International conventions and non-state actors: selection, signaling and reputation effects

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

Kristian Skrede Gleditsch, University of Essex
Simon Hug, Université de Genève
Livia Isabella Schubiger, London School of Economics and Political Science
Julian Wucherpfennig, University College London

Abstract

Whether international humanitarian norms are respected during and after civil conflict depends on the behavior of both governments and non-state actors (NSAs). However, international conventions on the protection of civilians generally do not address NSAs, as such conventions are open only to the representatives of states. In a pioneering initiative, the non-governmental organization Geneva Call has started to address this problem by soliciting NSAs to sign 'deeds of commitment' to ban particular activities violating humanitarian norms. Focusing on the case of anti-personnel mines, we examine why non-state actors would choose to sign conventions that limit their autonomy, and whether such conventions can change the behavior of governments and non-state armed groups. We propose a game-theoretic model of how the interaction between governments and NSAs shape their incentives to commit to and comply with international humanitarian norms. Our empirical evidence highlights the importance of these interdepencies between governments and NSAs in the realm of humanitarian engagements.

word count: approx. 11000

Introduction

Civil wars have been the deadliest type of armed conflict since 1945 and often see pervasive human rights violations. Evidence of deliberate violence against civilians, forced recruitment and child soldiering, inhumane treatment of prisoners, and use of banned weapons abound in human rights reports, legal proceedings, and scholarship on civil wars. Yet, international humanitarian norms intended to protect noncombatants in armed conflicts have only limited reach in civil wars. International humanitarian law is generally not geared towards armed non-state actors, but almost exclusively focused on the representatives of states. Whereas governments are legally bound by their international humanitarian commitments, armed non-state actors[1] (NSAs) are typically precluded from signing the relevant conventions.[2]

This discrepancy has not escaped the attention of the international community. The non-governmental organization (NGO) Geneva Call tries to fill this void by soliciting NSAs to sign specific conventions in which they commit to banning particular activities and consent to monitoring for compliance. The first convention offered is a *deed of commitment* banning the use of landmines, intended to parallel the interstate Ottowa Convention on the Prohibition of the Use, Stockpiling, Production and Transfer of Anti-Personnel Mines and on their Destruction (e.g., Moser-Puangsuwan, 2008).[3] Landmines are used extensively in contemporary civil wars such as the conflicts in Syria and Myanmar. The *2015 Landmine Monitor Report* records a global total of 3,678 direct fatalities caused by mines for 2014 alone (p. 1).[4] In addition to the immediate harm inflicted (much of which remains unrecorded), unremoved landmines threaten to kill or maim people for many years even after armed hostilities have ceased.

Several NSAs have responded to Geneva Call's initiative since its launch in 2000 and signed agreements limiting their use of landmines in much the same way as governments under the Ottawa convention. However, even a cursory look at contemporary

[1] By *non-state actors* (NSAs) we refer to actors currently or previously — engaged in armed conflict, in line with the terminology of Geneva Call (2007). Studies of human rights stressing non-state mobilization — e.g., Finnemore and Sikkink (1998) — often use the term much more broadly, including non-governmental organizations (NGOs).

[2] Note, however, that some authors argue that NSAs are still legally bound by such conventions (see Clapham, 2006).

[3] Geneva Call has meanwhile expanded to other areas, such as child soldiering and sexual violence.

[4] More specifically, this includes victims from landmines, victim-activated improvised explosive devices, cluster munition remnants, and other explosive remnants of war, see `http://www.the-monitor.org/en-gb/reports/2015/landmine-monitor-2015.aspx`.

conflict and post-conflict countries reveals considerable variation in whether governments and non-state actors actually commit to these agreements. In his 2010 inaugural address, Colombian President Santos implored the *Fuerzas Armadas Revolucionarias de Colombia* (FARC) to cease using landmines.[5] Geneva Call had since 2006 unsuccessfully attempted to get FARC to sign the deed of commitment on anti-personnel mines (Geneva Call, 2006), despite the Colombian government having joined the Ottawa landmine treaty in 2000. Geneva Call's activities were more successful in Sudan. In October 2001 the Sudan People's Liberation Movement/Army (SPLM/A) signed the proposed convention (Geneva Call, 2007), and the Sudanese government followed suit two years later in signing up to the Ottawa treaty. By the end of 2014, 162 state parties had signed the Ottawa mine ban treaty, and 50 non-state actors from 10 countries had done the same for Geneva Call's *deed of commitment*.[6] While some non-state actors have formally committed to the convention after their government signed the treaty, others have signed before the government pledged its support.

These actions by governments, non-state actors, and the NGO Geneva Call highlight an important dynamic currently neglected in the literature on humanitarian law in civil wars. Signing decisions by governments as well as their compliance with treaty obligations are typically studied without taking into account domestic non-state armed groups. Likewise, the behavior of NSAs is considered in isolation, neglecting the opportunities offered by NGOs like Geneva Call and the influence of governments.

Taking these overlooked dynamics into account, in this article we address two questions: First, why would state and non-state actors sign a constraining convention that limits their tactical repertoire? Second, what effects do such conventions have on the subsequent behavior of the signing parties? We offer an explanation that highlights the interdependence between NSAs and governments and the role of reputation effects relative to costs of forsaking a specific military technology. Our analysis extends current research on treaties in international relations to non-state actors and to the context of civil wars and their aftermath. We posit that the ongoing debate on the screening and constraining effects of international agreements is relevant beyond states and formal treaties (e.g., Simmons, 1998, 2010; Simmons and Hopkins 2005; von Stein 2005), and that analogous mechanisms can apply to NSAs and informal

[5] "Santos assumes Colombia's presidency amid conciliation with Venezuela, Ecuador" *Los Angeles Times* August 10, 2010.

[6] `http://www.icbl.org/index.php/icbl/Universal/MBT/States-Parties`, `http://www.genevacall.org/`.

agreements or commitment devices as well (see Bangerter, 2011; Jo and Thomson, 2014; Jo, 2015).

Indeed, although Geneva Call's efforts to engage non-state armed groups are quite specific and limited, the data on convention signing and compliance allow us to gain insights that extend beyond this particular application. Our analyses show how decisions by governments and non-state actors are mutually dependent. The decision of an NSAs to ratify Geneva Call's deed of commitment can influence the reputation costs of governments for not signing, and hence induce governments to sign the Ottawa convention in situations where they would not otherwise do so. Similar relationships are likely to be found in other areas where no NGO offers formal deeds of commitments to NSAs, but where NGOs try to influence governments and NSAs through "naming and shaming" (e.g., Hafner-Burton, 2008). Moreover, we find that selection effects are important for evaluating the impact of the Ottawa convention and Geneva Call's deed of commitment on subsequent conflict behavior. Although a naive analysis suggests that countries signing the Ottowa convention are more likely to refrain from using landmines, this effect disappears once we take the signing decisions by NSAs into account (in line with more critical assessments, e.g., Drezner, 2005). Accounting for these selection effects is also crucial when it comes to the behavior of NSAs, for which our findings point to the opposite: We find a restraining effect of Geneva Call's deed of commitment for NSAs, but only after accounting for the endogenous nature of their decision to formally endorse this convention.

Human rights and non-state actors

While humanitarian law originally developed to constrain the behavior of governments in international conflicts, scholars and experts of law started considering ways to legally constrain NSAs in the 1970s (Clapham, 2006; Barbelet, 2008, 82f), as both states and NSAs perpetrate human rights violations in intrastate wars. Human rights obligations for NSAs are still less clearly developed than for states, as NSAs are not signatories to standard international conventions.[7] Not surprisingly, the literature on compliance with international law remains largely state-centric (Simmons, 2010), and only recently have scholars become interested in whether other agents such as individuals or NSAs comply with international law, treaties, or less formal obligations

[7] See Clapham (2006) on the debates surrounding the application of international humanitarian law and human rights to NSAs.

such as norms or conventions (Bangerter, 2011; Jo and Thomson, 2014; Jo, 2015; Morrow, 2007, 2014).[8]

There has been some interest in compliance with the laws of war (Valentino, Huth and Croco, 2006; Morrow, 2007, 2014), as well as wartime civilian targeting (Valentino, Huth and Balch-Lindsay, 2004; Eck and Hultman, 2007; Downes, 2011) and child soldiering (Beber and Blattman, 2013; Lasley and Thyne, 2015). Bussmann and Schneider (2015), for example, analyze how the ratification of international humanitarian law influences access by the International Committee of the Red Cross in civil wars and violence against civilians. Closest to our focus and approach, Jo and Thomson (2014) assess compliance with humanitarian access in armed conflicts and its relationship to reputation and international organizations (see also Stanton, 2009; Bangerter, 2011; Jo, 2015).

All these studies, however, consider violations of humanitarian laws in isolation from the interactions among opponents in civil conflicts. The recent literature on human rights treaties and compliance is more attuned to the interaction between actors: Researchers drawing on a sociological institutionalist perspective see the increasing ratification of human rights treaties as evidence of the growing strength of human right norms (e.g., Finnemore and Sikkink, 1998; Risse, Ropp and Sikkink, 1999). However, other scholars reach more pessimistic conclusions, suggesting that governments often sign human rights conventions without actually enforcing them (e.g., Hafner-Burton, 2008; Hathaway, 2002; Hollyer and Rosendorff, 2011; Simmons, 2009; Vreeland, 2008). Skeptics thus question whether the proliferation of treaties actually entails more than lip-service to norms, and whether signing treaties by itself changes behavior.

