily influenced by an NSA’s use of landmines.

³⁸We employed Gelman and Hill’s (2006) Bayesian logit model, as in some instances we encountered problems of complete separation (we will note these when discussing the results). In addition we also controlled for time dependence by following Carter and Signorino’s (2010) suggestion. As these corrections do not affect our results we report these results in the appendix.

scholars have noted, the strategic nature of the decisions we wish to explain (signing of a treaty by government and NSA) creates statistical estimation problems. While estimators are rather well understood for complete information models (see for instance Signorino, 1999; Signorino, 2002; Signorino and Yilmaz, 2003; Signorino, 2003; Signorino and Tarar, 2006), for incomplete information models as ours only few models exist (see for instance Lewis and Schultz, 2003; Esarey, Mukherjee and Moore, 2008; Whang, 2010). In addition, our theoretical model is not a simple stage game of a more complex repeated one, as the former changes as a function of the outcome of the previous iteration. For this reason we rely essentially on simply estimating for each relevant decision node in figure 2 the effect of various variables on the decisions of a given actor.

Table 3: Bayesian Logit Estimates of Signatory Status of Government

sample:	government signs		government signs before NSA		government signs after NSA	
	<i>strict</i> Model 1	<i>lenient</i> Model 2	<i>strict</i> Model 3	<i>lenient</i> Model 4	<i>strict</i> Model 5	<i>lenient</i> Model 6
constant	(0.737)	(0.719)	(1.055)	(0.887)	-1.804	-2.074
democracy(lagged)	1.005	1.166	(0.980)	(0.864)	(1.756)	(1.800)
one-sided violence _G (logged)	(0.736)	(0.726)	-0.043	-0.014	-0.948	-0.806
one-sided violence _N (logged)	0.019	0.030	(0.039)	(0.034)	(1.513)	(1.562)
use of mines by government	(0.029)	(0.027)	0.014	0.001	(0.053)	(0.053)
	-0.012	-0.016	(0.040)	(0.039)	-0.020	-0.016
	(0.036)	(0.036)			(0.078)	(0.080)
	(0.790)	(0.774)	(0.959)	(0.926)	-0.623	-0.635
					(1.102)	(1.062)
<i>N</i>	185	223	151	177	39	51
<i>log L</i>	-13.214	-17.054	-0.444	-5.053	7.836	6.944

Standard errors in parentheses
* indicates significance at $p < 0.1$

Table 3 focuses on the government’s decision to sign the Ottawa convention, and models 1 and 2 simply regress this decision on a set of independent variables without taking into account where in the game tree the decision to sign is reached. Both for the strict (model 1) and lenient (model 2) sample definition the results reported in table 3³⁹ we find that especially the signing costs, i.e., whether the government has used landmines in the past, affect negatively the decision of the government to sign the Ottawa convention. We also find that the reputation costs as measured by the country’s status as democracy positively (though statistically not significantly) affects the government’s decision. The extent to which the government or the NSA commit one-sided violence, on the other hand, hardly seems to affect a government’s decisions.

In addition, given that our dyads are grouped by country we have also, in a previous version, employed clustered standard errors, which affected, however, unnoticeably our results.

³⁹Table 8 reports the results of the same models including, however, controls for time-dependency. The substantive insights from these modes are the same.

The estimates of these first two models rely on the assumption that governments decide independently of what the NSA is doing. The next set of models relax this assumption and focus on the government’s decision in two situations, namely when the NSA has not yet signed and when it considers signing after the NSA’s signing. As models 3-6 demonstrate, these situations differ quite importantly. First of all, the results of models 3 and 4 show that both signing (use of mines) and reputation (democracy) costs statistically significantly affect the government’s decision to sign before an NSA has done so. This is in line with our implications from the theoretical model, namely that high such costs should reduce, respectively increase, the likelihood of a government signing the Ottawa convention. When the latter decides, however, after the NSA has signed, then neither reputation nor signing costs seem to matter (models 5 and 6).⁴⁰ This seems to suggest that these costs serve as a screening device, i.e., they affect the government’s decision at the first node and the NSAs reaction (taking into account this decision) leads to situations where these costs have already screened out potential governmental signers.

Contrary to another implication from our theoretical model we find that one-sided violence does barely affect a government’s decision. According to our model one-sided violence by the NSA, proxying for the government’s prior belief of facing a “mean” type should have increased the likelihood of signing at the first decision node. While we find a positive effect as predicted (models 3 and 5), these do not reach statistical significance. We find also, however, that as governments commit more one-sided violence, they are more likely to sign if the NSA already has. This might be linked to the increased reputation effects that we have assumed in our theoretical model in situations where an NSA signs before the government does.

Table 4 presents the results for the NSAs. The implications that we have presented for the NSAs’ actions involve again the prior beliefs and the costs faced by government. Both should operate, however, strongly through the government’s prior decision⁴¹ Consequently we focus on the prior beliefs, proxied by one-sided violence, and the costs faced by the NSA. The first two models in table 4 are again based on all decision nodes where the NSA can decide on signing Geneva

⁴⁰In these models we encounter a problem of complete separation. Non-democratic countries never sign in this instance (but there are very few cases, and certain values of one-sided violence by the NSA also perfectly predict the outcome.

