RICE UNIVERSITY
Domestic Politics, NGO Activism, and Global Cooperation

By

Jessica Edry

A THESIS SUBMITTED
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE

Brett Ashley Leeds
Professor of Political Science, Department Chair

Eduardo Salas
Professor and Myer R. & Gladys M. Cline Chair of Psychological Sciences, Department Chair

Songying Fang
Associate Professor of Political Science

Rick K. Wilson
Herbert S. Autrey Professor of Political Science and Professor of Statistics and Psychology

HOUSTON, TEXAS
April 2020
ABSTRACT

Domestic Politics, NGO Activism, and Global Cooperation

by

Jessica Edry

Can international institutions affect state behavior, and if so, how? This question has motivated decades of research that has substantially deepened our understanding of how international processes affect states. This dissertation contributes to this body of research by focusing on global cooperation in environmental politics and human rights. This type of cooperation poses complex challenges because negotiators face incentives to water down treaty commitments in order to enhance participation and strategies for enforcing obligations are limited. I argue that the ability of these institutions to affect state behavior by encouraging domestic pressure for fulfillment of states’ international commitments makes certain types of institutions more likely to deliver intended effects on state behavior. I also suggest, however, that this mechanism may be in danger because of growing restrictions on non-governmental organizations (NGOs) whose work is critical to the mobilization process. Therefore, this dissertation contributes new insight into the link between international institutions, domestic political dynamics, and state behavior, and offers policy recommendations for enhancing the impact of such institutions.
Acknowledgements

I am eternally grateful to everyone who has supported me throughout this journey toward my PhD. First and foremost, I would like to thank my advisor, Ashley Leeds, for being a wonderful mentor and friend. From the moment I met Ashley, I could see how dedicated she is to her students, her research, and her department. Her enthusiasm made me want to come to Rice, and I’m so glad I trusted my instincts and accepted the offer. From day one, Ashley has challenged me to sharpen my thinking and be clearer in my writing. She has read countless drafts of my papers and yet somehow she is always ready to read more. She never wavered in her support for me, even as I struggled to articulate my career aspirations and eventually decided to pursue a non-academic path. She has shown me what it means to be a mentor and taught me lessons that I know will serve me well throughout my career.

Second, I would like to thank Rick Wilson for teaching me what it means to be an experimentalist and taking a chance on me by funding my first experiment in the Behavioral Research Lab. I am impressed by his ability to tease apart a research design and not only identify the threats to inference but also provide suggestions for how to move forward. His feedback has made my work immeasurably better, and I am grateful for his patience and generosity (even though I know he will roll his eyes when he reads this).

Third, I would like to thank Songying Fang for giving me helpful feedback throughout my time at Rice and challenging me to think harder about the claims I am trying to make. I would also like to thank Eduardo Salas for agreeing to serve on my dissertation committee and raising important questions that helped me make my work more accessible to readers outside of political science.

Next, I would like to thank the community at Rice political science for supporting me every step of the way both intellectually and emotionally. Faculty who were not on my dissertation committee consistently gave me valuable feedback and advice.
fellow graduate students have created such a wonderful sense of community that I will sorely miss. In particular, I am thankful for the strong support network formed by the amazing women in my cohort—Liana, Kaitlin, Kristin, and Carly. I can’t believe how lucky I was to come to Rice at the same time that they did.

Finally, I would like to thank my husband, Adam, for following me halfway around the world from Israel to Houston and greeting every challenge with a smile and an encouraging hug. His never-ending love and support has made this journey considerably easier, and I am forever grateful to him for his kindness, thoughtfulness, and lame jokes.

Last but certainly not least, I would like to thank my parents for believing in me and providing me with a solid foundation from which to grow and learn. I can only hope that one day I will be able to give my children the same loving foundation that will give them the confidence to believe they can do anything they set their minds to.
List of Figures

2.1 Substantive Effect of Democracy ........................................... 41
2.2 Substantive Effect of Report Consideration .......................... 43

3.1 Mean Contributions by Period and Session Type .................... 74
3.2 Mean Punishment Rates by Period and Session Type .............. 76
3.3 Substantive Effect of Government Contribution by Experimental Condition ................................................................. 79

4.1 Average Treatment Effects: Experiment 1 ............................ 108
4.2 Average Treatment Effects: Experiment 2 ............................ 110
4.3 Heterogeneous Effects by Ideology: Experiment 1 .................. 112
4.4 Heterogeneous Effects by Ideology: Experiment 2 ................. 114
List of Tables

2.1 Multilevel Logistic Analysis of Withdrawing Reservations ................. 40
2.2 Heckman Probit Models ...................................................... 45
2.5 Objective Level of Respect for Women’s Rights ............................... 58
2.6 Annual Change in Respect for Women’s Rights ............................... 59

3.1 Multilevel Logistic Analysis of Interest Group Punishment ................. 78
3.2 Summary of Pairwise Comparisons for Leader Contributions .................. 85
3.3 Summary of Pairwise Comparisons for Interest Group Punishments ........ 86

4.1 Summary of Experimental Conditions ........................................... 103
4.2 Hypothesized Statistical Tests .................................................... 106
4.3 Sample Sizes by Experiment, Condition, and Ideology ....................... 111
Chapter 1

Introduction

In November 2018, UN Secretary General António Guterres stood before the UN Security Council to make the case for increased multilateral cooperation to confront global problems (Guterres 2018). According to Guterres, cooperation on a global scale is crucial to mitigating existing threats in a number of areas, such as climate change and refugee flows. However, there are grounds for skepticism over whether global cooperation, specifically through the creation of new institutions that seek to engage the entire international community, can meaningfully contribute to these efforts.

This dissertation asks whether global cooperation, particularly in the contexts of environmental politics and human rights, can make a difference. I focus on environmental politics and human rights because the challenges associated with creating effective global institutions are particularly severe in these areas. In both environmental politics and human rights, widespread participation is not only desirable but essential to treaty effectiveness. Yet because participation is so important, negotiators face a pressing trade-off between designing strong institutions that encourage progress toward the treaty’s goals and creating institutions that will win sufficient support from a diverse community of states (Gilligan 2004). As a result, many global treaties in these areas include flexibility mechanisms that ease the treaty’s acceptance but threaten to undermine its objectives by enabling states to make shallow
commitments that promise little policy change (Helfer, 2013). Moreover, scholars of international cooperation have long recognized that traditional strategies of enforcing international commitments are less applicable to these issue areas (Guzman, 2008; Downs and Jones, 2002; Simmons, 2009). These challenges contribute heavily to the skepticism surrounding the utility of global cooperation in confronting environmental problems and improving respect for human rights.

Given these challenges, is there any hope for effective global cooperation in these areas? This dissertation is composed of three stand-alone essays, each of which deals with a separate piece of my answer to this question. In the first essay, I argue that the perceived trade-off between depth and participation is less severe than scholars have been led to believe because initially shallow commitments can and often do deepen over time in response to domestic mobilization. Previous research offers compelling evidence that international institutions can indirectly influence state behavior by encouraging mobilization among domestic groups that advocate for greater respect for international commitments (Simmons, 2009; Dai, 2007). I connect this literature to the debate over the effects of flexibility in international institutions and offer evidence to suggest the growth of domestic mobilization alleviates these concerns. Next, I argue that the existence of this domestic mobilization mechanism implies global institutions can deliver meaningful impacts on state behavior so long as they are designed with information-producing mechanisms. I use a laboratory experiment to generate causal evidence in support of my claim that international institutions have greater effects on state behavior when they consistently provide information for domestic audiences.

However, in the final essay I point to ways in which this indirect link between treaties and state policy change may be weakening as governments devise new strategies of undermining the civil society groups that are critical to making this process work. This chapter explores the growing phenomenon of states imposing restrictions on non-governmental organizations (NGOs) that receive foreign funding. I use a sur-
vey experiment to show how labelling NGOs as foreign-funded weakens their ability to influence public attitudes toward policies that violate international law. These findings suggest these restrictions and the discourse that surrounds them will make it more difficult to rely on domestic mechanisms of treaty enforcement in the future. For policymakers, this further underscores the need for international actors to push back against this trend.

In the following sections, I discuss the importance of global cooperation in environmental politics and human rights and the challenges posed by these types of treaties. Then, I consider the role domestic mobilization can play in alleviating these concerns. This indirect link between international institutions and state behavior through the mobilization of relevant domestic groups serves as the theoretical basis for this dissertation. Finally, I present the plan for the dissertation and outline its contributions to the literature on international cooperation and ongoing efforts to design effective global institutions.

1.1 The Challenge of Global Cooperation in Environmental Politics and Human Rights

This dissertation focuses on global cooperation in environmental politics and human rights because global treaties in these issue areas face particularly acute obstacles to effectiveness. Global cooperation refers to treaties that aim to unite all members of the international community in an effort to address a global concern. Examples of this type of cooperation include the Paris Agreement on Climate Change and the core international human rights treaties. This definition therefore excludes multi-lateral treaties that are not open to all states in the international system or have a regional focus, such as the North American Free Trade Agreement (Elsig, Milewicz and Stürchler, 2011).
Designing global treaties to deal with environmental and human rights-related concerns is especially difficult because widespread participation is viewed as critical to treaty effectiveness. In the context of climate change, the benefits of reduced emissions will only be realized if a sufficient number of states contribute to international efforts (Milinski et al., 2008). Moreover, many advocates and diplomats believe broad membership is critical to productive cooperation, in part because it confers greater legitimacy on the initiative (Victor, 2006). In fact, studies find higher levels of participation increase public support for global climate agreements (Bechtel and Scheve, 2013) and support for reducing emissions rises when other countries are said to be implementing similar measures (Tingley and Tomz, 2012). The perceived link between participation, legitimacy, and effectiveness is also evident in the human rights arena. For example, Dai (2007) suggests the humanitarian provisions in the Helsinki Final Act were especially powerful because they were endorsed by all countries in Europe.

When universal participation is viewed as essential to the treaty’s effectiveness, however, negotiators face a trade-off between designing strong institutions and broadening the treaty’s appeal. On the one hand, precise, legally binding obligations within a centralized institution make it easy to detect and punish violations. Therefore, these more legalized design features should enhance compliance with treaty provisions and increase the likelihood that the treaty will achieve its stated goals. Yet a strict treaty with high costs for noncompliance may discourage participation among states that are uncertain about their ability to comply with its provisions. Scholars refer to this predicament as a “broader-deeper trade-off,” where breadth refers to participation and depth refers to the extent to which states will need to modify their behavior to comply (Gilligan, 2004; Bernauer et al., 2013).

Importantly, scholars who study this dilemma generally contend that it is not as clear-cut as it may seem. Gilligan (2004) shows the broader-deeper trade-off only exists when the cooperation problem at hand dictates states set identical policies.
Once states are permitted to set their policies at different levels, the trade-off disappears. Meanwhile, Bernauer et al. (2013) argue that the depth of an agreement is multidimensional, and specific design features can have different effects on participation. Examining a dataset of global environmental treaties, they find agreements with more precise obligations attract fewer participants, but monitoring and enforcement provisions are not consistently related to participation rates.

These studies are important because they tell us we need to think more deeply about different ways in which the broader-deeper trade-off may manifest, but they leave a number of questions unanswered. If states set their policies at different levels, as Gilligan (2004) suggests, how can we expect the treaty to deliver its intended effects? Might states less committed to the treaty’s cause join and set their policies at the levels they would have kept if they had remained outside the institution? This is precisely the concern raised by critics of human rights treaties, who argue the focus on widening participation renders treaties so weak that they struggle to deliver positive effects (Hafner-Burton, 2013). Moreover, Bernauer et al. (2013) note that flexibility in treaty design should make agreements more palatable to states, but the only flexibility mechanism they investigate is imprecision. As Koremenos (2016) explains, there are a host of other mechanisms, such as escape clauses, withdrawal clauses, duration provisions, and reservations that can be used to make treaties more flexible.

Reliance on flexibility mechanisms to enhance treaty participation creates a new version of the broader-deeper trade-off that is even more difficult to resolve. On the one hand, increased flexibility should make the treaty more appealing to states uncertain about the benefits or costs of compliance (Koremenos, Lipson and Snidal, 2001; Koremenos, 2005; Rosendorff and Milner, 2001) and may even encourage states to make deeper commitments than they would otherwise be willing to accept (von Stein, 2008; Kucik and Reinhardt, 2008). On the other hand, flexibility may under-
mine the entire cooperative endeavor because it encourages “opportunistic behavior” in which states legally flout international commitments when compliance becomes inconvenient (Helfer 2013, 176). The overall question of whether flexibility helps or hinders the creation of effective global institutions remains unresolved.

These concerns are further compounded by the fact that traditional mechanisms of enforcing international obligations are less applicable to treaties in environmental politics and human rights. Scholars of international cooperation generally speak of three mechanisms—reciprocity, retaliation, and reputation—that make treaties self-enforcing, meaning that states comply with their obligations without being compelled by some external governing body (Guzman 2008; Simmons 2009). While the application of these mechanisms to economic and security issues is fairly straightforward, they are less potent when the subject of cooperation is environmental politics or human rights. I discuss each mechanism in turn.

In the reciprocity mechanism, states abide by their commitments because they know that reneging on a promise will lead other states to deny them cooperation in response (Axelrod 1980). In the context of trade, a state can raise tariffs on a violating country to deny them the benefits of cooperation. In the security realm, if a state’s ally in a mutual defense pact fails to fulfill its obligations during wartime, that state can break off the alliance, leaving the violator more vulnerable to attack. By contrast, refusing to cooperate with violating states has little meaning in a human rights context and may be impossible when environmental protection is the subject of cooperation. For example, refusing to respect human rights in response to another state’s violations simply means two populations suffer abuses rather than one. In the case of environmental protection, states often cannot prevent a violator from enjoying the benefits of cooperation between other states. For instance, if a state refuses to fulfill its commitment to reduce air pollution, the cooperating states cannot deny the violator access to cleaner air. In this case, then, reciprocity is of little use.
In the retaliation mechanism, states fulfill international commitments for fear of punishment from other states. One prominent form of retaliation is economic sanctions. Crucially, states are willing to incur the cost of imposing sanctions because it signals that violations of this or other agreements will not be tolerated and therefore induces further compliance (Guzman 2008). For example, the United States, United Nations, and European Union imposed sanctions on Iran after it violated the Nuclear Nonproliferation Treaty (Laub 2015). This retaliation aimed to bring Iran into compliance and send a message to other states that such actions would not be tolerated. Thus, the benefits of preventing the spread of nuclear weapons and preserving the power structure enshrined in the nuclear regime outweighed the costs of imposing sanctions.

When violations are related to environmental politics or human rights, however, it is less obvious why states would be willing to pay the costs of retaliation, especially given that improving environmental protection or respect for human rights in another country does not typically have direct effects on other states. This is not to say that retaliation is never used to enforce environmental or human rights obligations, but rather its use is less common compared to economic and security obligations. Tellingly, less than four percent of recorded cases of states threatening or imposing economic sanctions between 1945 and 2005 seek to change environmental policies in the target state and just under eight percent are related to human rights practices (Morgan, Bapat and Kobayashi 2014). By contrast, nearly fifty percent of sanctions episodes target states’ trade policies (Morgan, Bapat and Kobayashi 2014). There are also a number of recorded instances of states using military force to retaliate against states that violate human rights, but this costly action tends to be reserved for humanitarian emergencies when large-scale violations are rampant (Binder 2015; Choi 2013). Violations viewed as less dire, such as gender discrimination, child labor, and human trafficking, are highly unlikely to invite this kind of punishment.
In the final mechanism, reputation, states comply with international commitments because they seek to safeguard their reputation for credibility. Reneging on an agreement signals to observers that the state cannot be trusted to fulfill its promises, which makes it more difficult to conclude future agreements. In the economic and security issue areas, for example, studies show that states in compliance with their bilateral investment treaties enjoy higher levels of foreign direct investment (FDI) compared to states accused of violations [Allee and Peinhardt, 2011] and states that have honored previous alliance commitments are more likely to be chosen as alliance partners in the future [Gibler, 2008; Crescenzi et al., 2012]. However, many scholars believe reputation is less effective in motivating compliance in the context of environmental or human rights treaties [Downs and Jones, 2002]. Even scholars who advocate for reputation-based explanations of human rights practices acknowledge that when reputation is defined as “credibility of commitments,” it has little relevance in the context of human rights [Kelley, 2017].

