Do Self-Reporting Regimes Matter?
Evidence from the Convention against Torture

Cosette D. Creamer & Beth A. Simmons*

June 2015

Draft. Not for citation.

*Cosette D. Creamer is a Visiting Assistant Professor at Boston University School of Law and a Ph.D. Candidate in the Department of Government, Harvard University. She holds a J.D. from Harvard Law School. Beth A. Simmons is the Clarence Dillon Professor of International Affairs in the Department of Government, Harvard University. All inquiries should be directed to: creamer@fas.harvard.edu. For helpful feedback, the authors would like to thank Christopher Fariss, Courtney Hillebrecht, Katerina Linos, Yonatan Lupu, Gerald Neuman, Tonya Lee Putnam, Kathryn Sikkink, Anton Strezhnev, Geir Ulfstein, Erik Voeten, and participants in the WCFIA-HLS International Law-International Relations Workshop, the Conference on the Domestic Politics of International Human Rights Agreements held at the Niehaus Center for Globalization and Governance, Princeton University, and the Human Rights and Constitutionalism thematic working group at the University of Oslo, Faculty of Law and Norwegian Centre for Human Rights. The authors also thank Diana Li and Andrea Ortiz for providing invaluable research assistance, and the Weatherhead Center for International Affairs for providing generous funding to support this research.
Abstract

Self-reporting on implementation is common in international regulatory agreements and yet we know almost nothing about how (or whether) it works. We argue self-reporting provides information for international and domestic audiences, with the potential to create pressure for agreement compliance. Using original data on the quality and responsiveness of reports submitted to the Committee Against Torture, we test for the causal effects of the periodic review process on the pervasiveness of torture in the reporting country. Adopting a dynamic approach to strengthen causal inference, we find that the review process in fact does reduce the incidence of torture in self-reporting states. Moreover, local media attention in reporting states spikes during the review process, consistent with a domestic mobilization mechanism. This is the first study to demonstrate the effects of self-reporting on torture outcomes. Since many international agreements are based on self-reporting, the results have broad significance for international relations.

Keywords: self-reporting; international agreements; human rights; treaty implementation
Introduction

A large scholarly literature seeks to explain why states would voluntarily participate in legal regimes designed to monitor treatment of a government’s own citizens (Hathaway 2003; Simmons 2009). It is evident that formal participation in the international human rights regime is expanding and not limited merely to governments we might expect to be fully committed to human rights protections, such as stable and transitioning democracies or countries with strong domestic traditions of respecting human rights (Conrad 2014; Hafner-Burton 2012: 267-68; Vreeland 2008). Despite increased participation, the human rights regime continues to suffer from an ‘enforcement problem,’ making its effect on improving practices on the ground questionable. Some scholars argue that enforcement by the United Nations (UN) and human rights treaty bodies has been an utter failure (Posner 2014). But as many have noted, enforcement mechanisms in the human rights context are likely indirect and work through non-governmental actors, transnational organizations, information mechanisms, and domestic institutions (Dai 2014; Dancy and Sikkink 2012).

The primary goal of this article is to unpack and expand our understanding of the intermediary effects of treaty ratification. What does ratification do? What happens, at the domestic and international levels, as a result of treaty ratification that might subsequently impact practices? In particular, we focus on the fact that treaty ratification initiates an iterative and ongoing “constructive dialogue” between a polity and the international human rights regime about state reports on implementation submitted to the treaty monitoring body. Not only is reporting and periodic review a legal obligation within all human rights conventions, it is also the primary international method to generate information and increase transparency about implementation of and compliance with the treaty.

Claims about the importance of information and monitoring in facilitating compliance with international agreements are made on both sides of the spectrum. In the context of human rights treaties, some claim that information—about compliance in particular—is the human rights regime’s primary tool to ensure that states fulfill their obligations (Dai 2007).
Others disparage the entire process as a bureaucratic exercise with little to no substantive effect on compliance and further characterize the record of state reporting to treaty bodies as shamefully inadequate (Hafner-Burton 2013). Yet we still know relatively little about this process or its potential influence on states’ actual human rights practices. Case studies are interesting, but to date they have been few in number, limited in geographic scope, and generally inconclusive about the connection between reporting and rights outcomes (McQuigg 2011). This article examines the effects of reporting and periodic review within the treaty regime governing torture, with the goal of shedding light on whether and how self-reporting influences human rights outcomes.

Why might we expect self-reporting to influence rights outcomes? Reporting to treaty monitoring bodies initiates a dialogue with international “experts” and as such can contribute to socialization of domestic elites and bureaucrats responsible for human rights practices. It also provides information for domestic audiences who have a stake in their government’s implementation of international agreements. Finally, self-reporting may set in motion bureaucratic routines to gather, authenticate, and analyze information that might not have occurred in the absence of the obligation to report. It is even possible that reporting helps to develops an autonomous capacity to self-monitor and self-enforce.

While self-reporting may provide an opening for constructive engagement with the treaty monitoring body, the review process can only be expected to work well if governments take it seriously. For this reason, in evaluating the effect of self-reporting under the Convention against Torture (CAT), we draw on original data on the quality and responsiveness of reports submitted to the regime’s treaty monitoring body: the Committee against Torture (CmAT). A focus on report quality permits us to evaluate the level of states’ engagement with the periodic review process, by analyzing reporting not simply as a procedural obligation but as an opportunity for government officials and domestic audiences to learn about and become socialized into the international human rights regime (Keohane, Macedo and Moravcsik 2009).
This article proceeds as follows. The following section provides a brief summary of the periodic review process and the claims made about this system’s (un)importance. We then theorize the mechanisms through which self-reporting and periodic review could influence a government’s human rights practices, before presenting evidence on the direct effect of the history of a country’s engagement with the treaty monitoring body on torture practices. Finally, we explore one potential mechanism through which reporting could influence outcomes—via domestic political activation and mobilization—and evaluate the extent to which evidence supporting this mechanism exists within Latin American countries.

Reporting to the International Human Rights Regime

The legal regime for international human rights was designed to provide accountability through the monitoring of governments’ implementation of treaty obligations, as one potential pathway towards improved rights practices. To this end, every major human rights convention establishes an oversight committee, comprised of independent experts nominated and elected by states parties. By virtue of treaty ratification, states must submit to each committee periodic reports on the legislative, judicial, administrative, or other measures adopted to give effect to their human rights obligations.