⁴¹In preliminary analyses we also found that the signing and reputation costs of the government do not directly affect the NSA’s decision.

Table 4: Bayesian Logit Estimates of Signatory Status of NSA

sample:	NSA signs		NSA signs before government	
	<i>strict</i> Model 1	<i>lenient</i> Model 2	<i>strict</i> Model 3	<i>lenient</i> Model 4
one-sided violence _G (logged)	0.016 (0.036)	0.015 (0.036)	0.022 (0.041)	0.022 (0.041)
one-sided violence _N (logged)	0.017 (0.036)	0.020 (0.036)	-0.012 (0.043)	-0.012 (0.043)
territorial control \geq moderate	(0.974)	(0.978)	-1.689 (1.112)	-1.712 (1.112)
NSA troop size (logged)	(0.433)	(0.422)	(0.574)	(0.577)
constant	(3.951)	(3.844)	(5.126)	(5.158)
N	150	156	88	90
$\log L$	-2.338	-2.479	5.036	5.033

Standard errors in parentheses
* indicates significance at $p < 0.1$

Call’s deed of commitment. The results provide strong evidence that the costs of relinquishing landmines (territorial control and troop size) strongly affect an NSA’s decision to sign. If its territorial control is moderate or high, the likelihood of signing decreases considerably, while larger NSAs are more likely to sign. Again, however, one-sided violence, our proxy for prior beliefs, seems not to affect the NSA’s decision. When only considering the NSAs decision when the government has not yet signed, it seems that especially the size of the NSA increases the likelihood of the NSA signing. The effect of territorial control is slightly reduced and loses statistical significance.

4.4 Evaluating the effectiveness

Having established these patterns, we now turn to evaluating the effect of such conventions. In tables 5 and 6 we assess how various variables affect the governments’ and the NSAs’ decision to use landmines. As discussed above, we rely on the IISS Armed Conflict Database (see above) to measure compliance. While this is an imperfect measure, as IISS also includes improvised exploding devices in their coding, it is in our view the best measure available.⁴² In both tables we first report for the two sample definitions how signing a convention or Geneva Call’s deed of commitment affects the two actors’ behavior, while neglecting the fact that signatories are potentially endogenous to their effect. For the government’s decision to comply (table 5) we find that its signatory status and that of the NSA has no effect on compliance. We only find that prior use of landmines (before Geneva Call’s engagement) considerably reduces the likelihood of compli-

⁴²It is important to note here that when we used information on mine use to explain the signing decision, we only considered this up to the point of signing.

ance. Table 6 shows, however, that for the NSA signatory status has an impact. If an NSA signs it is much more likely to refrain from using landmines. Interestingly enough the signatory status of the government has the reverse effect. If the government has signed the Ottawa convention, then the NSA is much more likely to use landmines. The other variables we employed to explain compliance appear to have no statistically significant effect.

Table 5: Probit of Government's Compliance

sample: dependent variable:	Model 1	Model 2	Model 3		
	<i>probit</i> <i>strict</i>	<i>probit</i> <i>lenient</i>	trivariate probit (CMP) strict=lenient		
	mine use _G				
government signed Ottawa convention	0.312 (0.700)	-0.233 (0.306)	(0.460)		
NSA signed Geneva Call's deed of commitment	-0.562 (0.554)	-0.310 (0.256)	-0.199 (0.290)		
democracy(lagged)	-0.540 (0.559)	0.009 (0.254)	(0.297)	(0.242)	
use of mines by government	(0.584)	(0.242)	(0.266)	(0.241)	
one-sided violence _N				-0.013 (0.012)	-0.030 (0.019)
one-sided violence _G				-0.006 (0.010)	0.030 (0.015)
territorial control \geq moderate					(0.515)
NSA troop size (logged)					0.828 (0.240)
constant	(0.513)	(0.217)	(0.262)	(0.262)	(2.117)
ρ_{12}					(0.232)
ρ_{13}					-0.156 (0.257)
ρ_{23}					(0.197)
<i>N</i>	235	282	247		
log <i>L</i>	-47.719	-57.896	-183.766		

Standard errors in parentheses
* indicates significance at $p < 0.1$

These results, while interesting, neglect that the signing of a treaty is hardly exogenous and both governments and NSAs choose strategically to sign (or not sign) a convention. We take this into account in models 3 of tables 5 and 6 by estimating a trivariate probit using CMP (Roodman, 2007) which deals with the endogenous nature of the government's and the NSA's signatory status.⁴³ To model the latter we reuse the same independent variables as above and find largely similar results. Interesting to note, however, is that in both tables we find that one-sided violence by the NSA, i.e. our proxy for prior belief, affects negatively the NSAs signing decision. Although the two coefficients fail to reach

⁴³This trivariate probit model allows for correlations between the error terms of three equations, namely the ones corresponding to the two signing decisions and one for the compliance decision. As Wilde (2000) has shown for the bivariate probit, such a model also allows for endogenous regressors, in our case the two signing decisions (see also Roodman, 2007). For these models the strict and lenient sample definitions collapse, as exactly the same cases are selected.

statistical significance, they point in the right direction, namely if an NSA is likely to be a mean one, the governments are more likely to sign first,⁴⁴ which induces NSAs to refrain from signing.