The debate about the effects of human rights treaties is part of a more general controversy about the problems in assessing the effects of treaties, since signing a treaty is often influenced by the intended or expected compliance (e.g., Simmons, 1998; Simmons and Hopkins, 2005; von Stein, 2005; Simmons, 2010). That signatories of treaties behave differently may simply reflect the fact that parties sign treaties they are already (or more likely to be) compliant with rather than the effects of treaties themselves. Similar issues are at stake when it comes to the Ottowa convention and the Geneva Call deed of commitment. Some scholars have questioned the importance of the Ottawa treaty, as most signatories did not stock landmines at ratification and the

[8]In the context of the Ottowa convention some researchers emphasize the importance of NGOs and civil society to bring about this mine ban treaty (Price, 1998; Rutherford, 2000), but this body of research has paid less attention to how conventions affect non-state actors in conflict.

enforcement mechanisms remain weak (e.g., Drezner, 2005). Similarly, the SPLM/A in Sudan had already stopped using landmines when it signed Geneva Call's deed of commitment.[9] In short, to assess the causal nature of such conventions, we need to establish plausible counterfactuals for behavior in the absence of the treaty.

The initiative of the NGO Geneva Call provides a unique opportunity to assess the interactions between governments and non-state actors and their influence on each other, as it is the most extensive effort to propose human rights conventions to NSAs to date. Geneva Call was founded in 1998, immediately after the Ottawa convention, in response to the concern that the convention did not prevent armed NSAs from continuing using mines. Geneva Call engages NSAs on landmines through the "Deed of Commitment for Adherence to a Total Ban on Anti-Personnel Mines and for Co-operation in Mine Action," which bans the production, use, and transfer of landmines, and calls for participating in mine clearance and mine risk education. The convention entails verification missions by Geneva Call. Geneva Call is currently engaged in six regions, namely Africa (since 2000), Asia (since 2000), the Caucasus (since 2006), Europe (since 2001), the Middle East (since 2000), and Latin America (since 2003).[10]

Table 1 summarizes the number of countries and NSAs that have signed the Ottowa treaty and the Geneva Call convention up to the final year in our data, 2009. If states and NSAs in conflict signed independently, we should see no systematic pattern to when they sign. Likewise, if NSAs simply mimic states, we should only see them sign in cases where governments have already signed, as the Ottawa convention has been open longer. Table 1 indicates that more states have signed the Ottowa treaty than NSAs signing Geneva Call's deed of commitment. Of course, many states have not been affected by armed conflict, and Geneva Call has so far only been active in a limited number of war-torn countries.[11] However, we also have a substantial number of cases where NSAs sign before states do, or where states have not signed despite NSAs signing.

[9] Personal communication by Pascal Bongard, program officer Geneva Call, January 5, 2011.

[10] See `http://www.genevacall.org/where-we-work/` (accessed September 21, 2015).

[11] Our empirical analysis focuses only on countries where Geneva Call has been active. The fact that Geneva Call tends to focus on NSAs with mine use in these countries implies a selection of hard cases, where actors are more likely to have military advantages from mines and less likely to be compliant at the outset.

Table 1: Ratification of Landmine Ban Convention and Sequence

	Countries	number of NSAs that have signed
Country signed first, NSA afterwards	Burundi	1
	Phillipines	3
	Turkey	1
NSA signed first, country afterwards	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 (77 with internal conflict, 54 since 1997)	
Neither country nor NSA signed	34 countries (21 with internal conflict, 13 since 1997)	

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

A Model

We propose a simple game-theoretic model with incomplete and imperfect information to understand why governments and armed non-state actors would ratify a convention banning the use of landmines—the Ottawa and Geneva Call conventions, respectively—and when ratification will subsequently be adhered to in practice. Formal ratification is not necessarily followed by compliance, especially since such a constraint can undermine an actor's ability to achieve its objectives in the dispute against its opponent. Governments and non-state actors in conflict pursue opposing objectives, which makes their interaction strategic. This induces incentives to misrepresent. Thus, although prior beliefs about compliance exist, both parties are uncertain whether the other party will actually put commitments into action.

Sequence of Play

The game is played between a government G and an armed non-state actor A in an ongoing or unresolved conflict. Accounting for the possibility that signing[12] can be a bluff ("cheap talk") for some actors who will deliberately not comply despite formal ratification, we assume that both actors G and A can be of type "nice" or "mean" (explained in detail below), but their respective types remain private information throughout the game. The extensive form of the game is given in Figure 1:

1. Nature (N) determines G's and A's type (i.e., whether it is in G's and A's interest to comply and refrain from using landmines if they choose to ratify their

[12]For NSAs there is no formal ratification process for conventions, and we use the terms signing and ratification interchangeably here.

respective conventions). For simplicity, this move is omitted from Figure 1.

2. The government G chooses whether (or not) to sign the Ottawa convention.

3. The armed non-state actor A decides whether (or not) to sign the Geneva Call convention.

4. If only A has signed, but not G, then G gets another chance to sign or not sign the Ottawa convention.

5. If A and/or G have ratified the convention, they decide *simultaneously* whether to comply by refraining from using landmines.

This sequence of play mirrors the historical fact that the Ottawa convention preceded Geneva Call's activities. We model compliance as a simultaneous move for several reasons. First, unlike the historical trajectory in which the Ottawa convention predates Geneva Call's deed of commitment, there is no obvious first-mover for compliance. Given the operational mode of landmines, it is difficult for both actors and analyst to precisely date their installation. Second, we only have annual data on mine use, and there is currently no information available to model more detailed sequencing. Finally, compliance is an ongoing and multi-dimensional process that also involves the removal of mines (which is admittedly a strong point of Geneva Call's deed of commitment and a weak point of the Ottowa convention).

Payoff Structure

The full payoffs for the two actors are given in Table 2 and are composed of the following four elements. First, to capture the widely accepted notion that conflict is generally costly (e.g. Fearon, 1995), we highlight the costs of conflict $cw_i > 0$, with $i \in \{G, A\}$. Second, we focus on the reputation of the government. We assume that the public perception of ratification (not compliance) results in three levels of reputation costs (or benefits) linked to three possible sequences of actions, namely G signing the convention first, G refraining from signing, and G signing after A has done so. Without loss of generality we scale these costs (and benefits) such that the intermediary level is assumed to be zero, while the highest corresponds to a benefit and the lowest to a cost. More precisely, we assume reputation benefits $r_G \geq 0$ (if G signs

Figure 1: Extensive Form of the Model

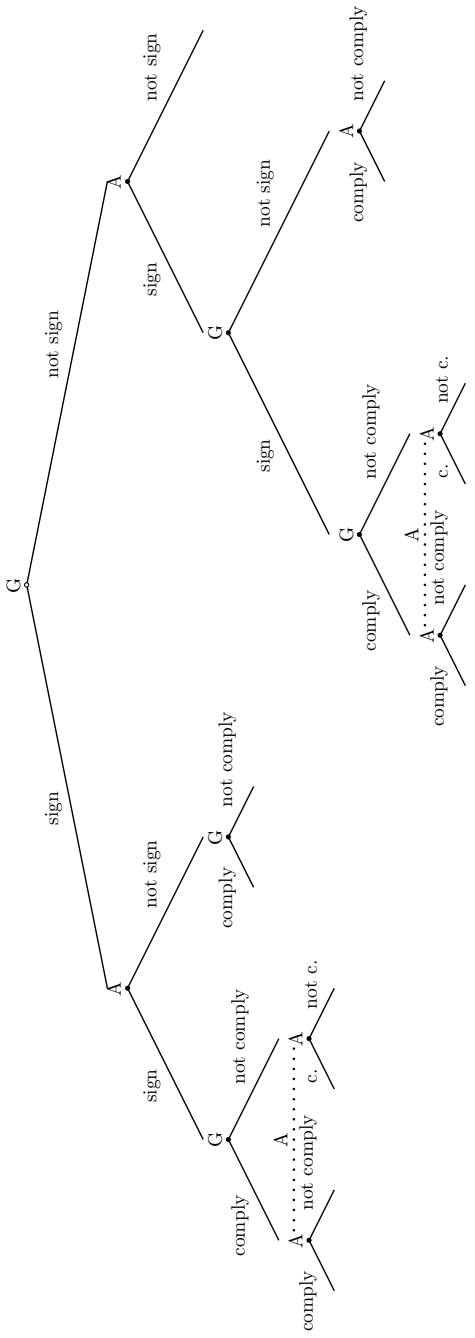

Table 2: Actions and payoffs

Actions		Payoffs	
G	A	G	A
sign, comply	sign, comply	$-cw_G - w_G + w_A + r_G$	$-cw_A - w_A + w_G$
sign, comply	sign, not comply	$-cw_G - w_G + r_G$	$-cw_A + w_G - c_A$
sign, not comply	sign, comply	$-cw_G + w_A + r_G - c_G$	$-cw_A - w_A$
sign, not comply	sign, not comply	$-cw_G + r_G - c_G$	$-cw_A - c_A$
sign, comply	not sign	$-cw_G - w_G + r_G$	$-cw_A + w_G$
sign, not comply	not sign	$-cw_G - c_G + r_G$	$-cw_A$
not sign, sign, comply	sign, comply	$-cw_G - w_G + w_A$	$-cw_A - w_A + w_G$
not sign, sign, comply	sign, not comply	$-cw_G - w_G$	$-cw_A - c_A + w_G$
not sign, sign, not comply	sign, comply	$-cw_G - c_G + w_A$	$-cw_A - w_A$
not sign, sign, not comply	sign, not comply	$-cw_G - c_G$	$-cw_A - c_A$
not sign, not sign	sign, comply	$-cw_G - 2 \times r_G + w_A$	$-cw_A - w_A$
not sign, not sign	sign, not comply	$-cw_G - 2 \times r_G$	$-cw_A - c_A$
not sign	not sign	$-cw_G$	$-cw_A$

first), $r_G = 0$ (if G signs after A), or costs ($-2 \times r_G \leq 0$ if A signs and G does not).[13] This acknowledges an expectation for governments to ratify human rights treaties, thereby signaling genuine commitment to respect international humanitarian norms to domestic and international audiences. Armed NSAs, however, face no comparable expectation to sign human rights conventions, precisely because such conventions have been introduced only very recently, and cover merely a limited number of issues. We hence assume that NSAs do not face reputation costs related to the signing of human rights conventions.[14]