Thus, global cooperation in environmental politics and human rights is especially challenging because the need for broad participation pushes negotiators to rely on flexibility in treaty design. Furthermore, because traditional mechanisms of enforcing obligations are less potent in these areas, it should be more difficult for these treaties to prevent states from abusing flexibility mechanisms. It is important, therefore, to understand whether these treaties can be expected to have any effect on state behavior and whether there are specific actions policymakers or negotiators can take to enhance their effectiveness. That is the task to which this dissertation seeks to contribute. To that end, in the following section I discuss how recognizing the domestic effects of international treaties can contribute to this endeavor.
1.2 The Role of Domestic Mobilization and NGO Activism

Recognizing the limitations of traditional enforcement mechanisms in the context of environmental politics and human rights, scholars have looked to domestic pathways through which international institutions may impact state behavior. One important finding from this literature is that international institutions encourage domestic mobilization in support of compliance with international obligations, often led by NGOs (Simmons, 2009; Dai, 2007). Hence, enforcement can come from within a state’s domestic political environment rather than from interactions with other states.

Domestic mobilization refers to coordinated efforts by domestic actors to pressure their government to change existing policies. This mobilization can take a variety of forms, such as circulating petitions, organizing protests, meeting with legislators, and publicizing information about government deficiencies in meeting international obligations. These actions force the issue of compliance with international commitments onto the government’s agenda and communicate to politicians that ignoring these issues could have negative repercussions. In democratic systems, domestic groups displeased with a government’s level of compliance can campaign against its reelection (Dai, 2007). Even in nondemocratic settings where free and fair elections are rare or nonexistent, public discontent is still undesirable because repression is costly and public dissatisfaction presents a window of opportunity for aspiring leaders to challenge the existing government (Bueno de Mesquita et al., 2003; Dimitrov, 2009).

Simmons (2009) offers a useful way to conceptualize the domestic effects of international institutions. She argues that human rights treaties increase the likelihood of domestic mobilization in favor of greater respect for human rights because they enhance the value of successfully changing government policy and raise the probability such success will occur. Mobilization can be quite costly—aside from the opportunity costs associated with spending time lobbying for policy changes, challenging govern-
ment policy may invite retaliation from government representatives or supporters. Thus, individuals will be unlikely to engage in this behavior unless they believe that they will benefit from these policy changes and that such change is feasible. In the context of human rights, individuals are most likely to mobilize when they value the rights in question and believe there is a relatively high probability that their demands will be met (Simmons 2009, 135).

Human rights treaties affect this calculus by introducing new ideas about the rights individuals should be able to claim and improving the chances that mobilization will lead to greater respect for these rights (Simmons 2009). The International Covenant on Civil and Political Rights (ICCPR), for example, states individuals are entitled to freedom of association, freedom of assembly, and equal protection of the law. In some settings, these rights are already respected. In some countries, however, conceiving of these concepts as “rights” may be new. In the case of the Convention on the Elimination of all forms of Discrimination Against Women (CEDAW), the idea that women should consider themselves equal to men and therefore deserving of rights traditionally afforded to men can be quite controversial precisely because it conflicts with the way many women have been taught to think of themselves. When a government ratifies these treaties, domestic groups such as NGOs often take the lead in educating people about these rights and building public support for their fulfillment (Simmons 2009; Merry and Stern 2005; Tsutsui and Shin 2008; Tsutsui, Whitlinger and Lim 2012). The value of successful mobilization rises as individuals become more aware of what they stand to gain from achieving greater respect for international human rights commitments.

In addition to increasing the incentives for mobilization, Simmons (2009, 144) argues that human rights treaties also improve the probability of successful mobilization by precommitting the government to be receptive, increasing the size of the pro-rights coalition, enhancing intangible resources, and expanding the range of strategies avail-
able to the movement. It is more difficult for a government to deny greater respect for rights it has formally acknowledged through ratification. If the government does resist domestic demands for policy change, the fact that the state has legally recognized the treaty’s authority should make it easier for pro-rights groups to find allies in lawyers and government opponents (Simmons, 2009, 145). Moreover, domestic groups can use treaties to legitimate their claims and provide focal points for action. Finally, the existence of a legal commitment opens up the potential for domestic actors to press their claims both in the public sphere and in the courtroom (Simmons, 2009).

Although Simmons (2009) presents her argument in the context of human rights treaties, the theory is broadly applicable to other issue areas. Dai (2007) suggests environmental institutions can supply information to domestic actors that helps them identify deficiencies in government environmental policy and tie electoral support to compliance. Within the trade regime, states that initiate formal disputes against other governments’ trade practices alert domestic audiences to their government’s alleged violations and raise the spectre of domestic punishment (Chaudoin, 2014a; Pelc, 2013; Matsumura, 2015; Mansfield, Milner and Rosendorff, 2002). In the security realm, approval or condemnation from the UN Security Council can influence public support for a government’s foreign policies and therefore constrain state behavior (Fang, 2008; Chapman, 2009, 2007).

Importantly, NGOs are widely recognized as critical actors in the domestic mobilization process (Hafner-Burton, LeVeck and Victor, 2016). NGOs collect and disseminate information to increase public awareness of government obligations and gaps in compliance (Keck and Sikkink, 1998; Sikkink, 2017; Simmons, 2009). From educating citizens about the content of international treaties to “naming and shaming” governments that violate their obligations (Murdie and Davis, 2012; Murdie and Peksen, 2015), NGOs communicate to citizens what they can expect to gain from mobilization and when such action is needed. Beyond simply relaying this information to the
public, however, NGOs also influence how domestic audiences interpret and interact with this information. In the human rights arena, NGOs “translate” global human rights standards into the local vernacular (Simmons 2009; Merry and Stern 2005). For example, NGOs helped galvanize support for policies to address violence against women in part by adapting international ideas to local contexts (Merry 2006; Htun and Weldon 2012). In environmental politics, NGOs draw attention to environmental abuse and create new frames through which audiences interpret these reports. For instance, NGOs opposed to whaling re-framed whales as sentient animals deserving of protection to drum up public support for government compliance with the International Convention for the Regulation of Whaling (Wapner 1995; Eilstrup-Sangiovanni and Bondaroff 2014). NGOs also help citizens use international treaties to press legal claims against their governments and hold them accountable for human rights violations (Sikkink 2011; Shelton 1994).

While scholars often point out that international institutions can spur mobilization among both supporters and opponents of compliance with treaty commitments, there is reason to believe that the presence of an international institution advantages pro-compliance groups more than anti-compliance groups. Survey experiments generally show that violating international law depresses support for government policies (Wallace 2013; Chaudoin 2014b; Kreps and Wallace 2016), which suggests domestic audiences are typically more sympathetic to the arguments of pro-compliance groups. Additionally, international institutions facilitate information and knowledge transfers directly to pro-compliance groups. For example, international human rights NGOs guide domestic NGOs through the periodic reporting process, helping them craft alternative “shadow” reports and training them on how to appear before the relevant committee and implement recommendations in their home countries (Zwingel 2005). These formalized interactions give pro-compliance groups an edge over anti-compliance actors.
Thus, the domestic mobilization mechanism offers an indirect pathway through which international institutions can affect state behavior that relies heavily on the work of NGOs. In this dissertation, I build on this research to derive expectations about variation in the effectiveness of international institutions and highlight ongoing trends that may serve to weaken this mechanism in the future.

1.3 Plan for the Dissertation

This dissertation consists of three stand-alone essays. The first two essays apply the logic of the domestic mobilization mechanism to the question of how international institutions can be designed to maximize both participation and intended effect. In the first essay, I argue that the process of domestic mobilization mitigates the supposed detrimental effects of flexibility in institutional design. I develop a general theory of how initially shallow commitments can deepen over time in response to domestic mobilization. Applying this theory to the human rights regime, I link domestic mobilization to states’ withdrawal of treaty reservations and identify the conditions under which such mobilization is most likely to occur. Finally, I use large-N statistical analysis to show how factors that support domestic mobilization are associated with states’ withdrawal of reservations to a major human rights treaty, the Convention on the Elimination of all Forms of Discrimination Against Women (CEDAW). This chapter suggests the trade-off between depth and participation is less severe than it may seem and negotiators should not shy away from incorporating flexibility when it is needed to expand participation. This is not to say that flexibility mechanisms have no downside, but all is not lost if a treaty allows states to make initially shallow commitments.

In the second essay, I offer causal evidence to support the assertion that international institutions have greater effects on state behavior when they consistently provide information for domestic audiences that can spur mobilization. A funda-
mental difficulty inherent to studying the effects of international institutions on state behavior is the endogeneity of treaty design and commitment. Because states decide what will be included in a treaty and whether to ratify it based on an assessment of their willingness and ability to comply with its terms, it is difficult to separate behavior induced by the treaty from behavior that would have occurred in its absence (Simmons and Hopkins, 2005; Simmons, 2000; von Stein, 2005; Braumoeller et al., 2018; Downs, Rocke and Barsoom, 1996). In this chapter, I use a laboratory experiment to investigate how information provided by an institution impacts contributions to a public good where some minimum threshold of contributions must be achieved before the public good is provided. By randomly assigning subjects to different institutional arrangements in a laboratory setting, I am better able to pinpoint the causal effect of information on public goods provisions. While the laboratory is certainly no replacement for observational data, it serves as a useful complement to observational research and provides further evidence for my claims.

While the first two essays suggest future international institutions should be designed with domestic mobilization in mind, the final essay adds an important caveat. Given the critical role that NGOs play in domestic mobilization, policies designed to hamper civil society pose a real threat to this mechanism. Government repression against civil society is not new, but it has evolved in recent years to include legal restrictions on NGO activity, especially those that receive foreign funding (Dupuy, Ron and Prakash, 2016; Christensen and Weinstein, 2013; Bakke, Mitchell and Smidt, 2019). The final chapter of this dissertation explores the growing phenomenon of states directing criticism toward foreign-funded NGOs and passing laws that require such organizations to identify themselves as “foreign-funded” in their publications. I argue that labelling an NGO as foreign-funded reduces its ability to influence public opinion toward policies that violate international law because individuals view foreign-funded NGOs as more biased and therefore less credible sources of informa-
tion. I find partial evidence for this argument in a set of survey experiments situated in the environmental politics and human rights contexts. Because NGOs advocating for compliance with human rights and environmental commitments often rely heavily on foreign funding, these dynamics threaten to weaken the domestic mobilization mechanism in the future and present a new challenge for scholars, activists, and policymakers who wish to see global cooperation succeed.

1.4 Contributions

This dissertation makes three central contributions. First, this dissertation advances the literature on international cooperation by connecting theories of domestic mobilization to the question of how global treaties can be designed to deliver intended effects. The domestic mobilization argument emerged as part of a research program focused on determining whether treaties matter at all (Simmons 2009; Dai 2007). More recent research, however, moves beyond this dichotomy to ask which treaties matter the most (e.g., Long, Nordstrom and Back 2007; Mattes 2018; Hafner-Burton and Tsutsui 2005; Büthe and Milner 2014). This dissertation makes a novel contribution by suggesting that domestic mobilization has implications for which treaties should be most effective, especially in the context of the debate over the merits of flexibility in international institutions.

Additionally, my findings with respect to agreement flexibility challenge recent work. For example, Linos and Pegram (2016a) suggest the trade-off between firm and flexible treaty language is particularly acute because politicians can use the vague language to argue that such action is unnecessary, while firmly worded obligations facilitate mobilization in support of compliance. They show that countries were more likely to implement strongly-worded recommendations from the 1991 Paris Principles on the Design of National Human Rights Institutions (NHRIs) compared to weakly-worded recommendations, and implementation of flexibly-worded tasks was
no different from omitted provisions. However, I argue this trade-off is less severe than it might appear because states’ commitments and levels of compliance change over time in response to domestic mobilization. I also highlight institutional features, such as periodic review, that can catalyze domestic mobilization processes, and I provide experimental evidence of the power of the information this procedure produces.

Second, this dissertation offers valuable advice to policymakers charged with designing future multilateral initiatives, especially in the context of human rights and environmental politics. Specifically, this dissertation suggests negotiators should not be overly concerned with limiting the use of flexibility mechanisms that promise to broaden participation at the expense of allowing for more shallow initial commitments. Instead, negotiators should focus on designing institutions that provide useful information for domestic audiences. This means incorporating procedures such as periodic review and creating mechanisms to engage domestic NGOs in the process of collecting and analyzing information on government practices. These insights are especially valuable in the current political environment, as international actors continue to seek ways to engage states in efforts to mitigate climate change and assess the impact of the human rights regime.

However, this dissertation also offers a word of caution. The final contribution of this project is to show how governments can dampen the effects of international institutions on domestic politics. As governments become more directly involved in the domestic mobilization process by criticizing and imposing restrictions on NGOs with ties to the international community, this mechanism of enforcing international obligations may be weakening. This dissertation provides the first experimental evidence to show the consequences of these developments for the link between NGOs and public opinion. Future research can build on this framework to investigate how these effects vary cross-nationally and among subgroups within a given population. Meanwhile, this evidence suggests policymakers are right to be concerned about this trend, and
efforts to prevent or repeal undue restrictions on civil society will be critical to the continued effectiveness of global cooperation.

Thus, this dissertation contributes significantly to our understanding of where and when international institutions are likely to have substantial effects on state behavior and how these dynamics may be changing as government policies evolve. These insights are most useful in the context of global cooperation for human rights and environmental protection, but the mechanisms I highlight here are broadly applicable to economic and security issues as well. Therefore, this dissertation helps answer the big question of whether global cooperation can improve the human condition.
Chapter 2

Shallow Commitments May Bite Deep: Domestic Politics and Flexibility in International Cooperation

When the Paris Agreement on climate change was reached in December 2015, then-UN Secretary General Ban Ki-Moon praised the development as an “historic moment” and remarked, “For the first time, we have a truly universal agreement on climate change, one of the most crucial problems on earth” (Davenport, 2015). However, as observers began examining the accord, skepticism about its ability to mitigate climate change began to surface. To get almost 200 countries to support the agreement, negotiators relied on a critical flexibility mechanism—allowing states to set their own mitigation commitments, known as “nationally determined contributions (NDCs)” (Victor, 2016). While some scholars view this flexibility as an asset that should make the agreement more durable and perhaps more effective (Victor, 2016; Stern, 2017), others worry that compromises in the agreement’s design intended to widen participation will ultimately undermine its objectives (Linos and Pegram, 2016).

The Paris Agreement highlights a vexing question in international cooperation—is watering down an agreement to enhance its appeal worth the cost? Allowing states to make half-hearted commitments may ensure near-universal participation, but critics counter that it would be more effective to force states to choose between meeting a higher standard and staying out of the regime altogether. This question is nearly
impossible to answer because doing so would require comparing existing treaties to counterfactual institutions with different design features. However, one can still ask whether allowing states to make shallow commitments that require little change in behavior is truly as damaging as critics suggest.

This paper argues the alleged trade-off between depth and participation is less severe than scholars and policymakers have been led to believe, because initially shallow commitments can and often do deepen over time. Even when states make half-hearted commitments to a treaty, the institution encourages mobilization among domestic actors who wish to see their governments make policy changes. This mobilization raises the costs of maintaining initially shallow commitments and pushes governments to deepen them. Therefore, governments should be most likely to deepen initially shallow commitments under conditions that facilitate such mobilization.