For the effects of the signatory status we now find partly different results. When taking into account the endogeneity of the signatory status it appears that the government's status actually increases the likelihood of its using landmines. On the other hand if the NSA has signed this reduces slightly (and statistically non-significantly) the use of landmines by government. For the NSA's compliance decision (table 6) we find as before that the signing of Geneva Call's deed of commitment reduces the likelihood of using landmines by an NSA. This effect just barely reaches statistical significance. On the other hand, the government's signatory status increases considerably the likelihood of the NSA using landmines.

Table 6: Probit of NSA's Compliance

sample: dependent variable:	Model 1 <i>probit</i> <i>strict</i>	Model 2 <i>probit</i> <i>lenient</i>	Model 3 trivariate probit (CMP) strict=lenient	
government signed Ottawa convention				
NSA signed Geneva Call's deed of commitment	(0.476)	(0.414)	(0.523)	
territorial control \geq moderate	(0.458)	(0.436)	(0.807)	
NSA troop size (logged)	-0.369 (0.361)	-0.464 (0.347)	-0.432 (0.418)	(0.515)
use of mines by government	0.114 (0.192)	0.031 (0.186)	0.217 (0.207)	(0.238)
democracy(lagged)	(0.364)	(0.314)	(0.368)	(0.239)
one-sided violence _N			(0.240)	-0.013 -0.036
one-sided violence _G			(0.012)	(0.022)
constant	-1.161 (1.578)	-0.077 (1.487)	-2.466 (1.618)	-1.118 (0.263) (2.092)
ρ_{12}				(0.296)
ρ_{13}				0.163
ρ_{23}				(0.452)
N	133	138	247	
$\log L$	-63.221	-71.767	-184.27653	

Standard errors in parentheses
* indicates significance at $p < 0.1$

⁴⁴This effect we fail to find in our empirical results, but this might be due to the fact that we do not distinguish between the government before or after the NSA.

5 Discussion and conclusion

This paper offers a first look and assessment of how NSAs decide whether to sign Geneva Call's convention on the ban of landmines and its effectiveness. We propose 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, for some of which we provide empirical tests of in the present paper.

Our empirical analyses provide clear evidence that the decisions by governments and NSAs to sign a landmine ban convention are not independent. Especially governments' decisions to sign appear to be notably affected by whether NSAs have already signed. Surprisingly, for the governments' decision to sign first its assessment of possible compliance by NSAs appears not 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 also instructive. We only find effects if we consider the decision of signing a convention as endogenous. When dosing so we can show that the decision by an NSA to sign a deed of commitment reduces its use of landmines. For the governments, however, after taking into account that it is mostly countries having not used landmines that signed the Ottawa convention, we find a negative effect on compliance.

These results speak to the broader literature on human rights conventions and the effectiveness of treaties. Our analyses clearly suggests that assessing the effectiveness of conventions needs to take into consideration the selection effects linked to the conscious decisions of actors signing these conventions. Our results also suggest that NSAs influence the domestic audience costs of government. In the empirical realm that we covered where an NGO actively intervenes to influence NSAs this is to be expected, but we surmise that NSAs and NGOs influence these audience costs also in other circumstances through more subtle indirect means.

Appendix

In this appendix we provide proofs of the propositions presented in the main text and some information on the data used in our empirical analysis.

Proofs

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.⁴⁵

1. Observation

If G signs at its first decision node, N will never sign, since it obtains the benefit of compliance by G for free, or cannot 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:

$$\begin{aligned}
 EU_N(\text{sign}) &= p' \times (-cw_N - q \times w_N + w_G - (1 - q) \times c_N) \\
 &= +(1 - p') \times (-cw_N - q \times w_N + (1 - q) \times c_N) \\
 &= -cw_N - q \times w_N - (1 - q)c_N + p' \times w_G \quad (5)
 \end{aligned}$$

$$\begin{aligned}
 EU_N(\text{not sign}) &= p' \times (-cw_N + w_G) + (1 - p') \times (-cw_N) \\
 &= -cw_N + p' \times w_G \quad (6)
 \end{aligned}$$

As w_N and c_N are both positive, independent of q N will never sign. *QED.*

2. Observation

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

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

⁴⁵For simplicity's sake we consider situations where actors are indifferent between two actions only when assessing whether semi-pooling equilibria may exist.