We also model the military costs and benefits, as banning landmines can potentially alter the battlefield capabilities of both actors. Compliance will limit the available military strategies, although the specific impact will depend on the advantages of using mines and differ considerably across conflicts. Thus, we consider the observable military disadvantage, or the increased costs of warfare $w_i \geq 0$ if a treaty is adhered to, with $i \in \{G, A\}$ (by assumed symmetry, these increased costs generate benefits for the adversary).

The model also considers the potential consequences of non-compliance if a treaty is formally ratified but not adhered to. Such costs will be substantial where actors seek legitimacy, either from domestic or international audiences (Jo, 2015). Similar to the signatory-related reputation costs outlined above, these costs also affect reputation,

[13]Thus, the costs of not signing the treaty after A'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.

[14]Note that this assumption does not imply that NSAs do not face reputation costs related to human rights violations; NSAs do indeed face such costs, albeit to varying degrees (see Stanton, 2009; Lasley and Thyne, 2015; Jo, 2015).

albeit for reasons related to non-compliance.[15] We denote the increased costs related to non-compliance by $c_i \geq 0$, with $i \in \{G, A\}$. While this component can be monitored and anticipated, we assume that non-compliance is not fully observable *ex ante*. Specifically, non-compliance by actor i removes the w_i term from both actors' payoffs, while the "offending" actor i pays a cost c_i if non-compliance is detected (possibly stochastically) by monitoring.

Finally, we use the conditions under which G and A will comply after having signed their respective engagements as elements to generate prior beliefs over the two types (i.e., "nice" or "mean") these two actors may be. Jo and Thomson (2014, 327) argue that "humanitarian compliance involves a tradeoff between military considerations and a motivation to comply." Following this logic, we use the relationship between w_i and c_i to characterize the actors' types as either compliant after signatory ("nice") or non-compliant after signatory ("mean"). For the "nice" type we assume that $c_i = \frac{w_i}{2}$, while for the "mean" type $c_i = 2 \times w_i$. This implies that a "nice" type will always be better off complying with a treaty or convention after signing, while the reverse is true for the "mean" type (cf. Jo and Thomson, 2014). Intuitively, this captures the central idea that for both actors compliance depends (in part) on the consequences of getting caught relative to the purely military costs. As we assume that G and A do not know the type of their adversary, the signing decision and the ensuing compliance will depend on their updated beliefs about each other. We denote the prior beliefs that the government and armed non-state actor, respectively, are of type "nice" as $p = Pr(c_G = \frac{w_G}{2})$ and $q = Pr(c_A = \frac{w_A}{2})$; conversely, $1 - p = Pr(c_G = 2w_G)$ and $1 - q$ = $Pr(c_A = 2w_A)$ denote the belief that the other party is "mean".

Equilibrium Analysis

Given incomplete and imperfect information, we use Perfect Bayesian Equilibrium (PBE) as our solution concept.[16] Using the values for c_i assumed for the "nice" and the "mean" types it can be easily seen that a "nice" type, when faced by a compliance decision, will always choose to comply, while the "mean" type in the same situation

[15]Note that in contrast to the reputation costs related to signing, we assume the costs for non-compliance to apply to NSAs as well. Both the Ottawa treaty as well as Geneva Call's deed of commitment foresee measures to make non-compliance publicly visible.

[16]In the appendix we also solve the game under the assumption of complete information to develop our intuitions.

will never do so.[17] Using our simple characterization of "nice" and "mean" types, we solve the game using PBE and state the following proposition:

Proposition 1. *The perfect Bayesian equilibria in pure strategies are:*

$$
\begin{aligned}
separating \quad &= \quad \begin{cases} \textit{mean G signs, nice G} \neg \textit{ signs, A} \neg \textit{ signs} & \textit{if } \frac{w_A}{2} > w_G > r_G > \frac{w_G}{2} \\ & \textit{if } 2w_G > w_A, w_G > r_G > \frac{w_G}{2}, \frac{w_A}{2w_G} > p \end{cases} \\
pooling \quad &= \quad \begin{cases} \textit{G} \neg \textit{ signs, A signs, G signs} & \textit{if } 2r_G > w_G, w_G > w_A, p > \frac{w_A}{w_G}, w_A > r_G, q > \frac{r_G}{w_A} \\ \textit{G} \neg \textit{signs, A} \neg \textit{ signs} & \textit{if } w_G > 2r_G \\ \textit{G signs, A} \neg \textit{ signs} & \textit{else} \end{cases}
\end{aligned}
$$

Proof. See Appendix. □

Proposition 1 shows that in pure strategies only a limited set of outcomes is possible. Several implications are worth highlighting here and inform our subsequent empirical evaluation:

1. If the government's costs of giving up landmines w_G exceed reputation benefits of signing r_G and the prior belief p that G is mean is low, only a "mean"(or noncompliant) G will sign, followed by both types of A not signing.

2. If the prior belief q that A is "mean" is high, then both types of G will first refrain from signing, but will follow suit after A has signed, provided that the military costs w_A are not too high.

3. If the military disadvantage w_G is sufficiently high compared to G's reputation benefits r_G, neither G nor A will ever sign.

4. If the prior belief q that A is "mean" is low, then both types of G will sign immediately.

Observable implications

We now translate the theoretical implications of our formal model into empirically testable hypotheses. Implications 1 and 3 suggest that higher reputation gains compared to the costs of relinquishing landmines for the government increase the likelihood for both actors G and A to sign their respective conventions:

Hypothesis 1. *Governments are more likely to sign a convention if the associated reputation gains outweigh the government's military disadvantage associated with compliance.*

[17] This follows from the systematic comparisons of the four cases listed above.

Hypothesis 2. *NSAs are more likely to sign a convention if the government's reputation gains associated with signing outweigh the government's military disadvantage associated with compliance.*

We assume that governments' domestic reputation costs are higher in democratic countries, and that the costs of refraining from employing landmines can be proxied by prior landmine use. The relationship between regime type and reputation costs with regards to human rights treaties may not always be straightforward (e.g., Hathaway, 2007). Still, we see regime type as good proxy for sensitivity to reputation costs here, since governments will care more about reputation costs when citizens have both the incentives and capacity to make their voices heard.[18] We assume that anti-personnel mines are a subject of particular concern for domestic constituencies, as landmines pose a long-lasting danger to civilians even after armed conflicts end, and that the majority of citizens will not support the use of anti-personnel mines in domestic conflicts even if they have no sympathy for the rebels. Moreover, we assume that democratic governments have more incentives than autocracies to signal responsiveness to these concerns, as reputation costs in democracies can translate into electoral losses. Thus, we expect that governments in democracies should be more likely to sign, especially in cases where landmines are militarily less beneficial. Similarly, prevailing international norms in regards to human rights are likely to impact states' signing decision. To capture international reputation costs we include the cumulative number of countries having signed the Ottowa convention in a particular year, assuming that the number of signatories is indicative of the international diffusion and strength of this particular norm (Finnemore and Sikkink, 1998).[19]

Implication 2 suggests that the likelihood of G signing a treaty decreases as the probability of facing a mean NSA increases. In this case, only if military costs for an NSA of relinquishing landmines are small will governments later sign-on, after NSAs have done so. This allows us to propose the following three hypotheses:

Hypothesis 3. *Governments are less likely to sign a constraining treaty if faced by armed non-state actors that are unlikely to comply with a convention.*

Hypothesis 4. *If the government believes itself to be facing an armed non-state actor that is unlikely to comply with a convention, the government is more likely to sign if*

[18] Simmons (2009), makes a similar assumption in her theory on human rights treaty compliance (see also Conrad and Ritter, 2013).

[19] We thank an anonymous reviewer for this suggestion.

the NSA's value of using landmines is small.

Hypothesis 5. *NSAs are more likely to sign a convention if the costs of compliance with a convention are small.*

We assume the expected cost of non-compliance to be private information. However, actors form their prior beliefs over the antagonist's costs of non-compliance with a convention by observing the use of one-sided violence against civilians. Thus, where a prior history of one-sided violence gives the government reason to believe that the armed non-state actor is "mean" or relatively unconcerned about potential costs of non-compliance, the act of signing by an NSA can convey a credible signal that induces the government to sign as well. By contrast, if the government believes that the armed non-state actor is "nice" as it has not targeted civilians, then it will move to ratify first.