Building on previous research that links domestic political mobilization to international cooperation, I focus on two factors that facilitate and encourage mobilization: democratic political institutions and influxes of information and assistance from international actors. First, I argue that democratic states will be more likely to deepen initially shallow commitments compared to nondemocracies because freedom of association, freedom of expression, and freedom of the press allow aspiring activists in these countries to meet and organize openly. In nondemocratic states, by contrast, additional conditions must be present for effective mobilization to emerge. I argue that when international attention is focused on a particular state’s implementation of the treaty, domestic groups receive much-needed assistance from international actors. Such aid can come in the form of information, which helps domestic groups identify deficiencies in government policies and gather support for their initiatives, as well as direct assistance, which helps domestic groups learn how to mobilize effectively. Groups operating in nondemocratic states stand to gain the most from this process because they are the most in need. Thus, nondemocratic states should be more likely
to deepen initially shallow commitments following periods of intense international attention toward their practices. There should be no relationship between international attention and deepening of commitments among democracies because these countries already have a high level of mobilization that is therefore less responsive to fluctuations in international scrutiny.

I test these expectations in the context of human rights, where states often use reservations to make their commitments more shallow by legally opting out of specific provisions. Using originally-collected data on the Convention on the Elimination of all Forms of Discrimination Against Women (CEDAW), I find that democracies are more likely to withdraw reservations compared to nondemocracies, and the likelihood of withdrawal among nondemocracies increases following consideration of the country’s periodic report. These results are robust to model specifications that account for factors that influence states’ propensities to enter reservations upon ratification.

This research has important implications for institutional design and international cooperation in human rights as well as other issue areas. Criticism of reservations implies a better path would have been to forbid reservations and require states to choose between committing to the full treaty or remaining outside the regime altogether, in the hopes that international and domestic pressure would eventually push hesitant states to accept the treaty in full. While we cannot compare CEDAW to the counterfactual regime in which reservations are not permitted, the extensive use of this flexibility mechanism suggests that allowing reservations enables states to enter and interact with the treaty regime sooner than they otherwise would. This paper shows that these initially shallow commitments can deepen over time, and it identifies the conditions under which this process is likely to occur. Hence, this study offers some solace to observers who worry that reservations strip human rights treaties of their bite. Deepening of initially shallow commitments is never guaranteed, but the data collected for this paper shows withdrawing reservations happens more often than crit-
ics would expect, and domestic actors working in conjunction with an international institution can hasten this process.

In the following sections, I discuss the role of flexibility in international agreements, present my argument for the connection between domestic mobilization and deepening commitments, and identify the conditions under which such mobilization is likely to emerge. Next, I present a research design to test my hypotheses using data from CEDAW and discuss the results from the analysis. Finally, I show that the results are robust to modeling techniques designed to account for selection into the sample of states with reservations. The paper concludes with a discussion of implications.

2.1 Flexibility and International Cooperation

Flexibility mechanisms present a persistent trade-off: On the one hand, flexibility provisions should help a treaty appeal to states that harbor serious concerns about the expected benefits of cooperation or costs of compliance (Bernauer et al., 2013; Swaine, 2006). On the other hand, too much flexibility may undermine cooperation by allowing states to flout international obligations or refuse to apply critical portions of a treaty. Consideration of this trade-off has spurred a productive research program that seeks to understand the merits and pitfalls of flexible treaty design. Scholars generally find that flexibility goes hand in hand with depth or legalization (von Stein, 2008; Kucik and Reinhardt, 2008; Baccini, Dür and Elsig, 2015), implying that flexible agreements enable states to commit to more stringent obligations than would otherwise be possible. Focusing on which states are most likely to use flexibility mechanisms and how, previous research suggests democracies and states wary of domestic enforcement of treaty provisions are most inclined to use flexibility mechanisms both during and after the ratification stage (Hafner-Burton, Helfer and Fariss, 2011; Simmons, 2009; Hill Jr., 2016; McKibben and Western, 2018).

Throughout this paper, I use the term “flexibility mechanisms” to refer to treaty
design features that allow states to unilaterally set the terms of their commitments. In other words, flexibility allows states to control the depth of their commitments, or the extent to which a treaty requires modification of existing practices. This definition is similar to Barbara Koremenos’s explanation of flexibility as pertaining to two questions: (1) How easy is it to modify rules, procedures, or obligations under the agreement? and (2) How easily can different positions be accommodated? (Koremenos, 2016, 41). These questions yield a broad definition of flexibility that encompasses five design features: duration provisions, which identify an end to the treaty’s legal status; escape clauses, which allow states to temporarily cease compliance without being considered in violation; withdrawal clauses, which allow states to disengage from the treaty at will; imprecision of substantive terms, which enables states to interpret vague obligations as they see fit; and reservations, which allow states to opt out of specific provisions. My definition omits duration provisions, which are “collective” rather than “unilateral,” because their use requires a collective decision by all treaty drafters or participants (Helfer, 2013, 178).

Previous research suggests flexibility alleviates states’ concerns about committing to restrictive treaties and encourages them to make deeper commitments than they would otherwise be willing to accept. Focusing on international agreements that seek to mitigate climate change, von Stein (2008) argues that the ability to create carbon sinks, or gain emissions credit by participating in activities that remove carbon dioxide from the atmosphere, was a flexibility mechanism that lowered the cost of commitment for states and led them to ratify the Kyoto Protocol more quickly than they would have done otherwise. Since the Kyoto Protocol involved greater legal obligation, precision of substantive terms, and delegation of monitoring powers to existing institutions compared to the United Nations Framework Convention on Climate Change (FCCC), von Stein (2008) concludes that flexibility enables treaty drafters to create more legalized institutions without compromising participation. In the context
of the international trade regime, Kucik and Reinhardt (2008) provide evidence that
the ability to utilize antidumping to temporarily impose duties on imports without
violating obligations within the World Trade Organization (WTO) encourages states
to join the trade regime and both agree to and implement lower tariffs.

The above examples are unique, however, in that states vary in their ability to
utilize the identified flexibility mechanisms. Due to the existence of appropriate
bureaucratic apparatuses, some states could expect to gain more carbon sink credits
or be better able to meet antidumping standards than others. Yet many flexibility
mechanisms, such as reservations, interpretation of imprecise language, or escape and
withdrawal clauses whose use does not require substantial bureaucratic capacity, are
open to all states parties. Indeed, scholars make similar claims that these flexibility
devices can also encourage deeper commitments (e.g., Swaine, 2006). Yet here, the
relevant question is not necessarily how flexibility mechanisms affect the behavior of
states who can use them but rather which states are most likely to use flexibility
mechanisms and how.

Existing research in this area points to two characteristics of domestic institutions
that are associated with the use of flexibility mechanisms: democracy and capacity
for enforcement. Hafner-Burton, Helfer and Fariss (2011) find that democracies and
countries in which courts can review and invalidate executive actions are more likely to
make use of derogations clauses in human rights treaties, which allow states to legally
suspend respect for individual rights during emergencies. Similarly, studies of reser-
vations to human rights treaties find that reservations are most common among states
whose existing legal standards diverge significantly from those mandated by the treaty
and whose judiciaries are powerful enough to make enforcement of those obligations
most likely (Simmons, 2009; Hill Jr., 2016; McKibben and Western, 2018). Looking
specifically at the strength of language used in the 1991 Paris Principles on the De-
sign of National Human Rights Institutions (NHRIs), Linos and Pegram (2016b) find
that states were more likely to institute strongly-worded recommendations compared to more flexible, weakly-worded recommendations, and this discrepancy was sharper for nondemocratic states, who face pressure to demonstrate progress on improving respect for human rights but fear delegating too much power to monitoring bodies. These studies suggest that domestic politics is crucial to understanding the extent to which flexibility mechanisms that encourage participation influence cooperation.

Still, the question of how the use of flexibility mechanisms changes over time remains unanswered. Hafner-Burton, Heller and Fariss (2011) show that derogating once is a strong predictor of future derogations, which suggests some states, especially those that cannot be held accountable domestically for violating human rights, may be abusing flexibility mechanisms to legally flout international obligations. Do these findings carry over to other types of flexibility mechanisms? If states use flexibility mechanisms to make initially shallow commitments, can these commitments deepen over time? In the following section, I present a theory to explain when and how this process can occur.

2.2 Domestic Politics and Deepening Commitments

I argue that initially shallow commitments deepen over time when domestic actors mobilize to pressure a government to broaden the scope of its commitments. Mobilization is most likely to emerge in democracies, where respect for freedom of association, freedom of expression, and freedom of the press facilitates the development of effective non-governmental organizations (NGOs) and other domestic groups. Nondemocratic regimes generally do not support the formation of powerful, independent domestic groups, but the prospects for effective mobilization are enhanced when international actors focus their attention on specific states parties. Still, the marginal effect of international assistance should be negligible in democratic regimes, because domestic groups in these environments already enjoy favorable conditions for mobilization.
Hence, democracies should be more likely to deepen initially shallow commitments compared to nondemocracies, and nondemocracies become more likely to deepen their commitments following intense international assistance to domestic actors.

The idea that domestic politics drives state behavior with respect to international obligations finds broad empirical support in existing studies of international cooperation, especially in issue areas such as human rights and environmental politics, where traditional mechanisms for inducing compliance, such as reciprocity, retaliation, and reputation (Guzman 2008) are less potent. With regards to human rights, Simmons (2009) finds that treaties have their greatest effects on partially democratic states, where citizens have much to gain from securing better respect for human rights but do not face crippling obstacles to mobilization. In the environmental context, Dai (2007) finds that governments exhibit higher levels of compliance with international treaties when pro-compliance domestic groups have greater electoral leverage and informational capacity. Even in issue areas such as trade, where strategies of reciprocity are available, domestic politics can make a difference. States where leaders can expect to be swiftly removed from power if they violate international commitments sustain higher levels of trade compared to states where replacing a leader is more difficult (McGillivray and Smith 2008).

Democratic institutions support the development of effective political mobilization, because they allow aspiring activists to meet and disseminate information freely. Indeed, indices of democracy from Freedom House and the Varieties of Democracy project include freedom of expression, press freedom, and the lack of restrictions on the operation of civil society organizations as core indicators of democratic governance (Freedom House 2017, Coppedge et al. 2018). Although Simmons (2009) argues that mobilization is less likely in stable democracies, because citizens in these regimes lack the motivation to demand greater respect for human rights, the motivation for mobilization exists where states’ commitments are deliberately made more
shallow than others’. For instance, when states enter reservations to consciously exempt themselves from respecting internationally-recognized human rights, citizens in all countries where reservations exist have reasons to demand their withdrawal, regardless of regime type. In the context of the Paris Agreement, citizens concerned about climate change in states that committed only to weak, largely symbolic attempts to cut emissions have much to gain from pressing for deeper commitments.

While groups opposing deeper commitments can also benefit from the openness of democratic societies, groups supporting deeper commitments can use the fact that the international community has spoken in favor of their goal to their advantage. Survey experiments consistently show that overall, citizens are more disapproving of government actions that violate international law, which suggests citizens are more inclined to favor policies supported by the international community when these conflict with their own government’s (Wallace 2013; Chaudoin 2014; Kreps and Wallace 2016). This research implies that domestic audiences are generally sympathetic to the argument that a government should change its policies to better comport with the demands of the international community. Even though flexibility mechanisms legally allow states to make shallow commitments if they so desire, the international community clearly prefers deeper commitments overall. In fact, members of the international community often resist governments’ use of legal flexibility mechanisms. For instance, human rights treaty bodies often question whether governments’ reservations and derogations are justified, and the WTO Dispute Settlement Body can rule on whether a state’s antidumping decisions are warranted (Kucik and Reinhardt 2008; Hafner-Burton, Helfer and Fariss 2011).

Compared to groups operating in democratic societies, domestic actors in nondemocratic societies who wish to press their government to deepen their commitments face additional obstacles that generally hinder the development of effective political movements. Still, citizens in nondemocratic countries are not helpless, for
domestic advocates can and often do turn to international actors for assistance (Keck and Sikkink 1998). International resources, which ebb and flow over time as states interact with a treaty regime, are especially useful to domestic actors in non-democratic states who face organizational challenges and struggle to effectively relay their demands to the government.

International resources that strengthen political mobilization include information and direct assistance. Dai (2007) proposes a general theory of how international institutions can enforce treaty obligations by publishing information on state behavior, which augments the power of pro-compliance domestic groups by enhancing their ability to identify deficiencies in government policy and mobilize support for change. Although Dai (2007) tests this theory in the realm of environmental politics, the idea that international institutions can provide information to domestic audiences applies to other issue areas as well. For example, the human rights treaty regime produces information through a process of periodic review, in which states parties submit public reports about their compliance with a treaty to the committee that monitors its implementation (O’Flaherty 2006). In the trade regime, initiating a formal dispute against another state’s trade practices alerts domestic audiences to potential treaty violations, an event that is often newsworthy and can motivate citizens to learn more about trade policy (Pelc 2013; Chaudoin 2014).

During times of increased information provision, domestic actors who wish to pressure their government to deepen its commitments can also receive direct assistance from international actors who share the same goals. For example, international NGOs advocating for human rights actively work to support their domestic counterparts and focus their attention on NGOs operating in the countries whose periodic reports are under consideration, offering them assistance in preparing alternative “shadow” reports and providing funds and training for their appearance before the relevant committee (Zwingel 2005). Likewise, Dai (2007) notes that one way in which the
Convention on Long-range Transboundary Air Pollution enhanced the power of pro-compliance domestic groups was by creating a venue for environmental activists to share information and strategies cross-nationally and help one another by spreading best practices. Because information provision does not occur constantly but rather follows a schedule set by the institution or other states parties, mobilization should be strongest and therefore most likely to influence government policy directly following periods of intense information production.

This information, together with the concomitant international assistance, is most valuable to domestic groups operating in nondemocratic regimes, because these groups are most in need of aid. Potent mobilization emerges on its own in democratic societies. While information can still aid the activities of groups operating in democracies, we should not expect to observe policy change following influxes of international resources in these states, because the marginal effect of such aid is minimal. By contrast, in nondemocracies, new information and the international assistance that comes along with it is more likely to enhance the strength of domestic political mobilization and increase the likelihood that domestic actors will succeed in their efforts to pressure a government to deepen initially shallow commitments. This argument, therefore, predicts that democratic states will be more likely to deepen initially shallow commitments compared to nondemocratic states, while nondemocratic states will be more likely to deepen initially shallow commitments following periods of intense international attention toward their practices.

2.3 Applying the Theory to the Human Rights Regime

I test my predictions in the context of human rights, where reservations enable states to opt out of specific treaty provisions. Reservations allow states to unilaterally limit their obligations under the terms of the treaty. In other words, states can use reservations to make their commitments more shallow, meaning that the treaty
requires less modification of existing policy than it would otherwise. The use of reservations is open to all states parties and does not require a specific bureaucratic apparatus. Once party to the treaty, a state can further modify the depth of its commitment by withdrawing reservations or adding new ones, although the legality of the latter procedure is disputed.

While few agreements prohibit reservations, their use is far more common in human rights than in other issue areas (Koremenos 2016). The Vienna Convention on the Law of Treaties allows states to make unilateral statements to “exclude or to modify the legal effect” of parts of a treaty (United Nations 1980). States may enter reservations to any provision, provided the reservation is not expressly prohibited, omitted from the treaty’s list of permissible reservations, or contradictory to the treaty’s “object and purpose” (United Nations 1980). Although other states may object to a signatory’s reservations, this practice has little direct effect in the human rights context, where states cannot use reciprocity to motivate compliance (Neumayer 2007). Moreover, the stipulation that reservations must not contradict the “object and purpose” of a treaty is the subject of much debate among legal scholars, who disagree as to the proper interpretation of the phrase as well as the appropriate course of action if a reservation is deemed unacceptable (Goodman 2002; Swaine 2006). Despite some efforts to formally judge the validity of reservations and sever those labeled impermissible, these policies are rarely utilized.