Each treaty monitoring body then considers these reports in the presence of government representatives, through a “constructive dialogue” during which it engages representatives, acknowledges progress made, and identifies areas for improvement. At the conclusion of this dialogue, the committee issues a set of concluding observations containing non-binding recommendations for legislative reforms and other efforts a government should undertake to address shortcomings in its treaty obligations (O’Flaherty 2006: 36). This entire process is known as “periodic review,” with all state reports and committee recommendations made public.

---

1 The Torture Convention requires states parties to submit an initial report within one year of ratification or accession, and subsequent periodic reports at least every four years. See Article 19(1), CAT.
Periodic review was intended to play a central role in encouraging treaty implementation and compliance. As Keller and Ulfstein note, “[t]he main responsibility for the international monitoring of national implementation is . . . entrusted to the UN human rights treaty bodies” (Keller and Ulfstein 2012: 2). Moreover, since it is a mandatory obligation, Kälin refers to the examination of state reports as “the key mechanism established at the universal level to monitor the implementation of treaty obligations by contracting states” (Kälin 2012: 16). But the system is often criticized as inadequate, ineffective, and even “in crisis” (Alston and Crawford 2000; Bayefsky 2001). Some point to the professional inadequacies of the “expert” committees (Hafner-Burton 2013: 102). Others note that states—even resource rich, democratic ones—don’t do what they are told to do by the experts (McQuigg 2011).

Moreover, there is a growing sense among critics that the system as a whole is breaking under its own unwieldy weight (Hafner-Burton 2013: 99; Posner 2014). As the body of treaties has grown, so too have the treaty bodies to which states are expected to report. One result may be reporting fatigue, and it is common to point out that late and non-reporting is fairly widespread (Hampson 2007; Schöpp-Schilling 2007). For example, of the 147 states parties to the CAT in June 2011, thirty (20.4%) had still not submitted their initial report (with Somalia’s initial report the latest at twenty years) and 122 (83%) had between one and five periodic reports overdue.

It is also debatable whether governments take the report drafting process seriously, with reports submitted to the CmAT varying considerably across countries and over time in their structure and quality (Creamer and Simmons 2015). Quality reporting requires an institutional capacity to provide factual knowledge of, expertise in, and familiarity with the treaty regime and the reporting process, which some states have developed through their

---

2Many governments must also provide information on implementation of the prohibition against torture in the context of reporting under the International Covenant on Civil and Political Rights. Given this partly overlapping procedural obligation, states may simply prioritize reporting to the Human Rights Committee over the specialized purview of the CmAT.

3The number of parties with overdue reports includes those instances where the Committee has indicated in its concluding observations on the prior report that a revised date of submission (usually in the near future) is permitted. This represents one way in which the Committee has attempted to address both systematic late reporting and its own increased workload, deviating from the periodicity mandated in the Convention.
National Human Rights Institutions (NHRIs), independent governmental bodies specifically mandated to promote human rights. In fact, the existence of an NHRI significantly increases the probability that a government will submit its CAT report and that the report will be more responsive to CmAT recommendations (Creamer and Simmons 2015).

Domestic political characteristics and regional imitation also partly explain (non) submission, with new democratic governments or those that have recently undergone a democratizing transition less likely to report as compared to countries that have not experienced a transition. However, when newly democratizing countries do report, the quality of their reports is exceptionally high (Creamer and Simmons 2015). Reporting density within a state party’s region substantially increases a government’s probability of reporting, suggesting that as more neighboring states engage with the CAT regime, the expectation that governments should take their reporting obligations seriously increases.

Understanding why states report, while valuable, provides no indication of whether the reporting regime fulfills its stated purpose—facilitating improved human rights practices and increased compliance (Alston 1997: 20). Critics assert that those most affected by treaty violations—and thus most likely to exert pressure on delinquent or non-compliant states—are rarely aware of the periodic review process, conducted in Geneva far removed from domestic media and non-governmental organization (NGO) attention. Hafner-Burton seems to reflect an informal (and untested) consensus among commentators: that “the reports often don’t seem to lead to results that matter” (Hafner-Burton 2013: 100). Others claim that the influence of the reporting and review process is positive, albeit diffuse and indirect, with NGOs, domestic actors, and other governments using the committees’ concluding observations to pressure governments.
Periodic Review and Substantive Compliance

Critics of the periodic review process are right about a number of its shortcomings. States neglect to turn in their reports on time, if at all; many are not self-critical; the oversight committees are swamped and render only unenforceable recommendations. However, there are three major mechanisms through which self-reporting could improve human rights practices: bureaucratic self-assessment and capacity development; elite socialization; and domestic political mobilization.

Internal self-assessment and capacity development

Self-reporting requirements prompt a government to collect and share information about human rights legislation, policies, and practices. Even if states are less than forthcoming within their reports, the domestic process of preparing a report could itself promote self-assessment (Kalin 2012: 39; Trindade 2000: 334). Ideally, all administrative bodies responsible for implementing a given treaty are involved in report preparation. In order to adequately fulfill their reporting duties, governments often “reorganize themselves in ways that may enhance the influence of individuals and bureaucratic units that are more sympathetic to external views,” providing further impetus for examination of the status quo (Keohane, Macedo and Moravcsik 2009: 18). Compiling a report requires a government to engage in a comprehensive review of national legislation and administrative practices, which may reveal previously unnoticed gaps. Because reporting duties encourage the state party to monitor its human rights situation on an ongoing basis, they facilitate internal capacities to collect and analyze information, raising the probability of uncovering deficiencies and finding ways to adopt best practices. To the extent that reporting bolsters the capacity to self-monitor, it may in turn lead to self-enforcement.
Socialization

Self-reporting can be a key transmission belt for elite socialization, especially when it is a step in a process of dialog with the international community. The multilateral nature of the treaty regime and the transnational network it engenders enable discussion and sharing of ‘best practices’ to address common implementation and compliance problems (Keohane, Macedo and Moravcsik 2009; O’Flaherty 2006). The treaty body draws on collective experiences to make recommendations that expand the range of options and information available to governments and publics (Keohane, Macedo and Moravcsik 2009: 18). The periodic review process thus might contribute to improved practices through a mechanism of learning and problem solving, both by government officials and committee members. Through interaction with the treaty body, states get advice about the technical aspects of implementation, various policy options, techniques, or other legal and technical details not previously known or considered. As governments report more frequently and render higher quality and more detailed reports, they engage in a greater number of learning opportunities to improve their policies and practices. As Chayes and Chayes (1993: 303) argue, such processes are central to eliciting compliance and effectiveness (see also Joachim, Reinalda and Verbeek 2008: 11; Mitchell 1998: 113).