For these three hypotheses we also consider the military benefit of landmines relative to the cost of treaty adherence. If landmines are of reasonable military and tactical value to an armed NSA, late signing by an NSA will be followed by the government signing the convention too. We argue that anti-personnel mines are of greatest military use when they can be used to defend controlled territory, and that larger NSAs will be less dependent on this type of weapon.

Implication 1 also allows us to relate the prior belief p of a "mean" G with the actions of the NSA. More specifically, the latter are much more likely to sign if this prior belief is high, allowing us to state the following hypothesis:

Hypothesis 6. *NSAs are more likely to sign a convention if they believe that the government will not comply after signing.*

The hypotheses stated so far focus on the signing decisions of the two actors. Implication 1, together with our basic assumption on compliance decisions, allow us to formulate the following additional hypotheses on compliance by the two actors:

Hypothesis 7. *Governments with high reputation gains and low costs from relinquishing the use of landmines are more likely to comply with a constraining treaty.*

Hypothesis 8. *Armed non-state actors are more likely to comply if the military value of landmines is small.*

Empirical analysis

Testing our model requires dyadic data, as we analyze the strategic interaction between armed non-state actors and governments. We focus on the regions and time periods where Geneva Call has been active, and include the following countries and years:[20]

- 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],[21] Georgia.
- Europe (2001 onwards): Turkey.
- Latin America (2003 onwards): Colombia.
- Middle East (2000 onwards): Iran, Iraq, Lebanon,[22] Yemen.

The relevant dyads clearly include all governments in countries where Geneva Call has been engaged. However, selecting NSAs is less straightforward. Geneva Call approaches predominantly NSAs that have used landmines in the past.[23] As there is likely to be some error in this assessment, we assume that, in principle, any armed non-state organization is a potential signatory of Geneva Call's deed of commitment. To ensure systematic selection criteria, we only include armed non-state organizations that (1) have been engaged in armed conflict according to the UCDP criteria for at least one year during the period 1989 through 2009 (Harbom, Melander and Wallensteen, 2008),[24] (2) are politically active, and (3) maintain an armed wing.

[20]Geneva Call only provides accurate start and end dates of engagement for a subset of countries, and we use the start year of engagement in regions for all others. See appendix for a full list of NSA-government dyads.

[21]UCDP does not consider Armenia a primary conflict party since it does not face an internal challenge but only supports the separatists in Nagorno-Karabakh against Azerbaijan.

[22]In Lebanon, Geneva Call is mainly in contact with organizations affiliated with Hezbollah, and we also include the Israel-Hezbollah dyad.

[23]Personal communication by Pascal Bongard (program officer Geneva Call, January 5, 2011). We also obtained a list of NSAs contacted by Geneva Call and use this subset of cases in a set of robustness checks of our main findings in the appendix. The NSA selection strategy is likely to bias our results against finding any effects of the proposed convention.

[24]UCDP defines armed conflict as "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

Table 3: Dyad-Years

	NSA active & armed		
Conflict active	Yes	Unclear	Total
No	215	212	427
Yes	239	0	239
Total	454	212	666

Non-state organizations that meet these criteria enter the dataset once they have been actively involved in an intrastate armed conflict, and following Geneva Call's engagement in the region. They are not considered prior to their active involvement in armed conflict, but once active, remain in the dataset up until 2009, when the coverage of our dataset ends, and regardless of whether the respective conflict has reached the 25-fatality threshold every year since.[25] However, consistent with the above-mentioned criteria, we only include NSAs as long as they qualify as politically active organizations that maintain their own armed wing (our coding effort). Table 3 summarizes the dyad-years in our dataset. We rely on both a strict coding (dyad-years in the first column) as well as a more lenient coding that also includes dyad-years with unclear activity in terms of either or both of these two criteria (second column).[26] We report the results based on the lenient coding below, but provide the ones based on the "strict" sample in the appendix.

We consider two outcome variables. The first indicates whether the NSA in a given dyad-year has signed Geneva Call's deed of commitment banning anti-personnel mines, or whether the government has ratified the international mine-ban treaty. The second pertains to actual compliance, based on information from the International Institute for Strategic Studies (IISS) Armed Conflict Database.[27] The variables *Mine use*$_G$ and *Mine use*$_A$ indicate whether the state (G) or non-state actor (A) used landmines and/or improvised explosive devices in given year. Lacking a better measure, we

of a state, results in at least 25 battle-related deaths in one calendar year" (`http://www.pcr.uu.se/research/ucdp/definitions/definition_of_armed_conflict/`). Some commitment signatories — usually small organizations and splinter groups — are not included in the UCDP data (see data appendix).

[25]To illustrate, the *Mouvement des forces démocratiques de Casamance* (MFDC) vs. Senegal dyad is first coded as active in 1990 in the UCDP dyadic dataset (v. 1-2010) and remained active in 2000. We include this dyad for all years after Geneva Call becomes active in the region (2000+), although the 25 battle-related threshold is not reached for every year after 2000.

[26]As can be easily seen, dyad-years that reach the 25 battle-related deaths threshold that defines conflict activity (according to UCDP) are always coded as active and armed opposition organizations.

[27]`http://acd.iiss.org/armedconflict/report/dsp_MainForm.asp`.

use an indicator for all types of mines and other 'improvised explosive devices.' This includes victim-activated devices as well as remote-controlled devices that detonate on demand,[28] but we consider it the best mine use measure currently available.[29]

We now detail our measures of the key elements in the theoretical model. We use a dummy variable *Territorial control* denoting whether the NSA exerts at least a moderate level of control over its main territory, adopted from Cunningham, Gleditsch and Salehyan (2009). We have argued above that this is related to the costs of treaty adherence (w_A). Since landmines are an effective way of securing territory, relinquishing their use will make the NSA more vulnerable. *Use of mines by government* indicates whether landmines and improvised explosive devices were used by states in a dyad after 1997 and prior to Geneva Call's engagement in the respective region.[30] To capture size-related effects, such as military capacity, we use an estimate of the *Troop size* of the NSA, adopted from Cunningham, Gleditsch and Salehyan (2009).

One-sided violence (*OSV*) captures the extent of deliberate and direct violence against civilians attributed to the NSA and the government by the UCDP.[31] More specifically, we use the best estimate of fatalities for all incidents of one-sided violence for a given actor and year.[32] We employ OSV as a proxy for the prior beliefs p and q, assuming that civilian targeting indicates lower costs for being caught in violation of a treaty (i.e., smaller c_i terms).

Democracy is measured using the binary Cheibub, Gandhi and Vreeland (2010) indicator, which we use as a proxy for the sensitivity of governments to domestic reputation costs. We are primarily interested in the difference between systems where

[28]Written communication by Hanna Ucko Neill, IISS Armed Conflict Database coordinator, December 14, 2010.

[29]Geneva 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). We do not use the "Landmine and Cluster Munition Monitor" elaborated by the International Campaign to Ban Landmines (see `http://www.the-monitor.org/` as it appears less systematic than the IISS source.

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

[31]The UCDP One-sided Violence Dataset (Eck and Hultman, 2007) records deadly attacks on civilians by governments and armed groups resulting in at least 25 deaths per calendar year.

[32]Where actors are involved in more than one conflict dyad, or where the reported figures do not distinguish between distinct organizations, we assign the same value to all relevant dyads. For example, the government of Burundi was involved in several conflicts during the period 1989-2009, and "Hutu rebels" encompasses more than one NSA in conflict (Harbom and Sundberg, 2009). In such cases, we assign the OSV fatality estimates to several dyads. Israel is an exception to this rule, as Geneva Call only covers the conflict with Hezbollah. Fatalities caused by militias publicly alleged to have acted on behalf of the state (Autodefensas Unidas in Colombia, Janjaweed in Sudan) are assigned to the state.

governments are chosen through contested elections and systems where this is not the case rather than the degree of democracy. As the original data end in 2008, we extrapolate for 2009. As a proxy for international reputation costs we use the cumulative number of countries having signed and ratified the Ottowa convention.[33]

Finally, we control for whether the conflict is active or not in a given year according to the UCDP 25 battle-related deaths criterion.

Ratifying Mine-Ban Treaties

We first examine the empirical implications of our model for the ratification of the mine-ban treaty, resp. convention, by the governments and NSAs. Table 4 reports the ratification sequence of both governments and NSAs in terms of the number of countries and dyads, as well as the relative percentage of countries and dyads in parentheses (dropping observations following signing). The first sub-table shows the initial decisions in line with the model where the government goes first. The upper left cell in the upper sub-table shows that we have 6 countries (and 54 dyads) where no NSA signs after the government does, and 2 countries and 2 dyads where they do. The lower right cell indicates that we have 8 countries (and 11 dyads) where NSAs sign when the government has not, a much higher relative share. This suggests that NSAs are more likely to sign the Geneva Call convention if the government has not signed the Ottowa convention. The lower part of the table indicates whether governments "follow suit," after observing the behavior of NSA. This shows that 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).

We proceed to a series of logit models to evaluate the theory's implications for the signing decisions of the parties. The strategic nature of these decisions creates statistical estimation problems. While there has been considerable work on estimators for complete information models (see, e.g., Signorino, 1999), only few models have been proposed for incomplete information models like ours (see, e.g., Lewis and Schultz, 2003). Moreover, these solutions tend to be demanding in terms of the underlying data. Therefore, as a more feasible alternative, we estimate separate models for each relevant decision node in figure 1. We employ Gelman and Hill's (2006) Bayesian logit model to address problems of complete separation, given the small number of

[33] Source: `http://www.unog.ch/80256EE600585943/(httpPages)/6E65F97C9D695724C12571C0003D09EF?OpenDocument`, accessed August 14, 2015).