Focusing on reservations is an appropriate way to test the theory outlined above, because reservations provide a simple way for states to formally alter the depth of their commitments over time. According to a strict interpretation of the Vienna Convention, states may only enter reservations upon ratification. In practice, however, some states enter so-called “late” reservations after ratification or denounce the treaty and re-accede with new reservations (Helfer 2006). Hence, withdrawing reservations indicates deepening of existing commitments, while adding new ones indicates that
commitments are being made more shallow.

My empirical analysis focuses specifically on the Convention on the Elimination of all Forms of Discrimination Against Women (CEDAW), which entered into force in 1981. Two qualities of CEDAW make it an appropriate case study and a hard test for my argument. First, states have entered more reservations to CEDAW than any other major human rights treaty (Tiefenbrun, 2012). The prevalence of reservations enables governments under pressure for withdrawal to argue that these caveats are a necessary means of adjusting the treaty’s requirements to accommodate different cultures. Second, many states offer rationales for reservations to CEDAW based on alleged conflicts with religious law or state constitutions that invoke religious law (Freeman, 2009). Therefore, governments who seek to resist pressure for withdrawal can justify their decisions in religious or cultural terms (Kabasakal Arat, 2003) and may find powerful political allies among religious authorities who oppose the norms promoted in the treaty. Hence, pressuring states to withdraw reservations from CEDAW should be especially difficult, and evidence that domestic mobilization makes a difference would indicate this mechanism is likely to work similarly for other treaties.

Adapting the theory’s predictions to this context is fairly straightforward. First, I expect that democratic states will be more likely to withdraw reservations compared to nondemocratic states. As discussed earlier, multiple features of democratic governance, such as respect for freedom of association and freedom of the press, support the development of domestic mobilization in favor of deepening treaty commitments. Therefore, democracy serves as a proxy for the strength of mobilization in favor of women’s rights. While a more direct measure of domestic political activity would be a more precise test of the connection between domestic mobilization and withdrawing reservations, comprehensive, cross-national data on this topic is currently unavailable. Still, there is some evidence to suggest that democracy is indeed correlated with the
strength of autonomous feminist movements, which would be most likely to advocate for deeper commitments to CEDAW (Htun and Weldon 2012).

For the second prediction, I define intense international attention as the consideration of a state’s periodic report by a treaty body. As discussed in the theory section, the process of periodic review produces important information that domestic actors can use to mobilize support and press their claims, and international actors often direct aid toward countries whose reports are under consideration. There is also notable anecdotal evidence that women’s movements, aided by the reporting procedure, were pivotal in the withdrawal of reservations to CEDAW, especially in nondemocratic countries. For example, in Singapore, the Association of Women for Action and Research (AWARE) submitted shadow reports to the CEDAW Committee each time the government became subject to periodic review. This organization also played a critical role in laying the groundwork for a constitutional amendment that allowed women to transmit citizenship to children born outside the country, which paved the way for Singapore to remove its reservation to Article 9 of CEDAW in 2007 (Tan 2008). Following additional shadow reports from AWARE and other groups that recommended replacing blanket reservations to Articles 2 and 16 with partial reservations that pertain only to specific provisions, the Singaporean government did precisely that in July 2011, withdrawing six reservations that it had maintained since 1995. This case illustrates how the reporting procedure can help catalyze domestic mobilization in nondemocratic settings. This discussion, therefore, leads to the following hypotheses:

**H1:** Democratic states are more likely to withdraw reservations compared to nondemocratic states.

**H2:** Nondemocratic states are more likely to withdraw reservations after their reports are considered by a treaty body.
2.4 Withdrawing Reservations to CEDAW, 1981-2015

This section describes the statistical approach to evaluating my hypotheses. To conduct the analysis, I utilize an original dataset of state behavior with respect to reservations to CEDAW (United Nations, 2017). The data cover all states parties to CEDAW with a population greater than 500,000 that entered at least one reservation upon ratification that qualifies the application of individual treaty provisions or the treaty as a whole. Although states could, in principle, add reservations after ratification, none have done so in the context of CEDAW.

I collect data on reservations and their withdrawal from the UN Treaty Collection (UNTC) (United Nations, 2017). This database records all state actions pertaining to the treaty and registered with the UN, including ratification or accession and reservations and/or declarations, as well as modifications and withdrawals of these statements. While declarations are technically distinct from reservations, some declarations nevertheless amount to reservations in accordance with the Vienna Convention, for they do in practice limit the treaty’s legal effects (Neumayer, 2007). For example, Pakistan entered a declaration upon accession to CEDAW in 1996 that states: “The accession...is subject to the provisions of the Constitution of the Islamic Republic of Pakistan.” With this declaration, Pakistan is effectively reserving its right to refrain from implementing provisions that can be argued to conflict with its constitution. I follow Neumayer (2007), Kearney and Powers (2011) and Hill Jr. (2016) in counting these types of declarations as reservations.

I also count as reservations declarations that restrict the interpretation of specific phrases, which in practice limit the treaty’s legal effect by limiting the range of mean-

---

1 I omit reservations to Article 29, paragraph 1, which calls for disputes between states parties to be submitted to arbitration, because this reservation is expressly permitted by the treaty. Therefore, there are little grounds for criticizing a government’s decision to maintain it.

2 The UN glossary of terms relating to treaty actions distinguishes between reservations and declarations in that the latter “merely clarify the state’s position and do not purport to exclude or modify the legal effect of a treaty” (United Nations, N.d.).
ing that can be derived from its text. For instance, a declaration entered by Qatar upon accession in 2009 states that “the question of the modification of “patterns” referred to in article 5 (a) must not be understood as encouraging women to abandon their role as mothers and their role in child-rearing, thereby undermining the structure of the family.” I consider this type of declaration to have a similar effect as that of a reservation and therefore count it as such. In contrast, I do not count as reservations declarations that have no meaningful impact on the treaty’s legal effect. For example, Australia entered the following declaration to CEDAW upon ratification in 1983: “Australia has a Federal Constitutional System in which Legislative, Executive and Judicial Powers are shared or distributed between the Commonwealth and the Constituent States. The implementation of the Treaty throughout Australia will be effected by the Commonwealth State and Territory Authorities having regard to their respective constitutional powers and arrangements concerning their exercise.” Such a declaration does not alter the treaty’s legal effect and does not constitute a reservation.

My coding rules depart from previous studies in that I disaggregate reservations by provision rather than article. Earlier studies count multiple reservations against a single article as one reservation and often frame the data in terms of the number of articles affected by reservations (Neumayer 2007; Kearney and Powers 2011; Simmons 2009; Hill Jr. 2016). Collapsing multiple reservations aimed at separate provisions of an article into one reservation makes little sense for this study, because states do not enter and withdraw reservations in this manner. Consider the case of Singapore, which entered reservations to Articles 2 and 16 of CEDAW upon accession in 1995. Article 2 includes seven provisions, and Article 16 includes nine. In 2011, Singapore withdrew its reservations with regards to six provisions: Article 2, paragraph (g), and Article 16, paragraphs 1(b) and 1(d) through 1(g). Reservations to the remaining 10 provisions remain intact. I therefore code these as 16 separate reservations, six of
which were withdrawn in 2011.

The vast majority of reservations are sufficiently precise so as to permit coding the affected provision(s) for each separate reservation. However, a small portion of reservations do not explicitly refer to a specific section of the treaty’s text. The declaration from Pakistan mentioned earlier is a relevant case in point. Neumayer (2007) proposes two strategies to deal with this issue: (1) counting these general reservations as reservations against all articles of the treaty, or (2) treating countries with general reservations as if they had not ratified the treaty at all. The logic behind these two approaches is that a general reservation may potentially alter the legal effect of any and all articles, which some would argue renders meaningless the state’s status as a party to the treaty.

Since my intent is to investigate under what conditions states parties withdraw reservations, neither strategy is particularly satisfying. It is difficult to determine whether a state entered a general reservation because it intended to limit the effect of all the treaty’s provisions or because it was unable to identify all the relevant provisions and concluded a general reservation would better express its wishes. In some cases, states enter general reservations because the treaty does not explicitly mention a specific concern about which it feels strongly. For example, New Zealand entered a reservation that states it reserves “the right not to apply the provisions of CEDAW in so far as they are inconsistent with policies relating to recruitment into for service in: (a) the Armed Forces...or (b) the law enforcement forces....” I do not purport to pass judgment upon which provisions may be relevant for these types of reservations. Hence, I regard such general reservations as a separate type of statement conceptually distinct from reservations that name specific articles or provisions, and I count each as a single reservation.

I code a reservation as withdrawn if the state has formally submitted to the UN a document specifying withdrawal of the reservation with respect to the specific pro-
vision or, in the case of a general reservation, the entirety of the reservation. Partial withdrawals in which a state modifies the wording of a reservation but does not withdraw it are not coded as withdrawals. In the rare event that a state enters a general reservation and then modifies it to apply to specific provisions, the general reservation is considered withdrawn, and the outlined provisions are recorded as separate reservations that remain intact. For example, Mauritania entered a general reservation upon accession in 2001. In 2014, the government of Mauritania informed the UN that it “partially withdraws its general reservation made upon accession, which shall continue to apply in respect of articles 13 (a) and 16 of the Convention.” In this case, Mauritania is coded as having entered eleven reservations: one general, one with respect to Article 13 (a), and one with respect to each of the nine provisions of Article 16. The first is coded as withdrawn in 2014, and the rest are coded as not withdrawn.

Based on these coding rules, I compile a list of 384 reservations entered by 46 countries. In the analyses presented below, the unit of analysis is country-year, with the observations for each state beginning with the year it ratified CEDAW and ending with the year it withdrew its last reservation or the end of the observation period, 2015. The dependent variable is coded 1 for each year in which the state withdraws at least one reservation and 0 otherwise. Thirty-three (71.7%) of the 46 countries withdraw some or all reservations during the observation period in 51 separate instances of withdrawal.

2.4.1 Independent Variables and Model Specification

To test my hypotheses, I estimate multilevel logistic regressions with random intercepts by country. This specification allows the observations for each country to be correlated with one another and assume a unique intercept based on country-level

---

3The full list of countries and reservations is provided in the appendix.
characteristics not captured by the variables in the model. I also address the stochastic dependence in the country-year dataset with the inclusion of time polynomials to emulate a duration model (Carter and Signorino 2010). Following Carter and Signorino (2010), I utilize a cubic polynomial of the number of years since ratification, or, for countries that withdrew a subset of their reservations, the number of years since the last withdrawal, because this functional form can capture hazard shapes common to parametric and semiparametric duration models without overfitting the data.

The first main independent variable of interest is democracy. I operationalize this variable as a dichotomous indicator because the theory describes a number of features of democratic governance, such as respect for freedom of association and freedom of the press, that work in concert to increase the likelihood of domestic mobilization. For this variable, I utilize the Revised Combined Polity Score (polity2) from the Polity IV project, a 21-point score that ranges from -10 (strongly autocratic) to +10 (strongly democratic) (Marshall, Gurr and Jaggers 2016). I code a state as democratic if its polity2 score in the previous year was greater than or equal to 6. Just over half (56.5%) of the observations in the dataset are coded as democratic.

The second independent variable codes whether the state’s report to the CEDAW Committee was considered in the previous year. Information from this variable is retrieved from the Committee’s website (UN Human Rights Office of the High Commissioner 2018). Almost all countries in the dataset have their reports considered at least once during the time they maintain reservations. This variable is coded 1 in 12.4% of the country-years in the sample. To test the second hypothesis, I include an interaction between this measure and the democracy variable.

I also control for variables that may impact the intensity of domestic mobilization or may be correlated with democracy and also influence a country’s propensity to withdraw reservations. First, I include a variable that captures changes in the govern-
ment’s level of respect for women’s rights since ratification. If reservations are indeed ways for states to commit to treaties without changing their practices, then it stands to reason that states would withdraw reservations when their level of respect for human rights improves such that their behavior can be defined as compliance without the reservations. Likewise, NGOs who notice improvements in women’s rights may sense that efforts to pressure states to withdraw reservations may be more successful and will therefore ramp up their activities.

To measure respect for women’s rights, I utilize two data sources. First, I use data from the Cingranelli-Richards (CIRI) Human Rights Dataset (Cingranelli and Richards, 2010). Following its use in previous studies (Cole, 2013; Murdie and Peksen, 2015), I combine three indicators of women’s social, political, and economic rights into a single index ranging from zero to nine. These indicators reflect the extent to which a variety of women’s rights, such as equality in choice of employment, the right to vote and participate in politics, and the right to own property are guaranteed by law and enforced in practice. I operationalize this variable as the lagged change in this index since ratification for each country-year. In other words, I subtract the index for women’s rights for each country in the previous year from its score at ratification. Theoretically, this variable can range from -9 to 9. In practice, it ranges from -5 to 4.

Because the CIRI data continue only until 2004, I employ a secondary measure that encompasses a narrower range of rights but covers the full observation period. This measure is the women civil liberties index from the Varieties of Democracy project, which ranges from 0 to 1 and captures the extent to which women enjoy freedom of domestic movement, the right to private property, freedom from forced labor, and access to justice (Coppedge et al., 2018). As with the CIRI index, I subtract each country’s women civil liberties index from its score at ratification. Within the dataset, this variable ranges from -0.379 to 0.32.

Next, I control for whether the country recently experienced a leadership transition
in which the new leader draws primary support from a societal group different from that of her predecessor (Mattes, Leeds and Matsumura, 2016). When new leaders come to power, domestic groups may intensify their efforts to pressure the government to change its policies. Likewise, new leaders who rely on different societal groups may be eager to enact policy changes to distinguish themselves from the previous government. Thus, countries should be more likely to withdraw reservations once the new leader enters into power. This concept is operationalized by a dummy variable that indicates whether the country experienced at least one leadership transition of this type that lasted longer than 30 days during the previous year. The variable comes from the CHISOLS (Change in Source of Leader Support) dataset by Mattes, Leeds and Matsumura (2016), which provides data up to 2008.

I also control for the total number of reservations entered by the state upon ratification. International actors may direct their resources toward countries with the highest numbers of reservations to maximize their impact. A large number of reservations, however, suggests a high level of conflict between the Convention and domestic legislation, which could make the government more resistant to pressure for withdrawal. To create this variable, I count the number of provisions affected by each state’s reservations. Roughly one-third of the states in the dataset entered under five reservations, while the remaining two thirds entered numbers ranging from five to 24.

Finally, I control for the number of recorded withdrawal events since the treaty entered into force. Calls to withdraw reservations may grow stronger as more states withdraw reservations over time, and governments may find it more difficult to justify their reservations as others withdraw. Therefore, I expect to observe a positive relationship between the total number of withdrawal events as of the previous year and states’ propensities to withdraw reservations. In the dataset, this variable ranges from 0 to 49.
2.5 Results

The results of the multilevel logistic analysis are reported in Table 2.1. In Models 1-3, the coefficient for democracy is positive and statistically distinguishable from zero. In other words, democratic states are more likely to withdraw reservations compared to nondemocratic states, even after controlling for whether the state’s periodic report has recently been considered. In fact, the coefficient for the report variable is positive but only reaches conventional levels of statistical significance in Model 2, when the temporal range is restricted to 1982-2005. A more nuanced picture emerges, however, once the interaction term is included. Models 4-6 suggest there is a positive association between the reporting procedure and withdrawing reservations, but this relationship only obtains for nondemocratic states. These findings lend support for both hypotheses.