$$\begin{aligned}
EU_G(\text{sign}) &= q' \times (-cw_G - w_G + w_N) + (1 - q') \times (-cw_G - w_G) \\
&= -cw_G - w_G + q' \times w_N
\end{aligned} \tag{7}$$

$$\begin{aligned}
EU_G(\text{not sign}) &= q' \times (-cw_G - 2 \times r_G + w_N) + (1 - q') \times (-cw_G - 2 \times r_G) \\
&= -cw_G - 2 \times r_G + q' \times w_N
\end{aligned} \tag{8}$$

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

If G is “mean” (i.e., $p = 0$)

$$\begin{aligned}
EU_G(\text{sign}) &= q' \times (-cw_G + w_N - c_G) + (1 - q') \times (-cw_G - c_G) \\
&= -cw_G - c_G + q' \times w_N
\end{aligned} \tag{9}$$

$$\begin{aligned}
EU_G(\text{not sign}) &= q' \times (-cw_G - 2 \times r_G + w_N) + (1 - q') \times (-cw_G - 2 \times r_G) \\
&= -cw_G - 2 \times r_G + q' \times w_N
\end{aligned} \tag{10}$$

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.⁴⁶ 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 N is identical, it can anticipate its payoff, namely if $2 \times r_G > w_G$ and $q = 1$

$$\begin{aligned}
EU_N(\text{sign}) &= p' \times (-cw_N + w_G - w_N) + (1 - p') \times (-cw_N - w_N) \\
&= -cw_N - w_N + p' \times w_G
\end{aligned} \tag{11}$$

$$EU_N(\text{not sign}) = -cw_N \tag{12}$$

⁴⁶This follows from the fact that $p = 0$ implies $w_G > c_G$.

Consequently a “nice” N signs in that case if $p' > \frac{w_N}{w_G}$. For a “mean” N the payoffs are as follows:

$$\begin{aligned} EU_N(\text{sign}) &= p' \times (-cw_N + w_G - c_N) + (1 - p') \times (-cw_N - c_N) \\ &= -cw_N - c_N + p \times w_G \end{aligned} \quad (13)$$

$$EU_N(\text{not sign}) = -cw_N \quad (14)$$

Consequently a “mean” N signs in that case if $p' > \frac{c_N}{w_G}$

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

$$EU_N(\text{sign}) = -cw_N - w_N \quad (15)$$

$$EU_N(\text{not sign}) = -cw_N \quad (16)$$

As $w_N > 0$ N will never sign. For a “mean” type

$$EU_N(\text{sign}) = -cw_N - c_N \quad (17)$$

$$EU_N(\text{not sign}) = -cw_N \quad (18)$$

is relevant and as $c_N > 0$ N will never sign.

Complete and imperfect information

Proof of Proposition 1 ⁴⁷

Based on the observations above the following subgame-perfect equilibrium can be established:⁴⁸

1. If $p = 1, q = 1$ and $w_G > 2 \times r_G, w_G > w_N, r_G > w_N$ $G : \{\text{sign}, \text{sign}, \text{comply}\},$
 $R : \{\text{not sign}, \text{sign}, \text{comply}\}$ ⁴⁹

⁴⁷We refrain from presenting the equilibria of the complete and perfect information game depicted in figure 1 as these are closely related to the ones presented here.

⁴⁸For simplicity's sake we omit cases where either of the two actors is indifferent between his or her to actions.

⁴⁹For simplicity's sake we shorten the strategies for both actors by only stating their actions at their first two decision nodes and indicating with the third element the action taken at their remaining decision nodes, as these do not vary.

2. If $p = 1, q = 1$ and $w_G > 2 \times r_G, w_G > w_N, w_N > r_G$ $G : \{not\ sign, sign, comply\}$,
 $R : \{not\ sign, sign, comply\}$
3. If $p = 1, q = 1$ and $w_G > 2 \times r_G, w_G < w_N$ $G : \{not\ sign, signcomply\}$,
 $R : \{not\ sign, not\ sign, .\}$
4. If $p = 1, q = 1$ and $2 \times r_G > w_G, r_G > w_G$ $G : \{sign, not\ signcomply\}$,
 $R : \{not\ sign, not\ sign, .\}$
5. If $p = 1, q = 1$ and $2 \times r_G > w_G, w_G > r_G$ $G : \{not\ sign, not\ signcomply\}$,
 $R : \{not\ sign, not\ sign, .\}$
6. If $p = 0, q = 1$ and $2 \times r_G > c_G, r_G > c_G$ $G : \{sign, sign, not\ comply\}$,
 $R : \{not\ sign, not\ sign, comply\}$
7. If $p = 0, q = 1$ and $2 \times r_G > c_G, c_G > r_G$ $G : \{not\ sign, sign, not\ comply\}$,
 $R : \{not\ sign, not\ sign, comply\}$
8. If $p = 0, q = 1$ and $c_G > 2 \times r_G, r_G > c_G$ $G : \{sign, not\ sign., not\ comply\}$,
 $R : \{not\ sign, not\ sign, comply\}$
9. If $p = 0, q = 1$ and $c_G > 2 \times r_G, c_G > r_G$ $G : \{not\ sign, not\ sign., not\ comply\}$,
 $R : \{not\ sign, not\ sign, comply\}$
10. If $p = 1, q = 0$ and $w_G > 2 \times r_G$ $G : \{not\ sign, not\ sign, comply\}$, $R :$
 $\{not\ sign, not\ sign, not\ comply\}$
11. If $p = 1, q = 0$ and $2 \times r_G > w_G, c_N > w_G, r_G > w_G$ $G : \{sign, sign, comply\}$,
 $R : \{not\ sign, not\ sign, notcomply\}$
12. If $p = 1, q = 0$ and $2 \times r_G > w_G, c_N > w_G, w_G > r_G$ $G : \{not\ sign, sign, comply\}$,
 $R : \{not\ sign, not\ sign, notcomply\}$
13. If $p = 1, q = 0$ and $2 \times r_G > w_G, w_G > c_N, r_G > w_G$ $G : \{sign, sign, comply\}$,
 $R : \{not\ sign, sign, notcomply\}$
14. If $p = 1, q = 0$ and $2 \times r_G > w_G, w_G > c_N, w_G > r_G$ $G : \{not\ sign, sign, comply\}$,
 $R : \{not\ sign, sign, notcomply\}$
15. If $p = 0, q = 0$ and $r_G > c_G$ $G : \{sign, sign, not\ comply\}$, $R : \{not\ sign, not\ sign, .\}$