Table 4: Signatories and "Follow-Suit" Signatories

	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)

Note: cell entries report number of countries, number of dyads (% of countries, % of dyads)

	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)

Note: cell entries report number of number of countries (% of countries)

cases.[34]

Table 5: Bayesian Logit Estimates of Signing

	Government			**NSA**	
	Model 1	Model 2	Model 3	Model 4	Model 5
	gov. signs	gov. signs before NSA	gov. signs after NSA	NSA signs	NSA signs before gov.
democracy (lagged)	-0.588	-0.406	-1.013	0.313	-0.517
	(0.674)	(0.716)	(2.277)	(1.072)	(1.287)
no. of ratifying countries	0.003	0.001	0.013	-0.40	-0.038
	(0.023)	(0.023)	(0.118)	(0.032)	(0.041)
use of mines by government	-1.113^{*}	-1.364^{*}	-0.000	(0.754)	0.941
	(0.677)	(0.741)	(2.500)	(0.983)	(1.316)
one-sided violence$_A$ (logged)	0.005	0.011	-0.061		
	(0.024)	(0.025)	(0.084)		
one-sided violence$_G$ (logged)				0.010	0.029
				(0.039)	(0.045)
territorial control $\geq$ moderate	-0.331	-0.421	1.013	-1.595	-0.824
	(0.638)	(0.661)	(2.277)	(1.076)	(1.284)
NSA troop size (logged)	0.507^{*}	0.391	0.365	$0.929*$	0.732
	(0.278)	(0.287)	(0.821)	(0.413)	(0.457)
active conflict	2.777^{*}	2.611^{*}	1.013	0.106	-0.678
	(1.353)	(1.379)	(2.277)	(1.114)	(1.264)
constant	-8.855^{*}	-7.386	-8.624	-6.552	-4.448
	(4.948)	(5.010)	(19.633)	(6.064)	(7.716)
N	217	209	8	318	209
$\log L$	-14.684	-12.395	23.047	3.441	11.415

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

Models 1-3 in Table 5 focus on the government's decision to sign the Ottawa convention. Model 1 looks at whether the government signs without taking into account the stage in the game tree. Hypothesis 1 states that higher reputation costs (democracy, as well as ratification behavior of other governments) and higher costs of relinquishing land mines (use) should increase, respectively decrease, the likelihood of signing the treaty. We only find partial support for this expectation. Use of land mines

[34] We offer a series of robustness checks in the appendix, including replications with standard errors clustered by country and controls for time dependence. These do not affect our substantive results.

significantly decreases the likelihood of signing, but democratic governments are not more likely to sign. In addition, the coefficient for the variable capturing international reputation costs (number of countries having ratified), while positive as expected, also fails to reach statistical significance. Similarly, while hypothesis 3 states that one-sided violence committed by an NSA should make governments less likely to sign the Ottawa treaty, we find a positive, although statistically insignificant, effect. We also find mixed results for hypothesis 4 as both the NSA's territorial control and troop size have the expected negative/positive effect, but only the latter coefficient reaches statistical significance. Finally, during times of active conflicts, governments appear significantly more likely to sign the Ottowa convention.

We then relax the assumption that governments sign independently of what the NSA is doing and focus on the government's decision in two distinct situations, namely when the NSA has not yet signed (model 2), and signing after the NSA has signed (model 3 — given the small number of cases, these results have to be interpreted with caution). The results suggest that these situations differ. Model 2 shows that signing costs (previous use of mines) have a statistically significant effect on the government's decision to sign before an NSA has done so, but much less so when an NSA has already signed. This is in line with the implications from our theoretical model in that high signing costs should reduce the likelihood of a government signing the Ottowa convention. Our proxies for reputation costs (democracy and number of ratifiers) remain insignificant, although the negative coefficient for democracy is larger after the NSA has signed. In addition, it appears that the reputation costs arising at the international level operate in the expected direction when an NSA has already signed, though the coefficient is associated with considerable uncertainty.[35] These results contradict our first hypothesis. When the NSA has not yet signed the convention, the pressure on governments to sign the Ottawa treaty seems especially high in active conflicts.

Contrary to our hypothesis 3, we also find that one-sided violence by the NSA, if anything, increases the likelihood of a government signing the treaty before an NSA. Our theoretical model implies that the government's prior belief of facing a "mean" type, proxied by one-sided violence by the NSA, should decrease the likelihood of

[35] In 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). This is addressed by using a Bayesian logit model, which also permits the estimation with the small number of cases in model 3.

signing at the first decision node. We find a slightly negative effect, however, only if an NSA has already signed.

Finally regarding our hypothesis 4 we find that the NSA's troop size increases the likelihood of a government signing, as we would expect based on our model. Both coefficients, however, fail to reach statistical significance. The same is true for the effect of the NSA's control of territory, for which we only find the expected sign if the government makes a decision to sign before an NSA.

Models 4-5 in Table 5 provide results for the NSAs.[36] Here, too, the implications for NSA's actions involve the prior beliefs and the costs faced by the government and the NSA. Model 4 simply considers whether NSAs sign Geneva Call's deed of commitment. The results provide strong evidence that the costs of relinquishing landmines (troop size, and to a lesser degree territorial control) clearly affect an NSA's decision to sign, which offers support for our fifth hypothesis. Larger NSAs are more likely to sign, while moderate territorial control decreases the likelihood of signing, which is in line with our hypothesis. Again, however, contrary to hypothesis 3, one-sided violence as a proxy for prior beliefs does not seem to affect the NSA's decision, although the coefficient has the expected positive sign. Finally, we find partial support for our second hypothesis, as higher reputation costs for the government (democracy) increase—though statistically not significantly—the likelihood of the NSA signing Geneva Call's convention. The same does, however, not hold for reputation costs linked to the number of ratifiers. Moreover, the costs of relinquishing landmines (proxied by the government's use) does not decrease the probability of an NSA to sign the convention as expected.

Examining the NSA's decision when the government has not yet signed, we find largely similar results, except that the coefficient for democracy becomes negative and the coefficient for an NSA's troop size becomes statistically insignificant.[37] Finally, an NSA's decisions appear not to be affected by whether a conflict is active or not.

[36]There are hardly any cases of NSAs signing after a government. A model for this decision mode generated estimation problems, and we omit it from table 5.

[37]These disappointing results might be linked to the last configuration under which proposition 1 envisions an equilibrium. More precisely, governments are more likely to sign after the NSAs, but it is mostly "mean" NSAs that will sign first. As this configuration is based on knife-edge conditions we refrained from using it as basis for a hypothesis.

How Effective are the Conventions?

We now turn to evaluating the effect of conventions in terms of landmine use by governments and NSAs. Models 1 and 3 in Table 6 report how signing either the Ottawa convention or Geneva Call's deed of commitment affects the respective actors' behavior, disregarding that signatory status is potentially endogenous to its effect.[38] For the government's decision to comply we find that its signatory status significantly reduces the likelihood of landmines use, while that of the NSA has no effect on compliance. Prior use of landmines and an active conflict reduce notably the likelihood of compliance, as does surprisingly also democracy, but not the number of ratifiers. Model 3 in table 6 shows that NSAs that sign do not behave differently from non-signatories in mine use. Interestingly, the signatory status of the government has the reverse effect. The NSA is much more likely to use landmines if the government has signed the Ottawa convention. NSAs with at least moderate territorial control appear to use landmines less frequently, while larger NSAs do so more frequently.

These simple probit results disregard the strategic dependence of the signing decisions by governments and NSAs implied by our theoretical model. We address this important issue in models 2 and 4 of tables 6 by estimating trivariate probit models to consider explicitly the endogenous nature of the government's and the NSA's signatory status, using the previous model specification.[39]

Taking into account the possibly endogenous nature of the signing of a convention or a deed of commitment affects both our estimates for the signing and the compliance

[38]For all analyses in this section we report a full set of robustness checks in the appendix with more restricted samples, excluding all cases from Somalia and restricted to NSAs engaged by Geneva Call. Since we have more than one NSA in some countries we also bootstrap our standard errors stratified by NSA. These robustness checks overwhelmingly confirm our main findings, with deviations arising occasionally when sample sizes become increasingly small for more restrictive definitions.

[39]Our empirical application does not suffer from partial observability, and models designed for this issue are not appropriate (e.g., Heckman-type models). The trivariate probit model allows for correlations among the error terms of three equations, corresponding to the two signing decisions and 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. Estimating simultaneously three interdependent equations for compliance and the signatory statuses of governments and non-state actors allows us to acknowledge and assess the extent of the endogeneity in terms of the magnitude and direction of the correlations between the error terms. Ideally, our signing equations should contain two components: the initial signing decision of the actors and the decision to remain part of the Ottowa/Geneva Call's convention regime. Since no single government nor NSA has left these regimes, the respective coefficients are, however, unidentified. Hence, the estimates for the signing decision conflate the effects of both the signing decision and maintaining one's status under the two conventions. We control, however, for the lagged signatory status in the signing equations.