Turning to the controls, the only variable that reaches conventional levels of statistical significance fairly consistently is the count of the number of reservations the state entered upon ratification. There are two possible explanations for this result. On the one hand, NGOs may focus their efforts on the most serious offenders, leading to more withdrawals among these states. On a more pessimistic note, states with a large number of reservations may be more willing to withdraw a select few to placate critics and gain the benefits of withdrawal without making substantial changes to domestic laws. Distinguishing between these two possibilities would require a deeper examination of the rationale behind individual reservations and the process of their withdrawal, which could prove a fruitful avenue for future research.

By contrast, there does not appear to be a statistical relationship between withdrawing reservations and recent leadership transitions that involve changes in the sources of leader support (SOLS). The coefficient for the SOLS change variable is positive in Models 2-3 and 5-6 but not statistically significant. The coefficient for
Table 2.1

| Dependent variable: Withdrawing at least one reservation |
|-----------|-------------|-------------|-------------|-------------|-------------|
|           | (1)         | (2)         | (3)         | (4)         | (5)         | (6)         |
| Democracy\(_{(t-1)}\) | 0.894**      | 1.050*      | 0.861**      | 1.208***    | 2.023**     | 1.243**     |
|           | (0.357)     | (0.552)     | (0.420)     | (0.402)     | (0.796)     | (0.501)     |
| Report Considered\(_{(t-1)}\) | 0.352        | 1.087**     | 0.634        | 1.382**     | 3.309***    | 1.678**     |
|           | (0.388)     | (0.456)     | (0.426)     | (0.585)     | (0.935)     | (0.693)     |
| Democracy\(_{(t-1)}\) * Report\(_{(t-1)}\) | -1.619**     | -2.996***   | -1.578*     |             |             |             |
|           | (0.806)     | (1.106)     | (0.895)     |             |             |             |
| Initial Number of Reservations | 0.067***     | 0.043       | 0.065**     | 0.065***    | 0.041       | 0.064**     |
|           | (0.024)     | (0.030)     | (0.026)     | (0.023)     | (0.031)     | (0.026)     |
| SOLS Change\(_{(t-1)}\) | 0.400        | 0.160       | 0.367        | 0.367       | 0.155       |             |
|           | (0.546)     | (0.521)     | (0.544)     | (0.521)     |             |             |
| \(\Delta\) Women’s Rights\(_{(t-1)}\) (CIRI) | -0.047       |             | -0.017      |             |             |             |
|           | (0.144)     |             |             |             |             |             |
| \(\Delta\) Women’s Rights\(_{(t-1)}\) (V-Dem) | 5.466*       |             |             |             |             | 5.105*      |
|           | (2.966)     |             |             |             |             | (2.967)     |
| Previous Withdrawals\(_{(t-1)}\) | -0.014       | -0.012      | -0.020       | -0.015      | -0.011      | -0.020      |
|           | (0.012)     | (0.024)     | (0.017)     | (0.012)     | (0.025)     | (0.017)     |
| Time | 0.045        | -0.767**     | -0.213       | 0.057       | -0.864**    | -0.222      |
|           | (0.181)     | (0.370)     | (0.246)     | (0.183)     | (0.375)     | (0.246)     |
| Time\(^2\) | 0.007        | 0.101**      | 0.030        | 0.005       | 0.110**     | 0.030       |
|           | (0.014)     | (0.044)     | (0.023)     | (0.014)     | (0.044)     | (0.023)     |
| Time\(^3\) | -0.0002      | -0.003**     | -0.001       | -0.0002     | -0.004**    | -0.001      |
|           | (0.0003)    | (0.001)     | (0.001)     | (0.0003)    | (0.001)     | (0.001)     |
|           | (0.807)     | (1.103)     | (0.909)     | (0.822)     | (1.231)     | (0.938)     |
| Countries | 46          | 41          | 44          | 45          | 41          | 44          |
| Withdrawals | 51          | 31          | 39          | 50          | 31          | 39          |
| Observations | 893         | 514         | 676         | 874         | 514         | 676         |

\(^*p<0.1; **p<0.05; ***p<0.01\)

Standard errors in parentheses

the count of previous withdrawals also fails to reach conventional levels of statistical significance across all six models.

Meanwhile, the two variables for change in the country’s level of respect for women’s rights since ratification offer only weak, somewhat mixed evidence that improvements in these measures correlate with withdrawing reservations.\(^4\) The coeffi-

\(^4\)These results hold for alternative codings of this variable, including the objective level of respect for women’s rights and the annual change in respect for women’s rights. See appendix for these results.
Fig. 2.1: The above graph shows the distributions of simulated values for the average predicted probability of withdrawing reservations for democratic states (black) and nondemocratic states (grey).

Coint for change in the women’s civil liberties index is positive and statistically significant at the 0.1 level, but the coefficient for change in the CIRI measure for women’s economic, political, and social rights is slightly negative with a wide standard error. At first glance, these results are somewhat surprising, given the overt connection between women’s rights and CEDAW. Yet this null finding comports with anecdotal evidence from campaigns seeking full implementation of CEDAW. For instance, Morocco, Algeria, Egypt, and Singapore all withdrew select reservations following legislative changes that brought state practice closer to compliance with CEDAW (Byrnes and Freeman, 2011; United Nations, 2011). However, in Malaysia and Bangladesh, NGOs complain that despite withdrawing a number of reservations, the government has yet to change laws or policies to meet these higher standards (WAO, 2012; CIC-BD, 2015). These examples suggest that the pathway from entering reservations to withdrawing them does not always flow through human rights improvements. Since activists often advocate for both withdrawal and behavioral change, the sequencing of these developments varies from case to case.
The substantive effect of democracy, based on estimates from Model 3, is displayed in Figure 2.1. Because the dependent variable is binary, I calculate predicted probabilities using the observed value approach recommended by Hanmer and Kalkan (2013). Rather than setting the values of all other covariates to their means and modes as is commonly done in the “average case” approach, the observed value approach involves calculating a predicted probability for each individual case given the observed values of the other variables and averaging over all the cases (Hanmer and Kalkan, 2013, 264). Despite the overlap in the distributions of simulated values, the two groups are statistically distinguishable ($p < 0.05$). The average predicted probability of withdrawal is nearly twice as large for democracies (10.75%) compared to nondemocracies (5.54%). In other words, the likelihood of withdrawing reservations is higher for democratic states than nondemocratic states, a finding in line with the expectations outlined in H1.

Figure 2.2 illustrates the substantive impact of report consideration among democratic and nondemocratic states, based on estimates from Model 6. The graph on the left shows how the average predicted probability of withdrawing reservations increases for nondemocratic states in the year after the state undergoes review of its periodic report. Among nondemocratic states, undergoing review increases the predicted probability of withdrawing reservations by 284% (from 4.05% to 15.5%). Again, the distributions are statistically distinguishable from one another ($p < 0.05$). The graph on the right, however, shows no discernible difference in the predicted likelihood of withdrawal among democratic states, as predicted by H2.

Overall, the evidence offers support for both hypotheses. Democracy is positively related to patterns of reservation withdrawal, and report consideration is associated with reservation withdrawal among nondemocracies. In the next section, I show that these findings are also robust to model specifications that account for selection into the sample of countries with reservations.
Figure 2.2: The above graphs show the impact of report consideration by democracy. The left graph depicts the distributions of simulated values for the average predicted probability of withdrawing reservations for nondemocratic states whose reports are considered in the previous year (black) and nondemocratic states whose reports are not considered (gray). The right graph shows the same series of distributions for democratic states.

2.6 Robustness Check: Heckman Probit Models

In order to withdraw reservations, states must first “select into” the sample by entering at least one reservation upon ratification. If democracies are more likely to enter reservations than nondemocracies, then observing a higher propensity for withdrawal among democratic states could simply reflect their ubiquity in the dataset rather than the effect of domestic mobilization processes. To address this concern, I conduct additional analyses that account for this process. The substantive results from the main analysis are identical.

To identify variables for the selection equation that satisfy the exclusion restriction and do not directly affect the decision to withdraw reservations, I exploit the temporal
features of the process of entering and withdrawing reservations. Deciding to withdraw a reservation implies that some factor that necessitated its entry has changed. Therefore, time-invariant state characteristics that encourage states to enter reservations should not directly impact patterns of withdrawal. One such time-invariant feature is the state’s legal tradition. Simmons (2009) argues that states with a common law legal system are more likely to enter reservations, because judiciaries have more leeway in interpreting treaty provisions and greater power to hold governments accountable. She also finds that Muslim countries more readily enter reservations, perhaps in an attempt to reconcile the treaty with their own understanding of justice (Simmons, 2009, 101). Therefore, I include dummy variables for four main types of legal systems as coded by Powell and Mitchell (2007): common law, civil law, Islamic law, and mixed. In the models, civil law is the reference category.

Next, I include a measure of the objective level of respect for women’s rights using the CIRI index of women’s rights and the V-Dem women’s civil liberties index, lagged one year. In theory, countries with greater respect for women’s rights should be less likely to enter reservations, because their policies are already more in line with the treaty’s requirements. After reservations are entered, however, the objective level of respect for women’s rights should no longer directly affect the decision to keep reservations—rather, change in respect for women’s rights should affect these decisions.

I also include a measure of wealth, operationalized as GDP per capita. Simmons (2009) finds a positive correlation between GDP per capita and the tendency to enter reservations, which she attributes to the capacity for wealthier states to expend the necessary resources to carefully compare the treaty with domestic legislation and find discrepancies that warrant reservations. Once the discrepancies have been found and reservations entered, wealth should no longer be a relevant factor.

Finally, I address the stochastic dependence in the data for the selection equation
Table 2.2: Heckman Probit Models

<table>
<thead>
<tr>
<th>Outcome</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Democracy_{(t-1)}</td>
<td>0.608** (0.306)</td>
<td>0.946* (0.523)</td>
</tr>
<tr>
<td>Report Considered_{(t-1)}</td>
<td>0.712** (0.358)</td>
<td>1.451*** (0.406)</td>
</tr>
<tr>
<td>Democracy_{(t-1)} * Report_{(t-1)}</td>
<td>-0.686 (0.439)</td>
<td>-1.271** (0.524)</td>
</tr>
<tr>
<td>Initial Number of Reservations</td>
<td>0.0291*** (0.0109)</td>
<td>0.0212* (0.0117)</td>
</tr>
<tr>
<td>SOLS Change_{(t-1)}</td>
<td>0.122 (0.251)</td>
<td>0.245 (0.294)</td>
</tr>
<tr>
<td>Δ Women’s Rights_{(t-1)} (CIRI)</td>
<td>2.387** (1.000)</td>
<td></td>
</tr>
<tr>
<td>Δ Women’s Rights_{(t-1)} (V-Dem)</td>
<td>-0.00667 (0.0134)</td>
<td>-0.00158 (0.0235)</td>
</tr>
<tr>
<td>Time</td>
<td>-0.108 (0.131)</td>
<td>-0.441*** (0.165)</td>
</tr>
<tr>
<td>Time^2</td>
<td>0.0140 (0.0135)</td>
<td>0.0555*** (0.0182)</td>
</tr>
<tr>
<td>Time^3</td>
<td>-0.000340 (0.000366)</td>
<td>-0.00178*** (0.000562)</td>
</tr>
<tr>
<td>Constant</td>
<td>-2.240*** (0.764)</td>
<td>-1.901 (1.242)</td>
</tr>
<tr>
<td>Selection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democracy_{(t-1)}</td>
<td>0.447* (0.250)</td>
<td>0.818*** (0.262)</td>
</tr>
<tr>
<td>Women’s Rights_{(t-1)} (CIRI)</td>
<td></td>
<td>0.199** (0.0913)</td>
</tr>
<tr>
<td>Women’s Rights_{(t-1)} (V-Dem)</td>
<td>3.008*** (0.748)</td>
<td></td>
</tr>
<tr>
<td>GDP per capita_{(t-1)} (log)</td>
<td>-0.0947 (0.0923)</td>
<td>0.00388 (0.0915)</td>
</tr>
<tr>
<td>Common Law</td>
<td>0.0589 (0.399)</td>
<td>0.190 (0.381)</td>
</tr>
<tr>
<td>Mixed Law</td>
<td>-0.347 (0.523)</td>
<td>-0.0709 (0.574)</td>
</tr>
<tr>
<td>Islamic Law</td>
<td>0.452 (0.406)</td>
<td>0.345 (0.491)</td>
</tr>
<tr>
<td>Time since CEDAW</td>
<td>0.158* (0.0895)</td>
<td>0.382*** (0.144)</td>
</tr>
<tr>
<td>Time since CEDAW^2</td>
<td>0.00442 (0.00573)</td>
<td>-0.0167 (0.0107)</td>
</tr>
<tr>
<td>Time since CEDAW^3</td>
<td>-0.000202* (0.000109)</td>
<td>0.000317 (0.000238)</td>
</tr>
<tr>
<td>Constant</td>
<td>-3.912*** (1.042)</td>
<td>-4.250*** (1.036)</td>
</tr>
<tr>
<td>ρ</td>
<td>0.0108 (0.515)</td>
<td>0.0453 (0.780)</td>
</tr>
<tr>
<td>Observations</td>
<td>1576</td>
<td>1228</td>
</tr>
</tbody>
</table>

Clustered standard errors in parentheses

* p < 0.10, ** p < 0.05, *** p < 0.01

by including polynomials of the time since CEDAW opened for signature. To account for dependence among observations from the same country, I report standard errors clustered on country.

Table 2.2 displays results from two Heckman probit models, each utilizing one of the two measures for women’s rights. The coefficient estimates from the selection stage do indicate that democracies are generally more inclined to enter reservations compared to nondemocracies, but the results from the main analysis are robust to
accounting for this tendency. Counterintuitively, it appears that greater respect for women’s rights correlates with a greater propensity to enter reservations. Perhaps governments that already respect women’s rights are especially concerned that any incongruities between treaty provisions and domestic practice will be more easily noticed. A full exploration of this finding is beyond the scope of this paper, but it presents a puzzle to be addressed in future research. None of the other variables included in the selection equation, with the exception of the time polynomials, show strong correlations with entry into the sample. Overall, this analysis provides further support for the hypotheses.

### 2.7 Conclusion

This paper addresses an important question for scholars and practitioners of international cooperation: under what conditions do initially shallow commitments deepen over time? I investigate this question in the context of the human rights regime and argue that pressure from domestic actors encourages states to withdraw reservations to human rights treaties, thereby deepening their formal commitments. Therefore, democracies with domestic political conditions that favor mobilization should be more likely to withdraw reservations than their nondemocratic counterparts. In nondemocracies, the process of periodic review provides domestic groups with much-needed information and assistance to help them mobilize effectively. The salutary effects of the reporting procedure are less applicable to democratic states, where domestic groups already operate in a political environment favorable to mobilization. Evidence from CEDAW lends support for this argument. Democracies are more likely to withdraw reservations compared to nondemocracies, and nondemocracies are more likely to withdraw reservations after their periodic reports are considered by the CEDAW Committee.

While the hypotheses and results apply directly to human rights, the theory is
general and applies across issue areas. For instance, the theory also predicts that, over
time, democracies will be more likely to commit to take greater action to mitigate cli-
mate change in the framework of the Paris Agreement. Nondemocracies, by contrast,
will be more likely to strengthen their pledges after their policies have undergone
periodic review, a procedure established in the original agreement.

Therefore, this study has direct implications for the study and practice of pro-
moting human rights internationally and also speaks to the ongoing discussion of the
merits and pitfalls of flexibility mechanisms in agreement design. With respect to
the human rights regime, this research challenges studies that frame reservations as
attempts to dodge international obligations and undermine the regime by avoiding
commitments that entail substantial policy change. This paper therefore provides
empirical backing to claims that allowing reservations may in fact be beneficial to
a treaty regime, for it promotes broad, deep participation \cite{Swaine2006}. Indeed,
this paper suggests that reservations are, for many states, a temporary restriction
on the scope of their commitment that is subject to future revision. In other words,
reservations do not permanently alter a treaty’s legal effect, and there is room for
domestic and international actors to mobilize and successfully demand their removal.
Hence, this paper suggests advocates for withdrawal should direct their efforts toward
motivating and facilitating domestic mobilization and encouraging NGOs, especially
those from nondemocratic states, to participate in the reporting process.