16. If $p = 0, q = 0$ and $2 \times r_G > c_G, c_G > r_G$ $G : \{not\ sign, sign, not\ comply.\}$,
 $R : \{not\ sign, not\ sign, not\ comply.\}$
17. If $p = 0, q = 0$ and $c_G > 2 \times r_G$ $G : \{not\ sign, not\ sign, not\ comply.\}$,
 $R : \{not\ sign, not\ sign, not\ comply.\}$

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 observations:

1. Observation

If G does not sign at the first decision node and N 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 N must evaluate the following expected utilities:

$$EU_N(sign|q = 1) = -cw_N - w_N \quad (19)$$

$$EU_N(sign|q = 0) = -cw_N - \frac{w_N}{2} \quad (20)$$

$$EU_N(not\ sign|q = .) = -cw_N \quad (21)$$

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (22)$$

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (23)$$

$$EU_G(\text{not sign}|p = .) = -cw_G \quad (24)$$

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 when the following conditions 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 N ’s decision to sign, while a “nice” type will not. As not signing by G has the same consequences for N as not complying, the expected utility calculations both for N 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 N ’s decision to sign. Consequently N evaluates the following expected utilities where the updated belief p' is identical to the prior belief, giving the assumption of a pooling equilibrium:

$$\begin{aligned} EU_N(\text{sign}|q = 1) &= p'(-cw_N - w_N + w_G) + (1 - p')(-cw_N - w_N) \\ &= -cw_N - w_N + p' \times w_G \end{aligned} \quad (25)$$

$$\begin{aligned} EU_N(\text{sign}|q = 0) &= p'(-cw_N - c_N + w_G) + (1 - p')(-cw_N - \frac{w_N}{2}) \\ &= -cw_N - \frac{w_N}{2} + p' \times w_G \end{aligned} \quad (26)$$

$$EU_N(\text{not sign}|q = .) = -cw_N \quad (27)$$

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (28)$$

$$EU_G(\text{not sign}|p = .) = -cw_G \quad (29)$$

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (30)$$

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (31)$$

$$EU_G(\text{not sign}|p = .) = -cw_G \quad (32)$$

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (33)$$

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 are in contradiction with the assumption that $2 \times r_G > w_G$ no pooling equilibrium exists.

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

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (34)$$

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

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

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (36)$$

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

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

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (38)$$

$$\begin{aligned} EU_G(\text{not sign}|p = 1) &= q(-cw_G - w_G + w_N) + (1 - q)(-cw_G - w_G) \\ &= -cw_G - w_G + qw_N \end{aligned} \quad (39)$$

Consequently, a “nice” G will prefer signing if $r_G > qw_N$ or if $q < \frac{r_G}{w_N}$. For a “mean” G the following is relevant:

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (40)$$

$$\begin{aligned} EU_G(\text{not sign}|p = 1) &= q(-cw_G - \frac{w_G}{2} + w_N) + (1 - q)(-cw_G - \frac{w_G}{2}) \\ &= -cw_G - \frac{w_G}{2} + qw_N \end{aligned} \quad (41)$$

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

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

Let’s next assume that $2 \times w_G > w_N > w_G$ and $p' = 1$, implying that only the “mean” N will sign, which implies the following:

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (42)$$

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

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:

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (44)$$

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

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

- $2 \times r_G > w_G$

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (46)$$

$$EU_G(\text{not sign}|p = .) = -cw_G$$

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (47)$$

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

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (48)$$

$$EU_G(\text{not sign}|p = .) = -cw_G \quad (49)$$

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (50)$$

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 the 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 N signs. But then N will neither sign. Hence, for this condition to allow for a separating equilibrium to exist the following has to hold:

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

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

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

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 N ’s signing, while the “nice” G will not. This induces N 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 N ’s decision to sign. As shown above in this situation a “nice” N will sign if $p' > \frac{w_N}{w_G}$ while a “mean” N will do the same if $p' > \frac{w_N}{2 \times w_G}$ holds. Consequently, a series of configurations have to be evaluated.