Treaty bodies also engage in a practice of “naming and shaming” by publicizing non-compliance and recommending proposals for reform. This negatively affects a state’s reputation for “good behavior” internationally (Guzman 2002; Keohane 1997) or increases social pressures to comply (Goodman and Jinks 2004, 2013; Risse and Ropp 1999). Reporting also enables acculturation, as “[t]he very process of identifying, describing, and controlling human rights practices helps the diffusion of the human rights discourse through global and local levels” (Goodman and Jinks 2004: 697). However, it is important to recognize that reporting may simply be part of the ‘script of modernity’ expected of states (Wotipka and Ramirez 2008). If reporting reflects simple mimicry of ‘appropriate’ behavior, it might not lead to any real normative or substantive change in human rights practices. Nonetheless,
governments that mimic reporting ‘scripts’ likely are susceptible to broader acculturation pressures for implementation and compliance. Even if reporting is initially *pro forma*, it exposes governments to identified best practices for treaty implementation and compliance, which might improve human rights practices on the ground.

Finally, genuine persuasion is a real possibility. Many regime participants believe that authoritative bodies such as the CmAT have impartial knowledge and specialized expertise, which may give them normative power ([Barnett and Finnemore 1999](#)) exercised through reasoned arguments to persuade elites that that they should comply with their human rights obligations ([Checkel 2005](#) [Risse 54](#)). At the very least, reporting generates an ongoing dialogue between treaty bodies and government representatives about the *meaning* of compliance ([Chayes and Chayes 1995](#)). Governments that engage more frequently, thoroughly, and candidly are likely to experience repeated persuasion attempts by the committee. On average, the frequency and intensity of persuasion attempts experienced by states should lead to improved human rights practices over time, as governments internalize the committees’ normative arguments. In particular, if a persuasion mechanism is at work we should see reports that are more responsive to committee recommendations and that actively engage in with committee concerns over time.

**Political Activation and Domestic Mobilization**

Many theories of compliance with international law rely on the availability of information about government activities and legal obligations to domestic publics. Arguably, a public commitment to comply with a human rights treaty raises domestic groups’ expectations that they can demand such compliance ([Simmons 2009](#)). Xinyuan [Dai 2007](#) argues that information produced by international bodies informs domestic audiences about the activities of their governments and whether a government has complied with its international legal obligations. This information allows domestic constituencies to apply electoral pressure on their government in order to hold them accountable.
If a government submits high quality reports about compliance shortcomings and efforts to remedy these, this information raises the expectations of domestic groups that they can legitimately demand these measures be put into practice. The very event of reporting serves to stimulate attention, discussion, and perhaps even participation in the process of report drafting. By mobilizing and empowering groups within and outside of government, reporting can have a catalytic effect in promoting internal policy reform. The committee’s concluding recommendations provide domestic constituencies with information needed to apply electoral and other forms of political pressure to encourage substantive compliance. Even when states are less than forthright, their reports provide a focal point for non-state actors to assess and criticize the information provided. A formal report submission presents opposition parties, NGOs, and other rights constituencies with a visible occasion and target for mobilization. In this way, even incomplete or inadequately analyzed information is better than none at all.

Evidence: Reporting, Review, and Torture

This section examines how the process of periodic review and the history of a government’s engagement with the CAT regime affect subsequent rights practices. Do states that regularly and promptly submit higher quality reports subsequently engage in less torture or cruel and inhumane treatment than states that fail to submit or submit lower quality reports? Does the degree of government responsiveness to prior Committee recommendations influence subsequent torture practices? Do states that ‘improve’ in their reporting (in terms of increased timeliness or quality of reports) demonstrate better compliance with the Convention Against Torture?

Self-reporting and periodic review is an ongoing and iterative process with potentially cumulative effects. It was never intended or designed to affect rights practices through a single report submission. Modeling the effects of a ‘single-shot treatment’ is thus not the best approach to analyzing the dynamic nature of periodic review to the CmAT, since both
the act of reporting and the evolving nature of the review process represent treatment variables of interest. For this reason, we are interested in the effect of a country’s reporting history on its human rights practices. However, estimating such treatment history effects requires a different approach from conventional regression estimators for single-shot effects. This is because time-varying variables can be both pre-treatment confounders in one period and post-treatment variables in another period. Including them in a conventional regression model induces post-treatment bias while failing to include them results in omitted variable bias (Blackwell 2013). To properly adjust for these confounders, we fit a Marginal Structural Model (MSM) developed by Robins, Hernan and Brumback (2000) and recently introduced to political science by Blackwell (2013). These models are estimated using an Inverse Probability of Treatment Weighting (IPTW) estimator, which permits adjustment for intermediate confounders while avoiding bias from conditioning on a post-treatment variable.

This method entails two steps. First, we estimate a model for treatment in each time period, conditional on time-varying confounders and past reporting history. Because report submission is a dichotomous decision, we estimate the probability of reporting with a logit model and the parameter vector for the model with a pooled logistic regression, with country-year as the unit of analysis. Using this model, we estimate the probability, in each period, that the unit received the treatment history that it did. In the second stage, we fit a regression model for the outcome given treatment and treatment history. We weight each observation by the inverse of its treatment probabilities calculated in step one. Robins, Hernan and Brumback (2000) show that this weighting scheme is sufficient to adjust for confounding by observed and time-varying confounders. Intuitively, weighting creates a hypothetical superpopulation where the link between treatment and confounders is broken. The parameters of this weighted regression model can therefore by interpreted causally, under a no omitted variable assumption.