More broadly, this paper contributes to our understanding of how states use flexi-
bility mechanisms once party to a treaty. While scholars generally agree that flexible
terms make treaties more attractive to potential signatories, predictions of the con-
sequences of these flexibility mechanisms for a treaty’s ability to affect behavior are
mixed. In the context of human rights, observers wonder whether human rights
treaties have any meaning if states are free to design their own commitments by
entering reservations to select provisions. If reservations were permanent and irrevo-
cable, the answer might be no. Yet this paper shows not only that reservations are susceptible to withdrawal, but also that their removal can be hastened by domestic mobilization. In other words, initially shallow commitments can deepen over time. This implication extends beyond the human rights arena to other issue areas as well. For instance, the Paris Agreement on climate change incorporates extensive flexibility by allowing states to set their own mitigation commitments but subjects states’ parties to periodic review of their progress toward meeting their stated goals (Keohane and Victor, 2016). This paper suggests that states’ contributions should increase over time if they face sustained domestic pressure to revise their pledges accordingly and involve NGOs in the review process.

Although this study cannot tell us whether allowing states to make shallow commitments is necessarily better than forcing states to choose between a deep commitment and no commitment at all, it strongly suggests the differences between these two design choices are smaller than they may seem. The conventional wisdom suggests flexible agreements should always be less effective at inducing behavioral change than more restrictive treaties (e.g., Köppel and Sprinz, 2019), but this simplistic view ignores the potential for depth to change over time. Future research should investigate precisely how behavioral change fits into this picture. This study focuses on states’ formal commitments under the assumption that formal commitments are at least weakly related to state behavior. Of course, domestic actors do not mobilize for deepening states’ commitments simply for its own sake. Rather, domestic activists seek deeper commitments so that they can hold governments accountable to more stringent legal obligations. In the context of human rights treaties, activists work simultaneously to pressure governments to withdraw reservations so they can be held accountable to deeper commitments and to bring domestic policy in line with international standards to obviate the need for reservations. Ultimately, these organizations seek greater respect for the rights outlined in the treaty, but they regard changes in
states’ formal commitments as important steps in this process.
## 2.8 Appendix 1: Reservations to CEDAW by Provision

<table>
<thead>
<tr>
<th>Provision</th>
<th>Reservations Entered</th>
<th>Reservations Withdrawn</th>
<th>Brief Description of the Provision’s Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>0</td>
<td>Definition of discrimination against women.</td>
</tr>
<tr>
<td>2:a</td>
<td>10</td>
<td>0</td>
<td>Embodiment of the principle of the equality of men and women in national constitutions and appropriate legislation.</td>
</tr>
<tr>
<td>2:b</td>
<td>9</td>
<td>0</td>
<td>Adoption of appropriate legislative and other measures that prohibit all discrimination against women.</td>
</tr>
<tr>
<td>2:c</td>
<td>9</td>
<td>0</td>
<td>Establishment of legal protection of the rights of women.</td>
</tr>
<tr>
<td>2:d</td>
<td>10</td>
<td>0</td>
<td>Avoidance of engagement in discrimination against women.</td>
</tr>
<tr>
<td>2:e</td>
<td>9</td>
<td>0</td>
<td>Elimination of discrimination against women.</td>
</tr>
<tr>
<td>2:f</td>
<td>15</td>
<td>3</td>
<td>Modification or abolishment of existing laws, regulations, customs and practices which constitute discrimination against women.</td>
</tr>
<tr>
<td>2:g</td>
<td>11</td>
<td>2</td>
<td>Repeal of national penal provisions which constitute discrimination against women.</td>
</tr>
<tr>
<td>5:a</td>
<td>6</td>
<td>3</td>
<td>Modification of social and cultural patterns to eliminate practices based on the idea of the inferiority or superiority of either of the sexes or on stereotyped roles for men and women.</td>
</tr>
<tr>
<td>5:b</td>
<td>2</td>
<td>2</td>
<td>Recognition in family education of maternity as a social function and the common responsibility of men and women in the upbringing of children.</td>
</tr>
<tr>
<td>7:a</td>
<td>5</td>
<td>5</td>
<td>Women’s right to vote and eligibility for election.</td>
</tr>
<tr>
<td>7:b</td>
<td>9</td>
<td>8</td>
<td>Women’s right to participate in functions of government.</td>
</tr>
<tr>
<td>7:c</td>
<td>4</td>
<td>4</td>
<td>Women’s right to participate in non-governmental organizations and associations in the public sphere.</td>
</tr>
<tr>
<td>9:1</td>
<td>7</td>
<td>5</td>
<td>Equal rights for women and men to acquire, change, or retain their nationality.</td>
</tr>
<tr>
<td>9:2</td>
<td>22</td>
<td>11</td>
<td>Equal rights for women and men with respect to the nationality of their children.</td>
</tr>
<tr>
<td>10:a</td>
<td>1</td>
<td>1</td>
<td>Equal access to educational facilities.</td>
</tr>
<tr>
<td>10:b</td>
<td>1</td>
<td>1</td>
<td>Equal access to educational curricula, examinations, teaching staff, and equipment.</td>
</tr>
<tr>
<td>10:c</td>
<td>2</td>
<td>2</td>
<td>Elimination of stereotyped concepts of the roles of men and women at all educational levels.</td>
</tr>
<tr>
<td>10:d</td>
<td>1</td>
<td>1</td>
<td>Equal opportunity to benefit from scholarships and other grants.</td>
</tr>
<tr>
<td>10:e</td>
<td>1</td>
<td>1</td>
<td>Equal access to continuing education programs.</td>
</tr>
<tr>
<td>10:f</td>
<td>1</td>
<td>1</td>
<td>Reduction of female student drop-out rates and the organization of programs for girls and women who have left school prematurely.</td>
</tr>
<tr>
<td>10:g</td>
<td>1</td>
<td>1</td>
<td>Equal opportunity to participate in sports and physical education.</td>
</tr>
<tr>
<td>10:h</td>
<td>1</td>
<td>1</td>
<td>Equal access to educational information to ensure the health and well-being of families.</td>
</tr>
<tr>
<td>11:1a</td>
<td>3</td>
<td>0</td>
<td>Women’s right to work.</td>
</tr>
<tr>
<td>11:1b</td>
<td>6</td>
<td>2</td>
<td>Women’s right to the same employment opportunities as men.</td>
</tr>
<tr>
<td>11:1c</td>
<td>4</td>
<td>0</td>
<td>Women’s right to a free choice of profession and employment, promotion, job security, and benefits.</td>
</tr>
<tr>
<td>11:1d</td>
<td>6</td>
<td>2</td>
<td>Equal remuneration and equal treatment for work of equal value.</td>
</tr>
<tr>
<td>11:1e</td>
<td>3</td>
<td>0</td>
<td>Women’s right to social security and paid leave.</td>
</tr>
<tr>
<td>11:1f</td>
<td>4</td>
<td>1</td>
<td>Women’s right to protection of health and safety in working conditions.</td>
</tr>
<tr>
<td>11:2a</td>
<td>2</td>
<td>0</td>
<td>Prohibition of dismissal on the grounds of pregnancy or maternity leave and discrimination in dismissals on the basis of marital status.</td>
</tr>
<tr>
<td>Section</td>
<td>No.</td>
<td>Code</td>
<td>Content</td>
</tr>
<tr>
<td>---------</td>
<td>-----</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td>11:2b</td>
<td>4</td>
<td>1</td>
<td>Introduction of maternity leave with pay or with comparable social benefits.</td>
</tr>
<tr>
<td>11:2c</td>
<td>2</td>
<td>0</td>
<td>Provision of social services to enable parents to combine family obligations with work responsibilities.</td>
</tr>
<tr>
<td>11:2d</td>
<td>2</td>
<td>0</td>
<td>Special protection to women during pregnancy.</td>
</tr>
<tr>
<td>11:3</td>
<td>1</td>
<td>0</td>
<td>Periodic reviews and revision, repeal, or extension of protective legislation.</td>
</tr>
<tr>
<td>13:a</td>
<td>4</td>
<td>2</td>
<td>Equal rights to family benefits.</td>
</tr>
<tr>
<td>13:b</td>
<td>2</td>
<td>2</td>
<td>Equal rights to financial credit.</td>
</tr>
<tr>
<td>13:c</td>
<td>2</td>
<td>2</td>
<td>Equal rights to participate in recreational activities.</td>
</tr>
<tr>
<td>14:2c</td>
<td>1</td>
<td>1</td>
<td>Rural women’s right to benefit directly from social security programs.</td>
</tr>
<tr>
<td>14:2h</td>
<td>1</td>
<td>1</td>
<td>Rural women’s right to enjoy adequate living conditions.</td>
</tr>
<tr>
<td>15:1</td>
<td>1</td>
<td>0</td>
<td>Equality for men and women before the law.</td>
</tr>
<tr>
<td>15:2</td>
<td>6</td>
<td>4</td>
<td>Identical legal capacity for women and men in civil matters.</td>
</tr>
<tr>
<td>15:3</td>
<td>5</td>
<td>4</td>
<td>Nullification of contracts directed at restricting the legal capacity of women.</td>
</tr>
<tr>
<td>15:4</td>
<td>13</td>
<td>6</td>
<td>The same rights for men and women with regard to laws relating to the movement of persons and the freedom to choose their residence and domicile.</td>
</tr>
<tr>
<td>16:1a</td>
<td>15</td>
<td>3</td>
<td>The same right for men and women to enter into marriage.</td>
</tr>
<tr>
<td>16:1b</td>
<td>12</td>
<td>4</td>
<td>The same right for men and women to freely choose a spouse and to enter into marriage only with consent.</td>
</tr>
<tr>
<td>16:1c</td>
<td>24</td>
<td>7</td>
<td>The same rights and responsibilities for men and women during marriage and at its dissolution.</td>
</tr>
<tr>
<td>16:1d</td>
<td>21</td>
<td>8</td>
<td>The same rights and responsibilities for men and women as parents in matters relating to their children.</td>
</tr>
<tr>
<td>16:1e</td>
<td>13</td>
<td>4</td>
<td>The same rights for men and women to decide freely on the number and spacing of their children.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The same rights and responsibilities for men and women with regard to guardianship, wardship, trusteeship and adoption of children.</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>16:1f</td>
<td>23</td>
<td>7</td>
<td>The same personal rights for men and women as husband and wife, including the right to choose a family name, a profession, and an occupation.</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>The same rights for men and women in respect of the ownership, acquisition, management, administration, enjoyment and disposition of property.</td>
</tr>
<tr>
<td>16:2</td>
<td>13</td>
<td>3</td>
<td>Assurance that the betrothal and the marriage of a child has no legal effect and specification of a minimum age for marriage.</td>
</tr>
<tr>
<td>General</td>
<td>12</td>
<td>6</td>
<td>Content varies.</td>
</tr>
</tbody>
</table>
## 2.9 Appendix 2: Reservations to CEDAW by Country