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (54)$$

$$EU_G(\text{not sign}|p = .) = -cw_G \quad (55)$$

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (56)$$

$$(57)$$

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:

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (58)$$

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

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:

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (60)$$

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

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:

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

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

$$EU_G(\text{sign}|p = 1) = -cw_G - c_G + r_G \quad (62)$$

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

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

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (64)$$

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

which implies that signing requires $r_G > q \times \frac{w_G}{2}$. This, however, is only possible if $q < \frac{r_G}{w_N}$, which is in contradiction with the condition for the “nice” G to not sign. Consequently, 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 is 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 N will also refrain from signing. Thus the following expected utilities becomes relevant:

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (66)$$

$$EU_G(\text{not sign}|p = 1) = -cw_G \quad (67)$$

Thus a “nice” G will sign if $r_G > w_G$ which is in contradiction with the assumption $w_G > 4 \times r_G$. Consequently 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 N will also refrain from signing, so that the same conditions should hold for a separating equilibrium, which are again in contradiction with $4 \times r_G > w_G > 2 \times r_G$. Consequently, no separating equilibrium can exist.

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

$$EU_G(\text{sign}|p = 1) = -cw_G - w_G + r_G \quad (68)$$

$$EU_G(\text{not sign}|p = .) = -cw_G \quad (69)$$

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (70)$$

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:

$$\begin{aligned} p(\text{not sign}|p = 1) &= 1 \\ p(\text{not sign}|p = .) &= s \end{aligned} \quad (71)$$

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

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

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

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

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

$$\begin{aligned}
EU_G(\text{not sign}|p = 1) &= q(t(-cw_G - w_G + w_N) + (1-t)(-cw_G)) \\
&\quad + (1-q)(-cw_G - w_G) \\
&= -cw_G - w_G + q \times w_G + q \times t \times w_N - q \times t \times w_G \tag{74}
\end{aligned}$$

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

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

$$\begin{aligned}
EU_G(\text{not sign}|p = 0) &= q\left(t\left(-cw_G - \frac{w_G}{2} + w_N\right) + (1-t)(-cw_G)\right) \\
&\quad + (1-q)\left(-cw_G - \frac{w_G}{2}\right) \\
&= -cw_G - \frac{w_G}{2} + q\frac{w_G}{2} + q \times t \times w_N - q \times t \frac{w_G}{2} \tag{76}
\end{aligned}$$

As the “mean” type has to be indifferent the following has to hold:

$$r_G = -cw_G - \frac{w_G}{2} + q\frac{w_G}{2} + q \times t \times w_N - q \times t \frac{w_G}{2} \tag{77}$$

which implies

$$t = \frac{r_G - q \frac{w_G}{2}}{2q(w_N - \frac{w_G}{2})} \quad (78)$$

t will be positive if $q < \frac{2 \times r_G}{w_G}$ which will always be the case given that we assume that $2 \times r_G > w_G$. And t will be smaller than 1 if $r_G < w_N$. 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:

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

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

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

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

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

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

Consequently, 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:

$$EU_G(\text{sign}|p = 0) = -cw_G - \frac{w_G}{2} + r_G \quad (82)$$

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

As the “mean” type has to be indifferent the following has to hold:

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

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.

Empirics

Table 7 provides a list of the countries covered in our empirical analysis, the number of observations and dyads from each country, as well as the start and end date of the period covered.

Table 7: Countries, NSAs and periods covered in empirical analysis

country	NSA	strict sample			strict sample		
		first year	last year	n	first year	last year	n
Burundi	CNDDFDD	2000	2003	4	2000	2003	4
Burundi	Palipehutu				2000	2003	4
India	NDFB	2000	2009	10	2000	2009	10
India	NSCN K	2005	2009	5	2005	2009	5
India	NLFT	2000	2009	10	2000	2009	10
India	NSCN IM	2000	2009	6	2000	2009	10
India	KNF	2000	2009	10	2000	2009	10
India	ULFA	2000	2009	10	2000	2009	10
Indonesia	GAM	2000	2005	6	2000	2005	6
Iran	KDPI	2000	2009	10	2000	2009	10
Iraq	PUK	2000	2007	8	2000	2007	8
Iraq	KDP	2000	2007	8	2000	2007	8
Morocco	POLISARIO	2000	2009	10	2000	2009	10
Myanmar	ABSDF				2000	2008	9
Myanmar	ARIF	2000	2006	7	2000	2009	10
Myanmar	KIO	2000	2009	10	2000	2009	10
Myanmar	KNPP	2000	2009	10	2000	2009	10
Myanmar	KNU	2000	2009	10	2000	2009	10
Myanmar	NMSP	2000	2009	10	2000	2009	10
Myanmar	RSO	2000	2006	7	2000	2009	10
Nepal	CPNM	2000	2006	7	2000	2009	9
Philippines	CPP	2000	2000	1	2000	2000	1
Philippines	MILF	2000	2000	1	2000	2000	1
Philippines	MNLF	2000	2000	1	2000	2000	1
Somalia	SPM				2000	2009	10
Somalia	SRRC	2001	2003	3	2001	2009	9
Sri Lanka (Ceylon)	LTTE	2000	2009	8	2000	2009	8
Sudan	SPLM/A	2000	2003	4	2000	2001	2
Turkey	PKK	2001	2003	3	2001	2003	3
total				179			218