Table 6: Results for Compliance/Mine Use

	Government				**NSA**			
	Model 1	Model 2			Model 3	Model 4		
	probit	trivariate probit			probit	trivariate probit		
DV:	G use	G use	G sig.	A sig.	A use	A use	G sig.	A sig.
government signed	-0.647^{***}	-0.547			0.507^{***}	0.804^{**}		
Ottawa convention	(0.230)	(0.440)			(0.138)	(0.343)		
NSA signed Geneva Call's	0.283	-0.185			0.042	-1.015^{**}		
deed of commitment	(0.327)	(0.436)			(0.305)	(0.466)		
government signed				1.633^{***}				1.696^{***}
Ottawa conv. (lagged)				(0.364)				(0.394)
NSA signed Geneva Call's			1.914^{***}				1.771^{***}	
deed of cmtnt. (lagged)			(0.340)				(0.411)	
democracy (lagged)	0.755^{***}	0.792^{***}	0.183	0.864^{*}			0.144	0.316
	(0.213)	(0.213)	(0.174)	(0.494)			(0.226)	(0.445)
no. ratifying countries	-0.038^{***}	-0.039^{***}	0.029^{***}	-0.007			0.030^{***}	-0.008
	(0.010)	(0.011)	(0.008)	(0.013)			(0.009)	(0.013)
use of mines by government	2.516^{***}	2.574^{***}	-1.011^{***}	1.442^{***}			-1.185^{***}	1.189^{***}
	(0.273)	(0.294)	(0.178)	(0.364)			(0.195)	(0.374)
territorial control $\geq$ moderate			0.356^{*}	-1.243^{***}	-0.319^{*}	-0.413^{**}	0.170	-1.090^{***}
			(0.193)	(0.372)	(0.185)	(0.180)	(0.206)	(0.376)
NSA troop size (logged)			0.112	0.956^{***}	0.135^{*}	$0.135*$	0.065	0.239^{***}
			(0.082)	(0.224)	(0.081)	(0.088)	(0.093)	(0.203)
one-sided violence$_G$ (logged)				0.017				0.020
				(0.014)				(0.012)
one-sided violence$_A$ (logged)			-0.010				-0.112	
			(0.008)				(0.009)	
active conflict	0.300	0.284	0.164	-1.000^{***}	0.645^{***}	0.464^{**}	0.639^{***}	-1.125^{***}
	(0.229)	(0.231)	(0.201)	(0.376)	(0.13)	(0.196)	(0.240)	(0.341)
Constant	$3.500**$	3.499^{**}	-5.570^{***}	-9.426^{***}	$-1.399**$	-1.799^{***}	-6.407^{***}	-6.696^{**}
	(1.445)	(1.479)	(1.535)	(2.911)	(0.636)	(0.661)	(1.679)	(2.633)
$atan(\rho_{12})$			-0.011				-0.185	
			(0.222)				(0.234)	
$atan(\rho_{13})$			0.423				1.008^{*}	
			(0.263)				(0.540)	
$atan(\rho_{23})$			-0.528^{**}				-0.435^{*}	
			(0.235)				(0.250)	
N	310		310		266		266	
$\log L$	-97.480		-304.0		-166.212		-339.0	

Standard errors in parentheses
$^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.1$

equation.[40] For the government's decision to sign we still find the strong negative effect of previous mine-use on signing the Ottawa convention (in support of our first hypothesis), while for the variables reflecting reputation costs we obtain positive coefficients, of which only the one for the number of ratifiers is significant. We still find no effect of prior beliefs (one-sided violence, hypothesis 3). The NSA's decision is now, as expected, affected both by territorial control, troop strength, one-sided violence by the government (though statistically insignificant) and democracy. More specifically, NSAs with at least moderate control over territory are much less likely to sign Geneva

[40]The difference is due to the considerable correlation in the error terms of the two signing equations. Apparently, omitted variables affect similarly the decisions of both governments and NSAs to sign or refrain from doing so. However, in the trivariate probit model we only implicitly take into account the sequence of signing decisions.

Call's deed of commitment. Larger NSAs, however, are much more likely to submit to this regime, while NSAs in democracies and those faced by "mean" governments, are more likely to sign.

Taking these endogenous signing decisions into account we find divergent effects of the treaty, resp. convention, on the use of mines. State actors not using mines are disproportionately likely to join the Ottawa convention, and failing to take this into account can be problematic in assessing the effects of the convention. Taking this into account suggests that the treaty by itself has no notable effect on the use of mines by governments (the insignificant, though negative coefficient also supports the pessimistic assessment, e.g., by Drezner, 2005). Surprisingly, we also find that democracy, and thus domestic reputation costs, increase the likelihood of a government using landmines. While this contradicts our hypothesis 7, the positive and significant coefficient for prior use of mines by the government as well as the negative and significant coefficient for the number of ratifiers lend support to this same hypothesis. We also find that NSAs signing the deed of commitment decreases (though not significantly) the probability of landmines being used by governments.

The compliance model for the NSAs also suggests that taking signing decisions into account is necessary. When doing so, we find that the government's signing of the Ottowa convention increases the likelihood that NSAs use landmines, while the deed of commitment by Geneva Call reduces it significantly. This latter finding points to the effectiveness of the deed of commitment, a result that only appears once selection issues are taken into account. In support of our hypothesis 8, the estimated coefficients for NSA characteristics, namely whether it is large and enjoys at least moderate levels of territorial control, remain statistically significant with the expected signs.

Discussion and Conclusion

We have proposed a simple game-theoretical model of the decisions of NSAs and state parties on whether to commit to a specific international humanitarian norm. In line with the theoretical model, our empirical analysis provides evidence that the decisions by governments and NSAs to sign a ban of anti-personnel mines are clearly interdependent. Governments' decisions to sign appear to be notably affected by whether NSAs have already signed. For NSAs, as expected, territorial control is an important factor influencing the costs of implementation and thus also the their

decisions to sign.

Our results are also instructive regarding the consequences of signing such conventions. When neglecting the endogenous nature of the decision to sign, we find that the Ottowa convention significantly reduces governments' use of landmines. However, this effect disappears when we take the singing decision into account. Accounting for the endogenous nature of actors' signing decision is equally important for assessing the effectiveness of the deed of commitment for NSAs, where our findings point in the opposite direction: Based on a naive analysis that neglects NSAs decision to sign, we fail to find a significant effect. When incorporating the signing decisions of both the treaty and the deed of commitment, however, we find a strong effect on NSA compliance. Thus, Geneva Call's deed of commitment seems to increase the likelihood that NSAs will remove landmines from their arsenal. Possible explanations for this contrasting finding include that signing a convention is a rare opportunity to gain formal (international) recognition for NSAs, or that NSAs are particularly vulnerable to monitoring.

While our finding regarding NSA compliance is very encouraging, other ways to achieve these goals are also possible. For example, whereas Geneva Call unsuccessfully tried to engage the FARC in its initiative, as outlined in the introduction, this NSA has nevertheless started to remove landmines in 2015 – even without a signed deed of commitment, and thanks to the ongoing peace process.[41] Future research should not only consider the logic of strategic interaction between governments and NSAs, but also examine the actors' internal dynamics and decision processes that motivate compliance with international humanitarian norms, as well as the sources of NSAs capacity to enforce compliance within their ranks.

To be sure, not all our theoretical expectations found empirical support. For instance, the expected compliance by NSAs does not appear to influence a government's decision to sign first. Moreover, we argued that governments' vulnerability to reputation costs, both domestic and international ones, plays a considerable role in their decisions to sign. While we found some evidence for the effect of international reputation concerns, our proxy for the vulnerability to domestic reputation costs, democracy, did not turn out to play the significant role that we expected. This result resonates with the contradicting findings on the effects of democracy in previous studies on human rights treaties and their effects.

[41] See `http://www.bbc.co.uk/news/world-latin-america-32941200`.

In summary, our results show that the actions of governments and NSAs in the realm of humanitarian law should be considered interdependent not only when it comes to the behavior of these actors in conflict, but also regarding their decisions to subject themselves to humanitarian law. These results speak to the broader literature on human rights conventions and the effectiveness of these treaties. Most importantly, our analysis clearly suggests that assessing the effectiveness of humanitarian conventions needs to take into consideration the selection effects linked to the conscious decisions of actors committing to these initiatives. The results also suggest that NSAs influence the costs and decision making of governments. In the empirical realm that we covered, where an NGO actively intervenes to influence NSAs, this is to be expected. We surmise that it is very likely, however, that NSAs and NGOs influence these costs also in other circumstances through more subtle and indirect means.

Appendix

In this appendix we provide proofs of the propositions presented in the main text.