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Algeria</td>
<td>18</td>
<td>2:a-g, 9:2, 15:4, 16:1a-h, 16:2</td>
<td>1</td>
<td>9:2</td>
</tr>
<tr>
<td>Australia</td>
<td>2</td>
<td>General, 11:2b</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Austria</td>
<td>2</td>
<td>7:b, 11:1f</td>
<td>2</td>
<td>7:b, 11:1f</td>
</tr>
<tr>
<td>Bahamas*</td>
<td>3</td>
<td>2:a, 9:2, 16:1h</td>
<td>1</td>
<td>16:1h</td>
</tr>
<tr>
<td>Bahrain</td>
<td>18</td>
<td>2:a-g, 9:2, 15:4, 16:1a-h, 16:2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Bangladesh</td>
<td>10</td>
<td>2:a-g, 13:a, 16:1c, 16:1f</td>
<td>2</td>
<td>13:a, 16:1f</td>
</tr>
<tr>
<td>Belgium</td>
<td>5</td>
<td>7:a-c, 15:2-3</td>
<td>5</td>
<td>7:a-c, 15:2-3</td>
</tr>
<tr>
<td>Brazil</td>
<td>5</td>
<td>15:4, 16:1a, 16:1c, 16:1g-h</td>
<td>5</td>
<td>15:4, 16:1a, 16:1c, 16:1g-h</td>
</tr>
<tr>
<td>Brunei*</td>
<td>1</td>
<td>9:2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Canada</td>
<td>1</td>
<td>11:1d</td>
<td>1</td>
<td>11:1d</td>
</tr>
<tr>
<td>Cook Islands*</td>
<td>4</td>
<td>General, 2:f, 5:a, 11:2b</td>
<td>4</td>
<td>General, 2:f, 5:a, 11:2b</td>
</tr>
<tr>
<td>Cyprus</td>
<td>1</td>
<td>9:2</td>
<td>1</td>
<td>9:2</td>
</tr>
<tr>
<td>Democratic People’s Republic of Korea</td>
<td>2</td>
<td>2:f, 9:2</td>
<td>2</td>
<td>2:f, 9:2</td>
</tr>
<tr>
<td>Egypt</td>
<td>17</td>
<td>2:a-g, 9:2, 16:1a-h, 16:2</td>
<td>1</td>
<td>9:2</td>
</tr>
<tr>
<td>Fiji</td>
<td>3</td>
<td>5:a, 9:1, 9:2</td>
<td>3</td>
<td>5:a, 9:1, 9:2</td>
</tr>
<tr>
<td>France</td>
<td>12</td>
<td>5:b, 7:a-c, 14:2c, 14:2h, 15:2-3, 16:1c-d, 16:1g-h</td>
<td>12</td>
<td>5:b, 7:a-c, 14:2c, 14:2h, 15:2-3, 16:1c-d, 16:1g-h</td>
</tr>
<tr>
<td>Germany</td>
<td>1</td>
<td>7:b</td>
<td>1</td>
<td>7:b</td>
</tr>
<tr>
<td>India</td>
<td>10</td>
<td>5:a, 16:1a-h, 16:2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Iraq</td>
<td>13</td>
<td>2:f-g, 9:1-2, 16:1a-h, 16:2</td>
<td>2</td>
<td>9:1-2</td>
</tr>
<tr>
<td>Ireland</td>
<td>11</td>
<td>9:1, 11:1b-d, 13:a-c, 15:3-4, 16:1d, 16:1f</td>
<td>5</td>
<td>9:1, 13:b-c, 15:3-4</td>
</tr>
<tr>
<td>Israel</td>
<td>10</td>
<td>7:b, 16:1a-h, 16:2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Count</td>
<td>Pages</td>
<td>Count</td>
<td>Pages</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
<td>-------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Jamaica</td>
<td>1</td>
<td>9:2</td>
<td>1</td>
<td>9:2</td>
</tr>
<tr>
<td>Jordan</td>
<td>5</td>
<td>9:2, 15:4, 16:1c-d, 16:1g</td>
<td>1</td>
<td>15:4</td>
</tr>
<tr>
<td>Kuwait</td>
<td>3</td>
<td>7:a, 9:2, 16:1f</td>
<td>1</td>
<td>7:a</td>
</tr>
<tr>
<td>Lebanon</td>
<td>5</td>
<td>9:2, 16:1c-d, 16:1f-g</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Lesotho</td>
<td>8</td>
<td>General, 2:a-g</td>
<td>1</td>
<td>General</td>
</tr>
<tr>
<td>Libyan Arab Jamahiriya</td>
<td>10</td>
<td>General, 2:a-g, 16:1c-d</td>
<td>1</td>
<td>General</td>
</tr>
<tr>
<td>Liechtenstein*</td>
<td>2</td>
<td>1, 9:2</td>
<td>1</td>
<td>9:2</td>
</tr>
<tr>
<td>Luxembourg</td>
<td>4</td>
<td>7:a-c, 16:1g</td>
<td>4</td>
<td>7:a-c, 16:1g</td>
</tr>
<tr>
<td>Malawi</td>
<td>2</td>
<td>5:a-b</td>
<td>2</td>
<td>5:a-b</td>
</tr>
<tr>
<td>Maldives*</td>
<td>10</td>
<td>7:a, 16:1a-h, 16:2</td>
<td>1</td>
<td>7:a</td>
</tr>
<tr>
<td>Malta*</td>
<td>14</td>
<td>11:a-f, 13:a-c, 15:1-4, 16:1e</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Mauritania</td>
<td>11</td>
<td>General, 13:a, 16:1a-h, 16:2</td>
<td>1</td>
<td>General</td>
</tr>
<tr>
<td>Mauritius</td>
<td>3</td>
<td>11:1b, 11:1d, 16:1g</td>
<td>3</td>
<td>11:1b, 11:1d, 16:1g</td>
</tr>
<tr>
<td>Micronesia*</td>
<td>14</td>
<td>2:f, 5:a-b, 11:1d, 11:2b, 16:1a-h, 16:2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Monaco*</td>
<td>6</td>
<td>General, 7:b, 9:1-2, 16:1e, 16:1g</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Morocco</td>
<td>18</td>
<td>2:a-g, 9:2, 15:4, 16:1a-h, 16:2</td>
<td>10</td>
<td>9:2, 16:1a-h, 16:2</td>
</tr>
<tr>
<td>New Zealand</td>
<td>3</td>
<td>General (2), 11:2b</td>
<td>3</td>
<td>General (2), 11:2b</td>
</tr>
<tr>
<td>Niger</td>
<td>7</td>
<td>2:d, 2:f, 5:a, 15:4, 16:1c, 16:1e, 16:1g</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Oman</td>
<td>6</td>
<td>General, 9:2, 15:4, 16:1a, 16:1c, 16:1f</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Pakistan</td>
<td>1</td>
<td>General</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Country</td>
<td>Text References</td>
<td>Count</td>
<td>Additional References</td>
<td></td>
</tr>
<tr>
<td>--------------------------</td>
<td>------------------------------------------------------</td>
<td>-------</td>
<td>-------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>Qatar</td>
<td>1, 2:a, 5:a, 9:2, 15:1, 15:4, 16:1a, 16:1c, 16:1f</td>
<td>9</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Republic of Korea</td>
<td>9:1-2, 16:1c-d, 16:1f-g</td>
<td>6</td>
<td>5, 9:1-2, 16:1c-d, 16:1f</td>
<td></td>
</tr>
<tr>
<td>Saudi Arabia</td>
<td>General, 9:2</td>
<td>2</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Singapore</td>
<td>General, 2:a-g, 11:1a-f, 16:1a-h, 16:2</td>
<td>23</td>
<td>7, General, 2:g, 11:1a-f, 16:1b, 16:1d-g</td>
<td></td>
</tr>
<tr>
<td>Spain</td>
<td>General</td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Switzerland</td>
<td>7:b, 15:2, 16:1g-h</td>
<td>4</td>
<td>2, 7:b, 16:1g</td>
<td></td>
</tr>
<tr>
<td>Syrian Arab Republic</td>
<td>2:a-g, 9:2, 15:4, 16:1c-d, 16:1f-g, 16:2</td>
<td>14</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>Thailand</td>
<td>7:a-c, 9:2, 10:a-h, 11:1b, 15:3, 16:1a-h, 16:2</td>
<td>23</td>
<td>23, 7:a-c, 9:2, 10:a-h, 11:1b, 15:3, 16:1a-h, 16:2</td>
<td></td>
</tr>
<tr>
<td>Tunisia</td>
<td>General, 9:2, 15:4, 16:1c-d, 16:1f-h</td>
<td>8</td>
<td>7, 9:2, 15:4, 16:1c-d, 16:1f-h</td>
<td></td>
</tr>
<tr>
<td>Turkey</td>
<td>15:2, 15:4, 16:1c-d, 16:1f-g</td>
<td>6</td>
<td>6, 15:2, 15:4, 16:1c-d, 16:1f-g</td>
<td></td>
</tr>
<tr>
<td>United Arab Emirates</td>
<td>2:f, 9:1-2, 15:2, 16:1a-h, 16:2</td>
<td>13</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>United Kingdom</td>
<td>1, 2:f-g, 9:1-2, 10:c, 11:1a-f, 11:2a-d, 13:a-c, 15:2-4, 16:1f</td>
<td>20</td>
<td>8, 2:f-g, 10:c, 13:a-c, 15:2, 15:4</td>
<td></td>
</tr>
</tbody>
</table>

* Small state omitted from analysis.
2.10 Appendix 3: Alternative Coding of Women’s Rights

The analyses below replicate Models 2, 3, 5, and 6 using the objective level of respect for women’s rights and the annual change in respect for women’s rights in lieu of the change since ratification. The overall substantive interpretation of these results is the same as that in the paper—respect for women’s rights is only weakly correlated with withdrawing reservations, and this relationship rarely approaches conventional levels of statistical significance.
2.10.1 Objective Level of Respect for Women’s Rights

Below are the results from replicating Models 2, 3, 5, and 6 where the variables for change in respect for women’s rights since ratification are replaced by the objective level of respect for women’s rights, lagged one year.

Table 2.5

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>Withdrawing at least one reservation</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(2)</td>
<td>(3)</td>
<td>(5)</td>
<td>(6)</td>
</tr>
<tr>
<td>Democracy_{(t-1)}</td>
<td>0.830</td>
<td>0.638</td>
<td>1.787**</td>
<td>1.033*</td>
</tr>
<tr>
<td></td>
<td>(0.605)</td>
<td>(0.536)</td>
<td>(0.832)</td>
<td>(0.604)</td>
</tr>
<tr>
<td>Report Considered_{(t-1)}</td>
<td>1.085**</td>
<td>0.568</td>
<td>3.347***</td>
<td>1.671**</td>
</tr>
<tr>
<td></td>
<td>(0.453)</td>
<td>(0.424)</td>
<td>(0.936)</td>
<td>(0.695)</td>
</tr>
<tr>
<td>Democracy_{(t-1)} * Report_{(t-1)}</td>
<td>-3.035***</td>
<td>-1.647*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.104)</td>
<td>(0.894)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Initial Number of Reservations</td>
<td>0.045</td>
<td>0.056**</td>
<td>0.042</td>
<td>0.055**</td>
</tr>
<tr>
<td></td>
<td>(0.030)</td>
<td>(0.026)</td>
<td>(0.031)</td>
<td>(0.026)</td>
</tr>
<tr>
<td>SOLS Change_{(t-1)}</td>
<td>0.489</td>
<td>0.321</td>
<td>0.469</td>
<td>0.305</td>
</tr>
<tr>
<td></td>
<td>(0.551)</td>
<td>(0.513)</td>
<td>(0.549)</td>
<td>(0.514)</td>
</tr>
<tr>
<td>Women’s Rights_{(t-1)} (CIRI)</td>
<td>0.050</td>
<td>0.079</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.134)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Women’s Rights_{(t-1)} (V-Dem)</td>
<td>1.116</td>
<td></td>
<td>1.081</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(1.233)</td>
<td></td>
<td>(1.239)</td>
<td></td>
</tr>
<tr>
<td>Previous Withdrawals_{(t-1)}</td>
<td>-0.008</td>
<td>-0.014</td>
<td>-0.006</td>
<td>-0.014</td>
</tr>
<tr>
<td></td>
<td>(0.024)</td>
<td>(0.017)</td>
<td>(0.025)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Time</td>
<td>-0.766**</td>
<td>-0.142</td>
<td>-0.858**</td>
<td>-0.154</td>
</tr>
<tr>
<td></td>
<td>(0.370)</td>
<td>(0.242)</td>
<td>(0.376)</td>
<td>(0.243)</td>
</tr>
<tr>
<td>Time$^2$</td>
<td>0.100**</td>
<td>0.023</td>
<td>0.108**</td>
<td>0.024</td>
</tr>
<tr>
<td></td>
<td>(0.044)</td>
<td>(0.023)</td>
<td>(0.044)</td>
<td>(0.023)</td>
</tr>
<tr>
<td>Time$^3$</td>
<td>-0.003**</td>
<td>-0.001</td>
<td>-0.004**</td>
<td>-0.001</td>
</tr>
<tr>
<td></td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
<td>(0.001)</td>
</tr>
<tr>
<td>Constant</td>
<td>-3.105***</td>
<td>-4.731***</td>
<td>-3.817***</td>
<td>-4.936***</td>
</tr>
<tr>
<td></td>
<td>(1.166)</td>
<td>(1.226)</td>
<td>(1.295)</td>
<td>(1.246)</td>
</tr>
<tr>
<td>Observations</td>
<td>529</td>
<td>676</td>
<td>529</td>
<td>676</td>
</tr>
</tbody>
</table>

*p<0.1; **p<0.05; ***p<0.01
Standard errors in parentheses
2.10.2 Annual Change in Respect for Women’s Rights

Below are the results from replicating Models 2, 3, 5, and 6 where the variables for change in respect for women’s rights since ratification are replaced by the annual change in respect for women’s rights. This variable is created by subtracting from each country-year’s objective level of respect for women’s rights the level observed in the previous year. This variable is also lagged one year.

Table 2.6

<table>
<thead>
<tr>
<th>Dependent variable:</th>
<th>(2)</th>
<th>(3)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Withdrawing at least one reservation</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Democracy (_{(t-1)})</td>
<td>0.952*</td>
<td>0.953**</td>
<td>1.966**</td>
<td>1.349***</td>
</tr>
<tr>
<td>(0.515)</td>
<td>(0.421)</td>
<td>(0.778)</td>
<td>(0.502)</td>
<td></td>
</tr>
<tr>
<td>Report Considered (_{(t-1)})</td>
<td>1.130**</td>
<td>0.643</td>
<td>3.343***</td>
<td>1.752**</td>
</tr>
<tr>
<td>(0.456)</td>
<td>(0.426)</td>
<td>(0.938)</td>
<td>(0.696)</td>
<td></td>
</tr>
<tr>
<td>Democracy (<em>{(t-1)}) * Report (</em>{(t-1)})</td>
<td>-2.971***</td>
<td>-1.661*</td>
<td>(1.103)</td>
<td>(0.895)</td>
</tr>
<tr>
<td>Initial Number of Reservations</td>
<td>0.043</td>
<td>0.062**</td>
<td>0.042</td>
<td>0.061**</td>
</tr>
<tr>
<td>(0.030)</td>
<td>(0.026)</td>
<td>(0.031)</td>
<td>(0.026)</td>
<td></td>
</tr>
<tr>
<td>SOLS Change (_{(t-1)})</td>
<td>0.480</td>
<td>0.261</td>
<td>0.440</td>
<td>0.251</td>
</tr>
<tr>
<td>(0.552)</td>
<td>(0.515)</td>
<td>(0.551)</td>
<td>(0.515)</td>
<td></td>
</tr>
<tr>
<td>Women’s Rights (_{(t-1)}) (CIRI)</td>
<td>0.285</td>
<td>0.289</td>
<td>(0.253)</td>
<td>(0.253)</td>
</tr>
<tr>
<td>Women’s Rights (_{(t-1)}) (V-Dem)</td>
<td>18.204*</td>
<td>17.676*</td>
<td>(9.688)</td>
<td>(9.688)</td>
</tr>
<tr>
<td>Previous Withdrawals (_{(t-1)})</td>
<td>-0.007</td>
<td>-0.019</td>
<td>-0.006</td>
<td>-0.019</td>
</tr>
<tr>
<td>(0.024)</td>
<td>(0.017)</td>
<td>(0.025)</td>
<td>(0.017)</td>
<td></td>
</tr>
<tr>
<td>Time</td>
<td>-0.749**</td>
<td>-0.176</td>
<td>-0.846**</td>
<td>-0.188</td>
</tr>
<tr>
<td>(0.378)</td>
<td>(0.244)</td>
<td>(0.380)</td>
<td>(0.245)</td>
<td></td>
</tr>
<tr>
<td>Time(^2)</td>
<td>0.097**</td>
<td>0.027</td>
<td>0.107**</td>
<td>0.028</td>
</tr>
<tr>
<td>(0.045)</td>
<td>(0.023)</td>
<td>(0.045)</td>
<td>(0.023)</td>
<td></td>
</tr>
<tr>
<td>Time(^3)</td>
<td>-0.003**</td>
<td>-0.001</td>
<td>-0.003**</td>
<td>-0.001</td>
</tr>
<tr>
<td>(0.002)</td>
<td>(0.001)</td>
<td>(0.002)</td>
<td>(0.001)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-2.973***</td>
<td>-4.072***</td>
<td>-3.594***</td>
<td>-4.315***</td>
</tr>
<tr>
<td>(1.104)</td>
<td>(0.917)</td>
<td>(1.234)</td>
<td>(0.949)</td>
<td></td>
</tr>
<tr>
<td>Observations</td>
<td>518</td>
<td>674</td>
<td>518</td>
<td>674</td>
</tr>
</tbody>
</table>

*p<0.1; **p<0.05; ***p<0.01

Standard errors in parentheses
Chapter 3

Information Provision and Domestic Sources of International Cooperation: The Utility of “Pledge and Review” in Climate Politics

How can international institutions effect change among self-interested states? Scholars often look to domestic political dynamics to explain patterns of cooperation and compliance with international commitments in areas such as environmental protection and human rights, where common strategies of reciprocity, retaliation, and reputation (Guzman, 2008) are less applicable. Domestic actors can mobilize to demand cooperative behavior from their governments and impose costs on them if they do not deliver (Simmons, 2009; McGillivray and Smith, 2008). In this scenario, international institutions play a critical role as providers of information. Information from international institutions helps domestic actors gauge the extent to which their government is fulfilling its obligations and facilitate mobilization when government policies are inadequate (Dai, 2007; Chaudoin, 2014a).

This paper focuses on the role of information generated by international institutions in environmental politics, specially the issue of climate change, which exhibits the basic characteristics of a threshold public good. In order to stave off the effects of climate change, states must agree to collectively contribute a sufficient amount of costly effort. If the threshold is not met, the consequences will be devastating for
the citizens of all countries, regardless of how much their governments contributed to global mitigation efforts. Therefore, domestic actors such as interest groups and non-governmental organizations (NGOs) who wish to see this crisis averted face a difficult informational environment where they cannot be sure whether their government’s contributions are sufficient and whether they should expend effort to punish a leader for inadequate contributions when such action is costly and success is not guaranteed.

I utilize an experimental design to examine the effects of information from international institutions in this scenario. An experiment allows me to explicitly manipulate the institutional setting to focus on the way information affects the willingness of domestic actors to exert costly pressure on a leader to change its behavior. While this is not the first experiment to explore this question (see Iris, Lee and Tavoni, 2016; McGillivray and Smith, 2008), my design incorporates important features of the institutional setting that other projects have not examined together. I show that when the institution provides domestic interest groups with information about their government’s contributions, government leaders contribute greater amounts to the public good, and domestic actors condition their decisions to try and remove a leader on the quality of leaders’ contributions.

These findings have important implications for the study and practice of international cooperation. Specifically, the results suggest international negotiators should direct their efforts toward information-providing mechanisms over other forms of enforcement. Scholars generally expect states will be wary to join international institutions that include stringent enforcement mechanisms and impose hefty penalties for noncompliance (Downs, Rocke and Barsoom, 1996; Rosendorff and Milner, 2001; Koremenos, 2005; von Stein, 2005). Hence, agreements designed to produce information rather than punish defectors can help attract more signatories to a treaty. Even if states suspect that information will be used by domestic actors to generate noncompliance costs, they may still be willing to sign if they expect they will be able
to withstand the pressure or if they face other incentives to join.