Table 8: Results for Signatory Status of Government with Control for Time-dependency (Polynomials)

sample:	government signs		government signs before NSA		government signs after NSA	
	<i>strict</i> Model 1	<i>lenient</i> Model 2	<i>strict</i> Model 3	<i>lenient</i> Model 4	<i>strict</i> Model 5	<i>lenient</i> Model 6
democracy(lagged)	0.917 (0.739)	1.042 (0.726)	(0.989)	(0.862)	-1.632 (1.753)	-0.763 (1.575)
one-sided violence _G	0.020 (0.029)	0.031 (0.028)	-0.044 (0.040)	-0.016 (0.034)	0.082 (0.059)	(0.054)
one-sided violence _N	-0.013 (0.037)	-0.020 (0.036)	0.019 (0.040)	0.003 (0.039)	-0.013 (0.083)	-0.013 (0.084)
use of mines by government	(0.802)	(0.777)	(0.978)	(0.937)	(1.217)	(1.108)
time	-0.109 (0.235)	-0.084 (0.227)	-0.183 (0.271)	-0.111 (0.253)	0.050 (0.632)	0.089 (0.498)
time ²	0.002 (0.034)	-0.002 (0.033)	-0.014 (0.041)	-0.014 (0.039)	0.065 (0.153)	0.020 (0.101)
time ³	0.000 (0.004)	-0.000 (0.003)	-0.001 (0.005)	-0.002 (0.005)	0.021 (0.031)	0.001 (0.017)
constant	(0.866)	(0.832)	(1.149)	(0.975)	-2.891 (2.125)	-2.538 (2.084)
<i>N</i>	188	226	154	180	42	54
log <i>L</i>	-4.017	-7.680	9.714	5.033	18.596	16.103

Standard errors in parentheses
* indicates significance at $p < 0.1$

Table 9: Results for Signatory Status of NSA with Control for Time-dependency (Polynomials)

sample:	NSA signs		NSA signs before government	
	<i>strict</i>	<i>lenient</i>	<i>strict</i>	<i>lenient</i>
	Model 1	Model 2	Model 3	Model 4
one-sided violence _G	0.014 (0.036)	0.013 (0.036)	0.019 (0.041)	0.019 (0.041)
one-sided violence _N	0.019 (0.037)	0.021 (0.037)	-0.013 (0.043)	-0.013 (0.043)
territorial control \geq moderate	(1.011)	(1.022)	-1.598 (1.160)	-1.610 (1.162)
NSA troop size (logged)	(0.452)	0.712 (0.433)	0.949 (0.638)	0.955 (0.640)
time	0.195 (0.275)	0.184 (0.272)	0.112 (0.281)	0.109 (0.278)
time ²	-0.009 (0.038)	-0.010 (0.038)	-0.010 (0.036)	-0.010 (0.036)
time ³	-0.002 (0.004)	-0.002 (0.004)	-0.002 (0.004)	-0.002 (0.004)
constant	(4.187)	(4.005)	(5.855)	(5.878)
<i>N</i>	153	159	91	93
log <i>L</i>	7.216	7.050	14.343	14.335