Within the first stage weighting models, we include all covariates (potential confounders) we expect to influence both the decision to report and torture practices. Creamer and
Simmons (2015) report that institutional capacity (existence of an NHRI), an Article 22 declaration accepting the jurisdiction of the CmAT to receive individual complaints, and reporting density within a state party’s region substantially increased a government’s probability of reporting. These variables all likely affect a government’s torture practices as well, so we include them and a number of additional time-varying covariates in the weighting model: logged GDP per capita and population (as collected by the World Bank); a country’s torture score (as measured by the Cingranelli-Richards (CIRI) Human Rights Database); and its Polity IV score. Finally, we include four binary variables intended to capture relevant domestic political characteristics: whether, since World War II or post-war independence, a country never scored an 8 or above on the Polity scale (never democratic); whether, since World War II or post-war independence, a country had never scored below an 8 on the Polity scale (stable democracy); whether the country is undergoing or underwent a democratic transition (moving from below an 8 to an 8 or above on the Polity scale); and whether a country had undergone a democratizing transition (+3 or more on the Polity scale) during either of the two previous years. All time-varying covariates are lagged one year.

In addition to these time-varying confounders, we include within the weighting models a set of treatment (self-reporting) history variables. This permits us to estimate within the second-stage outcome models the average treatment effect of both the single-shot treatment (reporting) and reporting history on a government’s human rights practices. For the decision to report, the treatment history variables include: whether a country reported in the previous year; the number of reports previously submitted; and the number of years since the last report due date. All weighting models include year fixed-effects.

Next, we estimate the effects of report submission (the treatment) and the reporting

---

4 This variable likely strongly predicts the non-submission of a report, as countries are only required to submit a report every four years. Although rare, however, there have been a few instances where a country submitted a report two years in a row, usually to catch up on delayed reports.

5 Following Blackwell (2013) and Cole and Hernan (2008), we conduct a preliminary model check based on the final distributions of the stabilized weights for each year. See Online Appendix Figure 1. The stabilized weights’ means at each point in time are all close to 1, with their upper bounds relatively low (observations for 2011 largely drop out of the outcome model analyses), indicating we have estimated a set of fairly well-behaved weights.
process (treatment history variables) on human rights practices with an ordered probit model that includes the IPT weights. Because we have few priors about when we should expect the reporting process to affect torture practices on the ground, we estimate these effects on a government’s torture score one and two years prior to the year of observation (to assess anticipatory or internal self-assessment effects of reporting), the treatment year, as well as one, two, three and four years following report submission (as significant reforms of legislation and practices identified as deficient often cannot occur instantaneously). All models include a linear (year) time trend.

Across all model specifications, the mere act of report submission has no significant effect on torture practices (see Online Appendix Table 1). However, as Figure 1 demonstrates, delayed reporting has a significant negative effect on the probability of observing lower levels of torture. As delay in report submission increases from 1 to 4 years, and all other variables (reporting and reporting history) are held constant at their means, the probability of observing frequent torture within a country increases by a little over seven percent. This implies the reverse as well. Prompt submission decreases the probability of observing higher levels of torture, suggesting that delinquent reporting—and long stretches between report submission—pose a real obstacle to the periodic review process having a positive effect on the ground. It also suggests that the frequent and iterative nature of the reporting process, rather than the mere act of submitting a single report, has the long-term potential to lead to improvements in rights practices.

It is not too much of a stretch to view the real treatment as the CmAT’s review and the set of recommendations issued at its conclusion, rather than a state’s report drafting and submission alone. The length of time between report submission and review has varied considerably over the years, ranging from twenty-six days to a little over three years and averaging 1.23 years. During the first two decades of the CAT’s existence, average delay between reporting and review slowly increased as the Committee’s workload grew, although within recent years it has sought to address this by consolidating reporting requirements
for governments with good compliance records. While the process of compiling the report could prompt internal self-assessment (though there is little evidence of that affecting torture practices within the first set of MSMS discussed above), until a government engages in dialogue with the CmAT about the ways in which it is falling short of treaty obligations, few if any external pressures to reform existing laws or practices are generated.

To evaluate whether this is the case, we re-estimate a set of weighting models with CmAT review as the treatment of interest. These models include the same time-varying confounders as the reporting-treatment weighting models, but replace report-specific with review-specific treatment history variables: whether the country had engaged in the in-person CmAT review in the previous year; the number of CmAT reviews received by country; and the number of

---

6In order to expedite review, the CmAT also began sending lists of issues to governments following report submission, to which the government responds in writing prior to in-person review.
years since the last review.

We estimate the effects of CmAT review and review history with the same set of year leads and lags described above. All models include a linear (year) time trend. As with reporting, the single-shot CmAT review has no significant effect on torture practices across all model specifications (see Online Appendix Table 2). Yet, as Figure 2 demonstrates, the number of times a country previously engaged in CmAT review significantly increases the probability that it will engage in less torture. Each additional review is estimated to decrease the probability of observing frequent torture within a country by around four percent. This suggests that it is the continuing constructive dialogue with the Committee and not merely the number of reports submitted that has the potential to positively affect rights practices on the ground.

To summarize, governments that promptly submit reports and that engage in more dialogue with the CmAT are more likely to subsequently engage in less torture over time. The very fact of reporting may provide an opening for constructive engagement with the treaty monitoring body. But do these results hinge on whether governments take reporting seriously? To capture report quality, we coded every submitted report along four dimensions: implementation, compliance, responsiveness, and inclusion of data. Treaty bodies have requested that state reports include information on changes in law and administrative procedures (Implementation) and concrete practices on the ground relevant to treaty obligations (Compliance). Most have expressly requested statistical information on outcomes relevant to treaty obligations (Data). Within our coding scheme, a report’s Quality represents the extent to which a government provided information and was transparent and forthcoming about shortcomings along these three dimensions. “Constructive dialogue” implies responsiveness; therefore every subsequent periodic report was also coded for how responsive a country was

---

7See Online Appendix Figure 2 for a preliminary model check based on final distributions of stabilized weights for each year.