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.[42]

1. Observation

   If G signs at its first decision node, A 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:

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

$$
\begin{aligned}
EU_A(sign) &= p' \times (-cw_A - q \times w_A + w_G - (1-q) \times c_A) \\
&= +(1-p') \times (-cw_A - q \times w_A + (1-q) \times c_A) \\
&= -cw_A - q \times w_A - (1-q)c_A + p' \times w_G \qquad (1) \\
EU_A(not\ sign) &= p' \times (-cw_A + w_G) + (1-p') \times (-cw_A) \\
&= -cw_A + p' \times w_G \qquad (2)
\end{aligned}
$$

As w_A and c_A are both positive, independent of q A will never sign. $QED.$

2. Observation

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

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

$$
\begin{aligned}
EU_G(sign) &= q' \times (-cw_G - w_G + w_A) + (1-q') \times (-cw_G - w_G) \\
&= -cw_G - w_G + q' \times w_A \qquad (3) \\
EU_G(not\ sign) &= q' \times (-cw_G - 2 \times r_G + w_A) + (1-q') \times (-cw_G - 2 \times r_G) \\
&= -cw_G - 2 \times r_G + q' \times w_A \qquad (4)
\end{aligned}
$$

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

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

$$
\begin{aligned}
EU_G(sign) &= q' \times (-cw_G + w_A - c_G) + (1-q') \times (-cw_G - c_G) \\
&= -cw_G - c_G + q' \times w_A \qquad (5) \\
EU_G(not\ sign) &= q' \times (-cw_G - 2 \times r_G + w_A) + (1-q') \times (-cw_G - 2 \times r_G) \\
&= -cw_G - 2 \times r_G + q' \times w_A \qquad (6)
\end{aligned}
$$

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

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

3. Observation

From observation 2 follows that if $2 \times r_G > w_G$ then independent of its type G will always sign at its second decision node.[43] 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 A is identical, it can anticipate its payoff, namely if $2 \times r_G > w_G$ and $q = 1$

$$\begin{aligned} EU_A(sign) &= p' \times (-cw_A + w_G - w_A) + (1 - p') \times (-cw_A - w_A) \\ &= -cw_A - w_A + p' \times w_G \quad (7) \\ EU_A(not\ sign) &= -cw_A \quad (8) \end{aligned}$$

Consequently a "nice" A signs in that case if $p' > \frac{w_A}{w_G}$. For a "mean" A the payoffs are as follows:

$$\begin{aligned} EU_A(sign) &= p' \times (-cw_A + w_G - c_A) + (1 - p') \times (-cw_A - c_A) \\ &= -cw_A - c_A + p \times w_G \quad (9) \\ EU_A(not\ sign) &= -cw_A \quad (10) \end{aligned}$$

Consequently a "mean" A signs in that case if $p' > \frac{c_A}{w_G}$

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

$$\begin{aligned} EU_A(sign) &= -cw_A - w_A \quad (11) \\ EU_A(not\ sign) &= -cw_A \quad (12) \end{aligned}$$

As $w_A > 0$ A will never sign. For a "mean" type

$$\begin{aligned} EU_A(sign) &= -cw_A - c_A \quad (13) \\ EU_A(not\ sign) &= -cw_A \quad (14) \end{aligned}$$

As $c_A > 0$ A will never sign.

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

Complete and imperfect information

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

Proof of Proposition 2 [44]

Based on the observations above the following subgame-perfect equilibrium can be established:[45]

1. If $p = 1, q = 1$ and $w_G > 2 \times r_G, w_G > w_A, r_G > w_A$ $G : \{sign, sign, comply\}$, $R : \{not\ sign, sign, comply\}$[46]

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

3. If $p = 1$, $q = 1$ and $w_G > 2 \times r_G, w_G < w_A$ $G : \{not\ sign, signcomply\}$, $R : \{not\ sign, not\ sign, .\}$

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

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

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 $c_G > 2 \times r_G, r_G > c_G$ $G : \{sign, not\ sign, not\ comply\}$, $R : \{not\ sign, not\ sign, comply\}$

[44] We refrain from presenting the equilibria of the complete and perfect information game as these are closely related to the ones presented here.

[45] For simplicity we omit cases where either of the two actors is indifferent between his or her two actions.

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

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 $2 \times r_G > w_G, c_A > w_G, r_G > w_G$ $G : \{sign, sign, comply\}$, $R : \{not\ sign, not\ sign, notcomply\}$

12. If $p = 1, q = 0$ and $2 \times r_G > w_G, c_A > w_G, w_G > r_G$ $G : \{not\ sign, sign, comply\}$, $R : \{not\ sign, not\ sign, notcomply\}$

13. If $p = 1, q = 0$ and $2 \times r_G > w_G, w_G > c_A, r_G > w_G$ $G : \{sign, sign, comply\}$, $R : \{not\ sign, sign, notcomply\}$

14. If $p = 1, q = 0$ and $2 \times r_G > w_G, w_G > c_A, w_G > r_G$ $G : \{not\ sign, sign, comply\}$, $R : \{not\ sign, sign, notcomply\}$

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 2 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. We solve the game by backward induction as, we will show below, either no private information is transmitted, or G's action reveal all information. We recall also, that whether G and A comply with their treaty/convention obligation is solely determined by their type.[47] Finally, we do not

[47]We note that a semi-pooling equilibrium exists in mixed strategies in which the mean G probabilistically signs the treaty, while the nice one does not, while the nice A after G's decision not sign will also refrain from signing, but the mean A signs probabilistically.

consider situations where G or A is indifferent between several courses of actions and thus also only consider pure strategy equilibria. Consequently, we will start by analyzing A's last left-most decision node before compliance decisions are reached. A compares the following expected utilities (with p' indicating the possibly updated prior belief):

$$\begin{aligned} EU_A(sign|q=1) &= p'(-cw_A - w_A + w_G) + (1-p')(-cw_A - w_A) \\ EU_A(sign|q=0) &= p'(-cw_A - c_A + w_G) + (1-p')(-cw_A - c_A) \\ EU_A(\neg sign|q=.) &= p'(-cw_A + w_G) + (1-p')(-cw_A) \end{aligned}$$

For both types of A it follows that they will never sign after G's decision at the first decision-node to sign the treaty.

Moving to G's right-most last decision-node before the compliance decisions, G will consider the following expected utilities:

$$\begin{aligned} EU_G(sign|p=1) &= q'(-cw_G - w_G + w_A) + (1-q')(-cw_G - w_G) \\ EU_G(sign|p=0) &= q'(-cw_G - c_G + w_A) + (1-q')(-cw_G - c_G) \\ EU_G(\neg sign|p=.) &= q'(-cw_G - 2r_G + w_A) + (1-q')(-cw_G - 2r_G) \end{aligned}$$

Consequently, the nice G will sign after As decision to sign if $q'w_A - cw_G - w_G > q'w_A - cw_G - 2r_G$ from which it follows that $2r_G > w_G$ has to hold. For the mean G to sign $q'w_A - cw_G - c_G > q'w_A - cw_G - 2r_G$ has to hold, implying that $4r_G > w_G$ (as $c_G = \frac{w_G}{2}$ for the mean type). From this it follows that three situations have to be distinguished: if $w_G > 4r_G$ no G will ever sign at its second decision node on the right of the game, if $4r_G > w_G > 2r_G$ then only the mean G will sign, while if $2r_G > w_G$ then both Gs will sign.

Using this knowledge we consider A's first decision node in the right part of the game and assume first that $w_G > 4r_G$. Consequently, A's expected utilities are the following:

$$\begin{aligned} EU_A(sign|q=.) &= -cw_A - w_A \\ EU_A(\neg sign|q=.) &= -cw_A \end{aligned}$$

It follows trivially that neither type of A will ever sign. Consequently, G considers at its first decision node the following expected utilites:

$$\begin{aligned} EU_G(sign|p=1) &= -cw_G - w_G + r_G \\ EU_G(sign|p=0) &= -cw_G - c_G + r_G \\ EU_G(\neg sign|p=.) &= -cw_G \end{aligned} \tag{15}$$

It follows directly that a nice G will sign if $r_G > w_G$ while a mean G will sign as long as $2r_G > w_G$. So under the condition $w_G > 4r_G$ three equilibria candidates exist, namely that both Gs sign at the first decision node when $r_G > w_G$, which contradicts our assumption $w_G > 4r_G$, that only the mean G signs at the first decision node if $w_G > r_G > \frac{w_G}{2}$, which again contradicts our assumption $w_G > 4r_G$, while if $\frac{w_G}{2} > r_G$ neither Gs nor As will ever sign. As this latter condition is not in contradiction with our assumption $w_G > 4r_G$ this establishes the first equilibrium of a pooling type.

Considering next the situation where $4r_G > w_G > 2r_G$, which implies that only a mean G will sign after A's signing, the following expected utilities for A are relevant

$$\begin{aligned} EU_A(sign|q=1) &= p'(-cw_A - w_A) + (1-p')(-cw_A - w_A) \\ EU_A(sign|q=0) &= p'(-cw_A - c_A) + (1-p')(-cw_A - c_A) \\ EU_A(\neg sign|q=.) &= -cw_A \end{aligned}$$

From this it trivially follows that neither type of A will sign after Gs decision not to sign. This implies that G considers at its first decision node the same expected utilities as those in equation 15. But as above, the conditions for signing on r_G and w_G are in contradiction with the assumption $4r_G > w_G > 2r_G$, implying that both Gs will refrain from signing, establishing a second pooling equilibrium.

The last conditions to explore is when $2r_G > w_G$, which implies that both types of G will sign if A has done so before them. Consequently, A considers the following expected utilities after G's decision not to sign at its first decision node:

$$\begin{aligned} EU_A(sign|q=1) &= p'(-cw_A - w_A + w_G) + (1-p')(-cw_A - w_A) \\ EU_A(sign|q=0) &= p'(-cw_A - c_A + w_G) + (1-p')(-cw_A - c_A) \\ EU_A(\neg sign|q=.) &= -cw_A \end{aligned}$$

Consequently, a nice A will sign if $p'w_G - cw_A - w_A > -cw_A$ implying that $p' > \frac{w_A}{w_G}$, while for the mean A the condition implies $p' > \frac{w_A}{2w_G}$. As these lower bounds for the updated belief p' can not exceed 1 for both Gs to sign, it is easy to establish that if $w_A > 2w_G$ neither of A's type will sign. This implies the same expected utilities for G as in equations 15, and consequently the nice G will sign if $r_G > w_G$ while the mean G will sign if $2r_G > w_G$. Consequently, if $r_G > w_G$ holds in equilibrium both Gs will sign, while A is not signing. If $2r_G > w_G > r_G$ then the mean G will sign and the nice G will not, and in both cases A will refrain from signing, which establishes an additional pooling equilibria and a separating one. If $2w_G > w_A > w_G$ and $\frac{w_A}{2w_G} > p'$ holds the same applies establishing another pooling equilibrium.