Standard errors in parentheses
* indicates significance at $p < 0.1$

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.
- Azam, Jean Paul. 2006. "On thugs and heroes: Why warlords victimize their own civilians." *Economics of Governance* 7:53–57.
- Azam, Jean-Paul and Anke Hoeffler. 2002. "Violence against Civilians in Civil Wars: Looting or Terror?" *Journal of Peace Research* 39(4):461–485.
- Balcells, Laia. 2010. "Rivalry and Revenge. Violence against Civilians in Conventional Civil Wars." *International Studies Quarterly* 54:2.
- Beber, Bernd and Christopher Blattman. 2010. "The Industrial Organization of Rebellion: The Logic of Forced Labor and Child Soldiering." NYU and Yale University.
- Bussmann, Margit and Gerald Schneider. 2011. "A Porous Humanitarian Shield: The Laws of War, the Red Cross, and the Killing of Civilians." Paper presented at the 4th Annual Conference on The Political Economy of International Organizations, Zurich, January 27 - 29, 2011.
- Carey, Sabine C., Mark Gibney and Steven C. Poe. 2010. *The Politics of Human Rights*. New York: Cambridge University Press.
- Carpenter, R. Charli. 2011. "Vetting the Advocacy Agenda: Network Centrality and the Paradox of Weapons Norms." *International Organization* 65:69 – 102.
- Carter, David B. and Curtis S. Signorino. 2010. "Back to the Future: Modeling Time Dependence in Binary Data." *Political Analysis* 18:271–292.
- Cheibub, Jos, Jennifer Gandhi and James Vreeland. 2010. "Democracy and dictatorship revisited." *Public Choice* 143:67–101.
- Clapham, Andrew. 2006. "Human rights obligations of non-state actors in conflict situations." *International Review of the Red Cross* 88:491–523.
- Cunningham, David E., Kristian S. Gleditsch and Idean Salehyan. 2009. "It Takes Two: A Dyadic Analysis of Civil War Duration and Outcome." *Journal of Conflict Resolution* 53(4):570–597.
- Dimaggio, Paul J. and Walter W. Powell. 1991. *The New Institutionalism in Organizational Analysis*. Chicago: University of Chicago Press.
- Downes, Alexander B. 2006. "Desperate Times, Desperate Measures: The Causes of Civilian Victimization in War." *International Security* 30(4):152–195.
- 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.
- Esarey, Justin, Bumba Mukherjee and Will H. Moore. 2008. "Strategic Interaction and Interstate Crises: A Bayesian Quantal Response Estimator for Incomplete Information Games." *Political Analysis* 16(3):250273.
- Finnemore, Martha and Kathryn Sikkink. 1998. "International Norm Dynamics and Political Change." *International Organization* 54(2):891.
- Gelman, Andrew and Jennifer Hill. 2006. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. New York: Cambridge University Press.
- 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 progress 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. 2010. "UCDP Dyadic Dataset Codebook Version 1-2010." Uppsala Conflict Data Program (UCDP), Department of Peace and Conflict Research, Uppsala University.
- Harbom, Lotta, Erik Melander and Peter Wallensteen. 2008. "Dyadic Dimensions of Armed Conflict, 1946-2007." *Journal of Peace Research* 45(5):697–710.
- Harbom, Lotta and Ralph Sundberg. 2009. "UCDP Non-state Actor Dataset Codebook. Version 1-2009." Department of Peace and Conflict Research, Uppsala University.
- Hathaway, Oona A. 2002. "Do Human Rights Treaties Make a Difference?" *Yale Law Journal* 111:1935–2042.
- Hill, Daniel W. Jr. 2010. "Estimating the Effects of Human Rights Treaties on State Behavior." *The Journal of Politics* 72(4):1161–1174.

- 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.
- Humphreys, Macartan and Jeremy Weinstein. 2006. "Handling and Manhandling Civilians in Civil War." *American Political Science Review* 100(3):429–447.
- 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.
- Kalyvas, Stathis. 2006. *The Logic of Violence in Civil War*. Cambridge: Cambridge University Press.
- Keck, Margaret E. and Kathryn Sikkink. 1998. *Activists Beyond Borders: Advocacy Networks in International Politics*. Ithaca and London: Cornell University Press.
- 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."
- Lewis, Jeffrey and Kenneth Schultz. 2003. "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information." *Political Analysis* 11:345367.
- 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.
- Morrow, James. 2007. "When Do States Follow the Laws of War?" *American Political* 101(3):559–572.
- Moser-Puangsuwan, Yeshua. 2008. Outside the Treaty Not the Norm: Non-state Armed Groups and the Landmine Ban. In *Banning Landmines. Disarmament, Citizen Diplomacy, and Human Security*, ed. Jody Williams, Stephe D. Goose and Mary Wareham. Rowman & Littlefield Publishers, Inc. pp. 163–178.
- 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.
- Roodman, David. 2007. "CMP: Stata module to implement conditional (recursive) mixed process estimator." Statistical Software Components, Boston College Department of Economics.
- 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. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13:273–296.
- 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: Theory and Methods." *American Political Science Review* 99(4):623–631.
- Stanton, Jessica. 2010. "Strategies of Restraint in Civil War." Paper prepared for presentation at The Program on Order, Conflict, and Violence Speaker Series, Yale University, November 3, 2010.
- Themnr, Lotta and Peter Wallensteen. 2011 (forthcoming). "Armed Conflict, 1946-2010." *Journal of Peace Research* 48:4.

- Valentino, Benjamin A., Paul Huth and Dylan Balch-Lindsay. 2004. "Draining the Sea: Mass Killing and Guerrilla Warfare." *International Organization* 58(2):375–407.
- Valentino, Benjamin A., Paul Huth and Sarah Croco. 2006. "Covenants without the Sword. International Law and the Protection of Civilians in Times of War." *World Politics* 58:339–77.
- 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.
- Whang, Taehee. 2010. "Empirical Implications of Signaling Models: Estimation of Belief Updating in International Crisis Bargaining." *Political Analysis* 18:381–402.
- Wilde, Joachim. 2000. "Identification of multiple equation probit models with endogenous dummy regressors." *Economics Letters* 69:309312.