8These findings largely hold for weighting and outcome models that additionally include report quality and responsiveness as treatment history covariates. See Online Appendix Figures 3 and 4 for preliminary model checks based on final distributions of stabilized weights for each year, including report quality history and report responsiveness history as treatment-history variables.
Figure 2: Effect of Additional Constructive Dialogues.
Simulated estimates of effect of number of previous CmAT reviews on probability of observing a given level of torture (CIRI Torture Score, two-years after observation-year). The circles represent estimates of the expected effect on the probability of observing a given level of torture as the number of previous reviews changes from 0 to 1, 1 to 2, and 2 to 3, and all other variables are held constant at their means. The lines are 90% confidence intervals. The circles and lines are solid when there is at least 90% confidence of a positive or negative effect. Otherwise, circles are open and lines are dotted. See Online Appendix for standard regression tables. IPTW balanced on treatment of CmAT review; model includes linear time trend and bootstrapped parameters.

to the Committee’s concluding observations on their previous report (Responsiveness)\(^9\).

When controlling for the average quality and responsiveness of past reports, the finding that delayed submission increases the probability of observing higher torture scores holds, with the substantive effect remaining largely the same (see Online Appendix Table 3). On the other hand, additional CmAT reviews no longer significantly and positively affect rights practices (see Online Appendix Table 4)\(^10\). Including report quality and responsiveness history variables within the weighting models also permits us to say something about their causal effect on torture practices. When balanced on year of report submission, a gov-

\(^9\)Online Appendix Figure 5 provides an overview of the instrument used to code each report to the CmAT along these four dimensions. For a more detailed discussion of the coding of report quality and responsiveness, see authors’ website.

\(^10\)When the linear time trend (years CAT regime in existence) is replaced by a country-varying time trend that measures the years a country has been a member of the CAT regime, however, the strong positive effect of additional reviews holds.
ernment’s past responsiveness scores significantly and positively affect its rights practices (Figure 3). Every additional one-unit improvement in past responsiveness scores decreases the probability of observing frequent torture within a country by around four percent. In surprising contrast, past quality scores significantly and negatively affect a government’s rights practices (Figure 3). A two-unit improvement in past report quality scores increases the probability of observing frequent torture within a country by around five percent.

When balanced on year of CmAT review, a government’s average quality scores (including for the report under review) have no discernible effect on rights practices (Figure 4 and Online Appendix Table 4). However, average responsiveness scores (including for the report under review) continue to significantly and positively affect a government’s rights practices (Figure 4 and Online Appendix Table 4). Every additional one-unit improvement in a government’s average responsiveness score decreases the probability of observing frequent torture within a country by close to ten percent. This provides strong support for the conclusion that it is the constructive dialogue with the CmAT—as an ongoing process—which holds the potential to improve rights practices on the ground. This is particularly true when a government engages with and responds to the CmAT’s concluding observations on its prior reports, thereby increasing the quality of the dialogues engendered by the periodic review process.
Figure 3: Effect of Past Report Quality and Responsiveness.
Simulated estimates of effect of average past report quality and responsiveness scores on the probability of observing a given level of torture (as measured by the CIRI Torture scale, two-years after observation-year). The circles in the first row represent estimates of the expected effect on the probability of observing a given level of torture, as the average quality score of past reports changes from 0 to 2, 2 to 4, and 4 to 6, with all other treatment and treatment history variables held constant at their means. The circles in the second row represent estimates of the expected effect on the probability of observing a given level of torture, as the average responsiveness score of past reports changes from 0 to 1, 1 to 2, and 2 to 3, with all other treatment and treatment history variables held constant at their means. The lines are 90% confidence intervals. The circles and lines are solid when there is at least 90% confidence of a positive or negative effect. Otherwise, circles are open and lines are dotted. See Appendix for standard regression tables. IPTW balanced on treatment of reporting; model includes linear time trend and bootstrapped parameters.
**Figure 4: Effect of Report Quality and Responsiveness.**
Simulated estimates of effect of average report quality and responsiveness on the probability of observing a given level of torture (as measured by the CIRI Torture scale, two-years after observation-year). The circles in the first row represent estimates of the expected effect on the probability of observing a given level of torture, as the average quality score of past reports (including the report under review) changes from 0 to 2, 2 to 4, and 4 to 6, with all other treatment and treatment history variables held constant at their means. The circles in the second row represent estimates of the expected effect on the probability of observing a given level of torture, as the average responsiveness score of past reports (including the report under review) changes from 0 to 1, 1 to 2, and 2 to 3, with all other treatment and treatment history variables held constant at their means. The lines are 90% confidence intervals. The circles and lines are solid when there is at least 90% confidence of a positive or negative effect. Otherwise, circles are open and lines are dotted. See Appendix for standard regression tables. IPTW balanced on treatment of CmAT review; model includes linear time trend and bootstrapped parameters.
Prompt and frequent engagement within the CAT regime significantly improves governments’ torture practices, but what mechanisms might account for the strong and significant findings of the previous section?

One potential pathway is through *capacity enhancement*. This involves developing the *internal capacity* to collect and analyze information by the state bureaucracy itself, as a first step toward self-enforcement. If this were true, it might not even be necessary for the oversight committee to respond; reporting alone could lead to more information, enhanced transparency, internal discussions about improvements, and eventually self-implementation of torture bans and safeguards. While a possibility, as we have seen the evidence for such a direct reporting effect is not strong. There is little measurable improvement in torture practices associated with reporting alone. But there is evidence of improvement (with a lag) associated with CmAT review and issuance of recommendations. This is suggestive of a mechanism that highlights *the dialogue*—not simply reporting, but justification, review, shadow reporting, recommendation making, and the provision of further information where it is wanting—as stimuli to policy change.

Another possibility is a *public mobilization* mechanism. That is, the process of dialogue initiated through self-reporting may mobilize domestic demands for closer scrutiny of torture laws and practices. One type of evidence for domestic political activation and mobilization would be the spread of discussion about the CAT periodic review process to uninformed publics. For this mechanism to be at work, we would expect the process to be publicly *visible* (Källin 2012: 41). If elite dialogues at the international level work because they become broadly politically relevant at the local level, we should observe their footprint in public discussions and debate.

To evaluate whether this is the case, we turn to an analysis of local media. We focus on Latin America, a region in which torture has historically been a serious issue and where reporting to the CmAT has varied across countries and over time (see Online Appendix Table...
5). Combined with its history of relatively democratic institutions, active civil society, and meaningful press freedom, Latin America is a plausible candidate to investigate the potential for the periodic review process to acquire some level of publicity and focus public attention on torture practices.