If $2w_G > w_A > w_G$ and $p' > \frac{w_A}{2w_G}$ then the mean A will sign after G's refusal to sign, while the nice A will not do so. Consequently G's expected utility is as follows:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G \\
EU_G(sign|p=0) &= -cw_G - c_G + r_G \\
EU_G(\neg sign|p=1) &= q'(-cw_G) + (1-q')(-cw_G - w_G) \\
EU_G(\neg sign|p=0) &= q'(-cw_G) + (1-q')(-cw_G - c_G)
\end{aligned}
$$

Consequently a nice G will sign at its first decision node if $-cw_G - w_G + r_G > -cw_G - w_G + q'w_G$ which implies that $\frac{r_G}{w_G} > q'$. For the mean G $-cw_G - c_G + r_G > -cw_G - c_G + q'c_G$ has to hold for a decision to sign, which implies that $\frac{2r_G}{w_G} > q'$. As we explore the case where $2r_G > w_G$, this condition holds independent of the value of q and q'. Consequently, $r_G > w_G$ then both Gs will sign at the first decision node, and A will not sign, establishing another pooling equilibrium. If $2r_G > w_G > r_G$ and $\frac{r_G}{w_G} > q'$ then again both types of G sign at the first decision node and A refrains from doing so. If $2r_G > w_G > r_G$ and $q' > \frac{r_G}{w_G}$ holds the mean G will sign, while the nice one will refrain, implying that $p' = 0$, which contradicts our initial assumptions so no pooling equilibrium exists in this case.

Next, still under the assumption that $2r_G > w_G$ we consider the case where $w_G > w_A$ and $p' > \frac{w_A}{w_G}$ leading both As to sign. In that case G's expected utilities are as follows:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G \\
EU_G(sign|p=0) &= -cw_G - c_G + r_G \\
EU_G(\neg sign|p=1) &= q(-cw_G - w_G + w_A) + (1-q)(-cw_G - w_G) \\
EU_G(\neg sign|p=0) &= q(-cw_G - c_G + w_A) + (1-q)(-cw_G - c_G)
\end{aligned}
$$

Consequently, a nice G will sign if $-cw_G - w_G + r_G > qw_A - cw_G - w_G$ implying that $\frac{r_G}{w_A} > q'$ and the same condition can also be derived for the mean G. Consequently, $\frac{r_G}{w_A} > q$ holds then both Gs will sign and A will not sign, while if $q > \frac{r_G}{w_A}$ holds then both Gs will not sign, but both As will sign, which induces both Gs also to sign, which establishes another pooling equilibrium.

Next consider the case where $2r_G > w_G$, $w_G > w_A$ and $\frac{w_A}{w_G} > p' > \frac{w_A}{2w_G}$ holds. As in this cases only the mean A signs after G's refusal to sign, the following expected utilites are relevant for G:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G \\
EU_G(sign|p=0) &= -cw_G - c_G + r_G \\
EU_G(\neg sign|p=1) &= q(-cw_G) + (1-q)(-cw_G - w_G) \\
EU_G(\neg sign|p=0) &= q(-cw_G) + (1-q)(-cw_G - c_G)
\end{aligned}
$$

Consequently the nice G will sign if $-cw_G - w_G + r_G > -cw_G - w_G + qw_G$ implying that $\frac{r_G}{w_G} > q$. For the mean G the condition is $-cw_G - c_G + r_G > -cw_G - c_G + qc_G$ implying that $\frac{2r_G}{w_G} > q'$. If $r_G > w_G$ applies then trivially both conditions are satisfied establishing a pooling equilibrium where both G will sign at the first decision node, followed by A's refusal to sign. If $w_G > r_G$ and $\frac{r_G}{w_G} > q$ still holds then another pooling equilibrium of the same type appears. If under $2r_G > w_G > r_G$ and $q > \frac{r_G}{w_G}$ then only the nice G will not sign, leading A to update its belief to $p' = 0$ making both types to refrain from signing. Consequently this establishes another separating equilibrium.

Finally, the last case to consider is if $w_G > w_A$ and $\frac{w_A}{2w_G} > p'$ holds. As in that case no A will sign after G's refusal to sign at the first decision node, the following expected utilities for G become relevant:

$$
\begin{aligned}
EU_G(sign|p=1) &= -cw_G - w_G + r_G \\
EU_G(sign|p=0) &= -cw_G - c_G + r_G \\
EU_G(\neg sign|p=.) &= -cw_G
\end{aligned}
$$

It follows trivially that if $r_G > w_G$ both Gs sign at the first decision node and A refuses to do so. If $w_G > r_G > \frac{w_G}{2}$ holds, we obtain a candidate separating equilibrium in which the mean G signs at the first decision node while the nice one refrains from doing so, which leads A to update its belief to $p' = 1$. But this value exceeds the upper limit of $\frac{w_A}{2w_G} > p'$ so that this equilibrium cannot exist. Finally, if we consider $\frac{w_G}{2} > r_G$, this latter expression contradicts our assumption that $2r_G > w_G$, so that no equilibrium can exist under these conditions.

As this exhausts all possible configuration, this proves our proposition. $QED.$

References

Bangerter, Olivier. 2011. "Reasons Why Armed Groups Choose to respect International Humanitarian Law or not." *International Review of the Red Cross* 93(882):353–384.

Barbelet, Veronique. 2008. "Engaging with armed non-state actors on humanitarian issues : a step towards peace?" Submitted for examination for PhD, University of York, Department of Politics.

Beber, Bernd and Christopher Blattman. 2013. "The logic of child soldiering and coercion." *International Organization* 67(01):65–104.

Bussmann, Margit and Gerald Schneider. 2015. "A Porous Humanitarian Shield: The Laws of War, the Red Cross, and the Killing of Civilians." *Review of International Organizations* p. in press.

Cheibub, José, Jennifer Gandhi and James Vreeland. 2010. "Democracy and Dictatorship Revisited." *Public Choice* 143(1-2):67–101.

Clapham, Andrew. 2006. "Human Rights Obligations of Non-State Actors in Conflict situations." *International Review of the Red Cross* 88(863):491–523.

Conrad, Courtenay R. and Emily Hencken Ritter. 2013. "Treaties, tenure, and torture: The conflicting domestic effects of international law." *The Journal of Politics* 75(02):397–409.

Cunningham, David E., Kristian Skrede 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.

Downes, Alexander B. 2011. *Targeting civilians in war*. Ithaca: Cornell University Press.

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

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

Fearon, James D. 1995. "Rationalist Explanations for War." *International Organization* 49(3):379–414.

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 Non-State Actors in a Landmine Ban: The Geneva Call Progress Report (2000-2007)*.

Hafner-Burton, Emilie M. 2008. "Sticks and Stones: Naming and Shaming and the Human Rights Enforcement Problem." *International Organization* 62(4):689–716.

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.

Hathaway, Oona A. 2007. "Why Do Countries Commit to Human Rights Treaties?" *Journal of Conflict Resolution* 51(4):588–621.

Hollyer, James R. and Peter Rosendorff. 2011. "Domestic Politics and the Accession of Authoritarian Regimes to Human Rights Treaties." *Quarterly Journal of Political Science* 6(3-4):275–327.

Jo, Hyeran. 2015. *Compliant Rebels. Rebel Groups and International Law in World Politics*. New York: Cambridge University Press.

Jo, Hyeran and Catarina Thomson. 2014. "Legitimacy and Compliance with International Law: Access to Detainees in Civil Conflicts 1991-2006." *British Journal of Political Science* 44(2):323–355.

Lasley, Trace and Clayton Thyne. 2015. "Secession, legitimacy and the use of child soldiers." *Conflict Management and Peace Science* 32(3):289–308.

Lewis, Jeffrey and Kenneth Schultz. 2003. "Revealing Preferences: Empirical Estimation of a Crisis Bargaining Game with Incomplete Information." *Political Analysis* 11(4):345–367.

Morrow, James D. 2007. "When Do States Follow the Laws of War?" *American Political Science Review* 101(3):559–572.

Morrow, James D. 2014. *Order within anarchy: The laws of war as an international institution*. New York: Cambridge University Press.

Moser-Puangsuwan, Yeshua. 2008. Outside the Treaty Not the Norm: Nonstate 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.

Rutherford, Kenneth R. 2000. "The Evolving Arms Control Agenda: Implications of the Role of NGOs in Banning Antipersonnel Landmines." *World Politics* 53(1):74–114.

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

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. 2010. "Treaty Compliance and Violation." *Annual Review of Political Science* 13:273–296.

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. 2009. "Strategies of Violence and Restraint in Civil War." Columbia University.

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(3):339–77.

von Stein, Jana. 2005. "Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance." *American Political Science Review* 99(4):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.

Wilde, Joachim. 2000. "Identification of Multiple Equation Probit Models with Endogenous Dummy Regressors." *Economics Letters* 69:309–312.