We are fully aware that press coverage is not the same thing as political mobilization. Due to space constraints, it is not possible to establish every causal step in the mobilization chain: the effect of media on civil society demands, the expression of these demands, their influence on policies, and policy influences on actual torture practices. Here, we want to establish the plausibility of the publicity/mobilization channel, which depends at least initially on publicly available information that self-reporting has taken place and the issues on which it focused.

For these reasons, we searched the major local press outlets for sixteen Latin American countries for awareness and discussion of the reporting and review process. Using five specific combinations of search terms in Spanish and Portuguese, we collected all articles mentioning the Convention Against Torture, the CmAT, and/or the periodic review process. We only collected and coded articles published by a newspaper within and about the reporting state. We did not collect media stories for states prior to their ratification of the CAT, but we did search for them regardless of whether a state had in fact met its reporting obligation(s).

Each article was then coded for whether it mentioned the focal state’s CAT obligations, its relations with the CmAT, and/or the reporting and review process specifically. If the ‘constructive dialogue’ with the CmAT is invisible domestically, we would expect little to no

---

11 The countries for which a searchable media database existed include: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Guyana, Honduras, Mexico, Nicaragua, Paraguay, Peru, Uruguay, and Venezuela. No reliable and searchable databases could be located for Cuba, Guatemala, or Panama. We identified the top three newspapers (by circulation) using [www.pressreference.com](http://www.pressreference.com) supplemented with queries to regional and country experts or citizens. It was not possible to search electronically all three outlets for every single year since the CAT entered into force, but in each case we searched as many years as possible from CAT inception to the present, supplementing with information from the Foreign Broadcast Information Service (FBIS) for earlier years. See Online Appendix Table 5 for further information on years and publications searched for each country.

reference to any of these topics in the local press. But if such dialogue matters to domestic audiences, we should see a spike in press references, followed by a somewhat higher degree of attention to CAT and the CmAT during or after the review year.

The media evidence suggests the reporting process reverberates in the national press throughout much of the region. Figure 5 demonstrates strong evidence of a spike in attention by the local press to the review process in particular during the reporting year (0) for each state. Moreover, attention to the CmAT continued after the formal review was concluded. In the year following review, the press continued to report on the recommendations issued by the committee, but also covered CAT obligations generally much more than was the case in the pre-reporting years.

![Figure 5: Domestic Media Coverage of Torture in Latin America.](image)

Figure 5: Domestic Media Coverage of Torture in Latin America.
Figure indicates the number of domestic newspaper articles that reference the Convention against Torture (CAT), the Committee against Torture (CmAT) or the in-person periodic review before the CmAT and/or the concluding observations and recommendations issued by the CmAT (Review/Recs). Total articles within each country’s searchable time period were summed and averaged over fifteen Latin American countries: Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, Ecuador, El Salvador, Honduras, Mexico, Nicaragua, Paraguay, Peru, Uruguay, and Venezuela. Articles are centered around the year of review (0).

Figure 6 takes a closer look at four states. Mexico and Venezuela top the totals with more than thirty articles in their respective national presses the year of their CAT review. The press in Mexico in particular sustained interest in the implications of the review process in the
following year. With every submission, the CmAT responded to Mexican self-reports, often praising Mexico on areas of progress (for example, efforts to train law enforcement relating to torture prohibitions and human rights protections in general), and making increasingly detailed recommendations for better implementation and compliance.\footnote{A recurring theme of this dialogue was the need for Mexico to give much more attention to detention policies and practices, and in particular the extensive use of pre-charge detention practices, at both the state and federal levels.}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{chart.png}
\caption{Domestic Media Coverage of Torture in Latin America.}  
Indicates the number of domestic newspaper articles that reference the Convention against Torture (CAT), the Committee against Torture (CmAT) or the in-person periodic review before the CmAT and/or the concluding observations and recommendations issued by the CmAT (Review/Recs). References are averaged over the number of CmAT reviews (CAT reports) undertaken by each country within its searchable time period. References are centered around the year of review (0).
For Colombia and Mexico, the searchable period is 1995-present, and covered 3 CmAT reviews for each country. For Nicaragua, the searchable period is 2009-present, and covered 1 CmAT review. For Venezuela, the searchable period is 1998-present, and covered 3 CmAT reviews.

Given such intense interactions, it is not surprising that the Mexican press has mentioned

\footnote{In response to Mexico’s combined 5th and 6th report (2012), the CmAT appears to have made at least 56 recommendations.}
the CmAT extensively over the years. Since 2005, when systematic electronic search became possible for the major Mexican papers, *El Universal*—Mexico’s moderate left paper—has published twenty-one articles that referenced the CAT in some way. Eleven of these referenced the CmAT, of which eight were specifically about the periodic review process or CmAT recommendations. Again, we see the media coverage clustered in time around the CmAT’s list of recommendations. For example, *El Universal* covered the CmAT’s critical questioning of the Mexican representative during the periodic review of November 2006. Follow-up articles noted CmAT (and other) pressures on Mexico to address homicides and disappearances of women in Ciudad Juarez. A flurry of press reports covered CmAT review of Mexico’s report in 2012, covering the content of the report in detail and discussing concerns raised by civil society organizations as well.

Slightly to the political right and aimed more squarely at an elite business audience (though still decidedly independent and often critical of the government), *La Reforma* published thirty-one articles referencing the CAT since 1995. Thirteen of these mentioned the CmAT, and eleven were specifically on the periodic review process or CmAT recommendations. Many were critical of the government’s submission or explicitly highlighted the differences between the government’s and ‘shadow’ reports.* La Prensa*—with a wider circulation but with a reputation toward the more sensational—published fifteen articles referencing the CAT in some way since 2011, the first year for which systematic electronic search is possible. Fourteen of these articles reference the CmAT, all of which commented on the periodic review process and/or CmAT recommendations. Some articles are highly critical; others mentions the CmAT’s praise and examples of official progress. Several

---

20 *La Prensa*, “En Mexico crecio la tortura y el Estado la tolera, denuncian ONGs,” 28 November 2012
21 *La Prensa*, “Niega PGR que haya politica del Estado para la commission del delito de tortura,” 3
articles link domestic discussions of specific reforms with recommendations of the CmAT. In Venezuela, media attention appears to have fallen off drastically after review years, but note that in this case twenty-five articles reported on the recent 2014 review (combining Venezuela’s third and fourth reports); thus, follow-on coverage is effectively right-censored. Venezuela’s reporting has been spotty. Two years elapsed between its first and second reports, and then a dozen passed before Venezuela was heard from again. *El Universal* reported that the government had not complied with the recommendations of the CmAT in 1999, but it was the lone article we were able to find which noted the government’s delinquent reporting. The CmAT and the periodic review process remained virtually invisible in the media until 2014, when the government’s combined 3rd and 4th reports were under examination and some *twenty-five articles* covered the review. The pages of both *El Universal* (a somewhat conservative, business-oriented newspaper) and *El Nacional* (slightly more to the left) were saturated with news of the proceedings, reporting on claims and counter-claims. Since Venezuela had not undergone review in a dozen years, many articles were simply explanatory but shadow reports, church criticisms, and other critical voices were represented as well. Many articles covered the government’s view, justifying plans for public safety and criticizing some of the evidence presented in shadow reports—especially relating to government actions against ‘violent’ demonstrations earlier in the year—as ‘baseless allegations.’ Plenty of editorial comment thought that the government and police had a lot to answer for within the review process.

---


Nicaragua appears to display a pattern opposite to Venezuela’s. In this case we have a problem of left-censoring, since both of that country’s major national newspapers, *El Nuevo Diario* and *La Prensa*, were only searchable from 2009, the year of Nicaragua’s first and only CmAT review. Many of the articles that reference the CAT regime are highly critical; uncharacteristically, several of them link the CmAT with the view that Nicaragua’s 2006 ban on abortions under all circumstances is in violation of its obligations under the CAT. About a dozen of the articles graphed in Figure 6 for Nicaragua center around women’s issues, in particular, the state’s draconian abortion laws.

Press reports in Colombia exhibit the pattern we would expect to observe if the reporting process matters to domestic audiences: a spike in the year of review, strong attention the following year, and then moderate sustained attention that exceeds pre-review levels thereafter. Colombia has submitted four reports to the CmAT, each one more overdue than the previous one, with their most recent submission in 2008. *El Tiempo* has a prolific history of publicizing the reporting and response process over the last twenty years. It has published eighteen articles on the CAT, fourteen of which made specific mention of CmAT recommendations for Columbia, criticizing the state in particular on out-of-control security forces, practices by the military that amount to torture and the failing judicial system. Other news sources were more difficult to search systematically, but there is still evidence they covered the periodic review process as well. When dialogue over Colombia’s 5th report became quite drawn out between 2008 and 2012, references to the CAT became more common. An article in *El Spectator* in August 2011 explicitly publicized the reporting process and criticized the information that the government had forwarded to the Committee. Similar

29 Although Nicaragua was one of the first countries to sign the CAT in 1985, it did not complete the ratification process until July of 2005, with its first report due in 2006 and submitted in 2007.


to Argentina (below), in 2014 the Colombian press cited CmAT’s criticism of Taser use in relation to their use by the Bogota police force.\textsuperscript{35}

Several other states demonstrate similar patterns to those displayed in Figure 6. In Argentina, some fifteen articles in \textit{La Nacion} (from 2004 to the present) and ten in \textit{Clarin} (from 1997 to the present) referenced torture. Ten of these articles referenced the CmAT, and twelve discussed the process of periodic review or the CmAT’s specific torture recommendations. These are Argentina’s two most important papers, well respected, and with significant circulation.\textsuperscript{36} The timing of these articles is consistent with the mobilizing power of review and dialogue: they are clustered in years just after Argentina’s reports are discussed at the CmAT and approximately when recommendations were issued (1997/98, 2004/05, and 2010). Newspapers reported on the CmAT’s criticisms of prison conditions (2004) and its 2010 disapproval of Taser-type weapons as a violation of bodily integrity rights.\textsuperscript{37} The Committee’s concluding observations and recommendations have repeatedly made the news, which is consistent with mounting public pressure to comply with expert interpretations of the Torture Convention.

The Chilean press has also reported on the periodic review process, often mere days after Chile’s reports were examined by the CmAT. In May 2009—two days after CmAT’s review of Chile’s report—\textit{La Tercera} reported on an exchange between the International Federation of Human Rights and the government about Chile’s compliance with international standards and obligations under the CAT.\textsuperscript{38} In 2010, CmAT recommendations that Chile reform its anti-terrorism law were noted in \textit{El Mercurio}.\textsuperscript{39}

\textsuperscript{35} \textit{El Spectador}, “En Concejo de Bogota tambien critican a Policia por utilizacion de pistolas Taser,” 29 July 2014.


\textsuperscript{38} \textit{La Tercera}, “Federaci´on de DDHH acusa ‘rebajas en exceso’ de penas a cr´ımenes de lesa humanidad en Chile,” 7 May 2009; “Ministerio de Justicia refuta dichos de federaci´on de DDHH sobre tortura en Chile,” 7 May 2009.

\textsuperscript{39} \textit{El Mercurio}, “Relator de la ONU reitera llamado para reforma a Ley Antiterrorista,” 25 September 2010.
Paraguay turned in four reports, and has received a lot of advice from the CmAT, but it has proved difficult to access news media coverage systematically until quite recently. When the CmAT discussed the fourth report in 2011, however, there was fairly significant coverage (nine articles). Some predominantly praise the government’s policies,[40] while others are more critical. Similar to the cases of Argentina and Colombia, one article noted that the use of Tasers is considered torture by the CmAT. Another points to a shadow report on the treatment of indigenous persons in Paraguay.[41] Several articles point out the litany of deficiencies and accusations against the Paraguayan police and the widespread practice of preventive detention.[42] Overall, Paraguay represents a case of dense media coverage concentrated right around the time of CmAT review and recommendations.

Cases of far less coverage of the CmAT review process included Brazil, Ecuador, Peru, and Bolivia. Brazil has generally been much less cooperative than Argentina with the CmAT, and Brazilian media has pretty much ignored the lone set of recommendations by the CmAT. There have been indirect references in the press to the recommendation to ratify the Optional Protocol authorizing individual complaints,[43] as well as criticism of Brazil’s reporting delinquency.[44] In Ecuador, within the past decade we could find seven articles in El Universo that mention the CAT and four that specifically covered CmAT review and recommendations. An article in November 2010 discusses the detailed questions Ecuador faced with respect to treatment of indigenous people and the weakness of the judicial system in handling torture cases.[45] Another discussed the CmAT’s criticism of treatment of Colombian refugees in Ecuador by the security forces of both countries.[46] El Comercio in late 2005 covered Ecuador’s shortcomings as listed by the CmAT, but also took the opportunity in March of

---


2006 to deflect criticism toward the United States for its torture in Guantanamo.\footnote{\textit{El Comercio}, “No tiene derecho,” 20 March 2006.}

Peru has had superficial but improving engagement with the CmAT over time. Peru’s reporting quality improved noticeable from the 1990s to the 2000s, although their reports tended to be a couple of years late, to the point that it became necessary to combine reports in 2012 in order to ‘catch up.’ We found only five references to the CmAT review process in the Peruvian press. Three were in 2006 when the 4th report was under examination; all were found in \textit{Peru.21} (founded in 2002) and were largely explanatory rather than critical. Two articles covered review of the combined 5th and 6th reports in 2012, noting the CmAT’s criticism of deplorable prison conditions in Peru.\footnote{\textit{Peru.21}, “ONU pide al Perú que evalúe cierre de los penales de Yanamayo y Challapalca,” 23 November 2012; Ojo, “ONU ordena a Perú cerrar dos carceles,” 24 November 2012.}

Next door, Bolivia has turned in only two reports (the second was seven years late, but of higher quality than the first). Four media articles have mentioned the CmAT, with the three in \textit{El Deber} the most explicit about the 2013 periodic review, including the Committee’s specific recommendations.\footnote{\textit{El Deber}, “ONU pide a Bolivia tipificar delito de tortura conforme a ley internacional,” 31 May 2013; “Tóásó se queja en la CIDH y la ONU exige ley contra la tortura,” 1 June 2013.} One is highly critical of the fact that human rights were in retreat in Bolivia and used the CmAT recommendations to bolster demands for changes in the penal code to conform to the CAT.\footnote{\textit{El Deber}, “Aumentan torturas y abusos a manos de policías y militares,” 8 September 2014.}

This review of the media suggests that the causal mechanism linking the reporting process to domestic information, awareness, and eventual mobilization is highly plausible. The press records for Latin America generally demonstrate the expected spike in reporting within local news media around the time the CmAT reviews government reports and issues its recommendations for how a country can improve its implementation of and compliance with the treaty.\footnote{We found little to no evidence of coverage of CmAT dialogue in some of the smaller states in Latin America, but had severe problems with electronic access to news databases. See Online Appendix for notes on Guyana, Panama, Costa Rica, El Salvador, Honduras and Guatemala.} Interestingly, we found little publicity when reports were actually submitted. Certainly, there is a good deal of variance across countries. Larger states, such as Mexico, tend to have higher capacity media organizations able to cover international processes of all.
kinds. And yet the media in some polities that possess similar capacities, such as Brazil, fail
to do so. Somewhat surprisingly, press coverage was fairly critical of the state in several cases
in which democratic institutions overall are somewhat weak (Venezuela, for example). But it
is clear that the periodic review process under the Torture Convention has not gone unnoticed
within the region. On the contrary, in many cases and indeed on average, newspapers
have given the process of review and CmAT recommendations the kind of airing one would
expect for the process to mobilize domestic pressures for change. While this is not proof
of every step in the mobilization mechanism—a discussion we cannot pursue here due to
space constraints—this evidence is certainly consistence with the importance of dialogue in
stimulating public discussions upon which such mobilization could plausibly develop.

Conclusion

Criticisms of the reporting and review process of the various human rights treaty bodies
are far more common than rigorous assessments of their actual consequences. While we do
not dispute there are weaknesses in the reporting system, the evidence presented here is a
striking contrast to the bulk of the literature on state interactions with expert human rights
treaty bodies. We have found evidence to suggest that dialogue between state representatives
and international experts may indeed generate new ideas, advice, and potentially pressure
for change in practice. The dialogue engendered through self-reporting may well have been
important to making improvements on the ground that reduced the pervasiveness of torture
in a number of countries (indeed, on average). Importantly, we found evidence that prompt
interaction with the CmAT reduces noticeably the probability of high torture scores. As
anyone would expect, the shift in the probability of torture due to dialogue is not massive,
but it is all the more believable for its modest size.

It is very important as well to stress what we did not find. We found evidence that
interactive histories are important to such improvements; modeling one-shot effects proved
useless. The number, the density, and the timing of dialogue are crucial to the process of rights improvements. This is certainly far more realistic a finding than to expect last year’s conversation with the experts to yield one-shot effects in the following year or so. We also found that iterative and interactive histories are key: reporting per se does not produce the same results as a more intense back-and-forth between governments and the Committee. Review histories had a causal influence on the probability of frequent torture; reporting histories alone did not. This suggests there are limits to self-assessment but potential for constructive engagement.

One reason for this finding is that the review process gains a domestic audience through the national media. Far from finding that no one pays attention to this process outside the halls of Geneva, it turns out that in Latin America at least the review process literally piques (or peaks, as in Figures 3 and 6) the media’s, and potentially the public’s, interest. The national media in this region is replete with discussions and debates about what governments are telling the experts, how shadow reports shape the conversations, what the CmAT has asked, and how governments have responded. There is plenty of official excuse-making going on, but a surprising amount of criticism as well. These patterns are consistent with a theory that treaties matter because discussing human rights engages interested domestic publics, who are in a better position, armed with legal rights and better information, to hold their governments accountable.

We hasten to add that the reporting regime is not a comprehensive solution to the world’s worst human rights abuses. For one thing, we have only examined states that have ratified the CAT; even among the ratifiers it has proved impossible to coerce a meaningful conversation out of unwilling states. Constructive dialogue only has effects when it actually takes place. That said, the results of this research suggest that the reporting and review system should be supported rather than disparaged. We agree with the critics who point out the problems of stretched resources and redundant processes. But a look at the evidence suggests that self-reporting has an important causal role to play in starting conversations
that reverberate domestically and open possibilities for change. Given the lack of a central international authority dedicated to information production on a broad scale, these findings on self-reporting are of central significance for international governance.
References