The ability to attract a large number of signatories is especially important in climate change, where widespread involvement is necessary to yield benefits. The experiment in this paper is modeled after the “pledge and review” process included in the 2015 Paris Climate Agreement (Keohane and Oppenheimer 2016). Therefore, the findings from this study also support the notion that this design feature can be an effective means of achieving broad cooperation to mitigate climate change. These results complement a recent experiment by Barrett and Dannenberg (2016), which shows a small improvement in contributions when a review process is implemented. However, this experiment explicitly identifies the mechanism through which pledge and review works to increase contributions and incorporates a design that allows this dynamic to play out.

This paper proceeds as follows. First, I explain how information provision empowers domestic actors to influence state behavior in international institutions. Based on this literature, I expect domestic actors will be more likely to punish leaders when they have more information about leader performance and that this increased risk of punishment motivates greater contributions to a public good. Next, I describe the experimental design to test my expectations and present results from its execution. The final section concludes with a discussion of the study’s contributions and implications.

3.1 The Role of Information in International Institutions

Information from international institutions enables domestic constituencies to enforce compliance with international commitments because it helps domestic actors recognize when to mobilize and gives them the tools to succeed when they do. International institutions can “sound the alarm” when states fail to live up to their international obligations and alert domestic actors to their government’s deficiencies
This information further enhances domestic actors’ ability to condition support on government policies and successfully pressure the government to change course (Dai 2007). Based on this logic, I argue that international institutions that provide information about government activities should be more effective in eliciting cooperative behavior from signatories compared to institutions that do not provide information.

Domestic constituencies have a clear informational disadvantage when it comes to evaluating government policies in the context of international cooperation. Identifying violations of international obligations requires extensive knowledge of the treaty’s provisions and how its requirements compare to existing policies (Dai 2007). If domestic actors wish to hold their government accountable to international commitments, they must know precisely what was promised and to what extent the government is fulfilling its obligation. Meanwhile, governments have incentives to disguise their inadequacies or blame their inability to meet certain requirements on external factors beyond their control (Dai 2007).

International institutions alleviate these problems by providing easily accessible information about the quality of government policies. For instance, members of the World Trade Organization (WTO) can initiate disputes against other countries to alert domestic audiences when their government has violated treaty obligations (Chaudoin 2014a). In diverse issue areas such as environmental protection, money laundering, intellectual property rights, and human trafficking, international actors routinely release performance indicators that condense comprehensive information about state practices into simple, comparable measures that domestic groups can use to demand attention to a given issue (Kelley and Pevehouse 2015, Kelley and Simmons 2019). Within the United Nations, the major human rights treaties require states parties to submit periodic reports detailing their efforts to comply with treaty provisions, which are then reviewed by a committee that releases a set of
recommendations for improvement (Creamer and Simmons, 2016). These recommendations clue domestic actors into where government deficiencies exist and legitimate their demands for change (Creamer and Simmons, 2016). Hence, information enables domestic actors to identify when mobilization may be necessary to alter the government’s activities.

Information from international institutions can also increase the likelihood of successful domestic mobilization, because it enables domestic actors to directly tie political support to government policies. Groups with substantial information about policy can make credible threats to withdraw support if the policies are not changed (Dai, 2007). In other words, governments can be confident that ignoring such demands will incur substantial political costs. The experience of the Convention on Long-Range Transboundary Air Pollution (CLRTAP) illustrates these points well. Dai (2007, 110-111) describes how the CLRTAP motivated and facilitated environmental activism by disseminating research on acid damage, information about emission trajectories, and compilations of abatement strategies. As environmental groups acquired information about the substance of the convention and government actions, they became better equipped to demand policy changes from their governments.

Citizens interact with information from international institutions through a variety of mediums. Opposition parties have incentives to disseminate information that portrays the ruling party in a negative light to drum up support for their party as a credible alternative. Non-governmental organizations (NGOs) that advocate for greater compliance with the treaty disseminate information about government inadequacies to motivate individuals to support the organization and join their efforts (Simmons, 2009; Dai, 2007). Moreover, international criticism directed at a government is often a newsworthy event within that country, and heightened media attention increases the likelihood that ordinary citizens will encounter this information. For instance, Creamer and Simmons (2016) show that local press outlets in Latin America
publish more articles referencing the Convention Against Torture when the government’s report is under review and immediately following the release of the committee’s recommendations.

In sum, information from international institutions encourages and empowers domestic audiences to hold their governments accountable to international promises. Yet institutions differ in the extent to which they provide this information. For instance, just 60% of treaties in a random sample of international agreements drawn from the UN Treaty Series database include mechanisms for monitoring compliance with treaty provisions (Koremenos, 2016). Hence, a substantial proportion of treaties provide limited or no information to domestic constituencies.

The argument for the link between information provision and compliance implies institutions that provide more information should lead to greater conformity with treaty standards. Investigating this prediction, however, is nearly impossible with observational data. Existing data on monitoring provisions in international agreements are subject to two strong selection effects that make causal inference about the power of information provision particularly problematic. First, negotiators base the decision to include monitoring provisions on the characteristics of the problem the treaty is intended to address and expectations about state incentives to violate treaty provisions (Koremenos, 2016). Second, in the case of multilateral treaties, states decide whether to join an institution based in large part on the treaty’s design, including considerations of the extent to which the treaty requires close monitoring of state activity. It is therefore conceivable that negotiators only include monitoring provisions when states are less likely to comply with the treaty’s requirements, but states only ratify agreements with stringent monitoring provisions when they expect to comply. Thus, the effects of information provision cannot be convincingly separated from expectations about compliance.

To address this issue, I use an experiment in which I directly manipulate the
extent to which the institution provides information. Laboratory experiments enable researchers to exert full control over the research setting to directly manipulate the independent variable of interest and produce clean estimates of the intervention’s effects (Morton and Williams 2010). In an experimental setting, I can hold constant all extraneous features of the institution’s design and focus only on the features that matter for the theory. I explain the experimental design in the following section.

3.2 Modeling Climate Change Politics

To test the effects of information provision on domestic mobilization and state behavior in international institutions, I design an experiment modeled after the “pledge and review” process in the 2015 Paris Climate Agreement (Keohane and Oppenheimer 2016). According to this treaty, parties’ declared emission reduction targets and subsequent emissions will be periodically reviewed by technical experts to determine whether the target was reached and provide suggestions for improvement (Barrett and Dannenberg 2016).

I focus on the issue of climate change, because it is an area in which nearly universal international cooperation is essential to averting a devastating outcome for the planet. Further complicating international cooperation in this area is the fact that other mechanisms for motivating compliance, such as reciprocity, retaliation for bad behavior, and reputational considerations (see Guzman 2008), are less likely to be useful in this context. Reciprocal actions do not produce much benefit, for one state’s refusal to contribute to mitigation efforts has no direct effect one another state’s wellbeing. Meanwhile, retaliating against states who refuse to contribute would most likely involve issue linkage that dramatically increases the costs of such retaliation, thereby decreasing its appeal. Finally, it is not clear what types of benefits a state could accrue from developing a reputation as a “reliable contributor” to mitigation efforts. Unlike issue areas such as finance, where a reputation for repaying loans
attracts future creditors (Tomz, 2007), there are few material benefits associated with being conscientious about environmental protection.

Moreover, “compliance” is an elusive concept in the Paris Agreement. The treaty ambiguously commits signatories to engage in “ambitious efforts” to keep global temperature rise below two degrees Celsius above pre-industrial levels (United Nations, 2015), but it does not outline specific emissions reduction goals or other metrics of compliance. Hence, whether this design can lead to meaningful progress towards mitigating climate change remains an open question that is important to investigate both from an empirical and normative perspective.

3.2.1 Experimental Design

The experiment occurs within a game played between five teams of two subjects each and an additional group of three subjects. Therefore, each session involves 13 subjects. A total of 130 individuals participated in 10 sessions of the experiment. Subjects were recruited from undergraduates at Rice University in the Behavioral Research Laboratory subject pool.

Subjects earned experimental currency units (ECUs) during three periods of play, each period consisting of five rounds of the game. All players were informed at the outset that one of the three periods would be randomly selected for payment. The exchange rate was $0.08, or $0.80 for every 10 ECUs. In addition to the payment from the game, subjects received a show-up fee of $5.00. On average, subjects earned $19.23 for one hour of play.

The Game

Subjects were randomly assigned to the role of government leader or interest group leader and kept that role throughout the experiment. Government leaders were referred to as “yellow players” and interest group leaders as “green players” to avoid
introducing normative considerations. At the start of the experiment, subjects were asked to guess the number of dots on a computer screen. They were then assigned a color based on their response time during this task. There were eight government leaders (yellow players) and five interest group leaders (green players) in each session. Interest group leaders were randomly assigned to one of the five teams, and government leaders were randomly assigned either to one of the teams or to the additional group, which served as a candidate pool for replacing government leaders who were removed during the game. Therefore, each team consisted of one government leader and one interest group leader.

The single interest group leader in each team mimics the “pro-compliance” constituency in a given polity (Dai, 2007). The domestic actors that are most likely to play the role of directing pressure toward government leaders in response to information from international institutions are interest groups, such as NGOs and opposition groups. While there will likely be multiple such groups in any given polity, I represent these groups with a single leader in the experiment because they should all respond similarly to information from international institutions given the overlap in their interests.

Because my theory is about international institutions delivering information to pro-compliance constituencies, I do not include an “anti-compliance” actor in the experiment. While it is indeed the case that domestic actors opposed to treaty compliance could be mobilized by information from international institutions (see Dai, 2007; Chaudoin, 2016), treaty bodies and international NGOs directly engage with pro-compliance constituencies through processes such as periodic reporting (Zwingel, 2005). Additionally, survey evidence suggests pro-compliance groups have an advantage in mobilizing the general public because citizens generally respond positively to arguments that emphasize following through on international commitments (Wallace, 2013; Chaudoin, 2014b; Kreps and Wallace, 2016). Therefore, focusing on the rela-
The relationship between information and pro-compliance actors is an appropriate test of my theory.

The game is a modified version of a threshold public goods game sometimes referred to as a “collective-risk social dilemma,” where subjects contribute to a pool aimed at preventing a loss that arises if the target is missed (Milinski et al., 2008; Tavoni et al., 2011). Within each team, the government leader was solely responsible for contributing to the fund, but only the interest group leader suffered the consequences if the threshold was not met. This feature reflects the tendency for government leaders to enjoy a higher socioeconomic status than their constituents, which will allow them to withstand the consequences of climate change largely unscathed. For example, wealthier individuals can purchase property in desirable areas less affected by extreme weather and outfit their homes with expensive protective equipment.

At the start of the game, all players were informed that in each round, a single card would be drawn from a deck of 19 blue cards and 1 red card. If a red card is drawn, the interest group leaders keep all of their endowment for the round. If a blue card is drawn, the interest group leaders lose half their earnings. This extreme loss symbolizes the dramatic effects of climate change. However, if the government leaders collectively contribute at least 70 ECUs, then half of the cards in the deck will be replaced by red cards. This modification increases the interest group leaders’ chances of averting the loss from 5% (1 out of 20 cards) to 50% (10 out of 20 cards). As in climate change, meeting the threshold (i.e. successfully soliciting meaningful contributions from all countries) decreases the likelihood of a catastrophic outcome but does not eliminate the possibility. Government leaders contributed ECUs from their own earnings, but they were not affected by the draw.

Government leaders began each round of the game with 60 ECUs, and interest group leaders began with 50 ECUs. The game proceeded in four steps. First, the
government leaders decided how much to contribute to the fund. Each government leader selected between a large contribution of 20 ECUs, a smaller contribution of 10 ECUs, or the player could choose not to contribute. If the five government leaders collectively contributed at least 70 ECUs, half of the cards were replaced by red cards. If the five government leaders contributed any amount less than 70 ECUs, the deck remained unchanged. This design feature reflects the consensus among climate advocates that merely symbolic contributions will not suffice to avert the deleterious effects of climate change. Therefore, in the game, half-hearted contributions of just 10 ECUs from all five players were insufficient to meet the threshold. Still, to meet the 70 ECU threshold with maximum efficiency, some government leaders would need to contribute more than others. If all five contributed the maximum amount, 30 excess ECUs would be wasted. Hence, the game produced incentives to free ride on the contributions of others and also raised concerns about equality in contributions, both of which have been identified as impediments to cooperation aimed at mitigating climate change (Bechtel and Scheve 2013; Falkner 2016).

Interest group leaders were not informed of their government leader’s contribution decision, but all players were told whether the threshold was met. Information about government leaders’ contributions was withheld because interest groups are not always well-informed of their government’s efforts to meet their stated emissions reduction goals, either because this information is not readily available or because acquiring this information would be costly. It is more likely, however, that domestic actors will be aware of the general state of international efforts to mitigate climate change, which is commonly discussed in media reports.

Second, a card was drawn to determine whether the interest group leaders lose half of their earnings. All players were informed about which card was drawn. Next, each interest group leader evaluated their government leader’s performance. An interest group leader that was dissatisfied with its government leader could spend 5 ECUs
to try and replace them with another yellow player from the candidate pool. The cost of punishing a government leader reflects the costs interest groups must incur to express their dissatisfaction with a government, which can take the form of mobilizing voters, organizing protests, making campaign contributions to opposition candidates, or other political activities. If an interest group leader tried to remove its government leader, the government leader would lose 10 ECUs, and the interest group leader would roll a die to determine whether the government leader will be removed (1/3 chance) or continue to the next round (2/3 chance). Therefore, as in any polity, paying the costs of expressing dissatisfaction with a leader does not guarantee their removal.

Finally, any government leaders selected for removal were replaced by players randomly selected from the candidate pool. Players in the candidate pool completed a mundane task during each round to keep them occupied and accrue earnings comparable to those of the other players. Yellow players in the candidate pool counted the number of times a letter appears in a randomly-generated string of text and earned 30 ECUs for providing the correct answer. Hence, yellow players in the candidate pool always earned fewer ECUs than yellow players on a team. This ensured that yellow players will prefer being a government leader to being a candidate, which reflects the incentive structure of political office-seekers, where the benefits of holding office outstrip those of being a candidate.

**The Treatment**

The treatment involved an additional step before the interest group leaders decided whether to try and remove their government leaders. All players were informed that the experimenter would publish two lists of yellow players based on the quality of their contributions. Government leaders who contributed 20 ECUs or, if the threshold is met, contributed 10 ECUs would appear on the “A” list. Those who contributed
nothing or, if the threshold is not met, contributed only 10 ECUs would appear on the “B” list. Therefore, interest group leaders in the treatment condition received an additional piece of information regarding their government leaders’ contributions. Interest group leaders were only informed of their government leader’s placement, and they received no information about other leaders’ behavior.

To address the potential for ordering effects, subjects were randomly assigned to ABA or BAB designs, where A implies no treatment and B implies treatment. Half of the subjects played one period of five rounds in the control condition (no treatment) followed by one period with the treatment and a final period in the control condition. The other half played one period with the treatment followed by one period in the control condition and a final period with the treatment.

**Expectations**

The two outcomes of interest were the mean level of contributions government leaders made over the course of each period and the mean proportion of rounds in which interest group leaders attempted to remove their government leaders in each period. Based on the expected relationship between information provision, domestic mobilization, and compliance with international commitments detailed above, I preregistered two hypotheses stipulating that both measures should be greater under the treatment condition:

\[ H_1: \text{Government contributions will be higher when the review procedure is in place.} \]

\[ H_2: \text{Interest group punishments will occur more frequently when the review procedure is in place.} \]

However, upon further thought, I added a third hypothesis related to changes in interest group behavior under the treatment condition. The logic laid out in the theory suggests government contributions should be higher when the review procedure is
in place, because government leaders expect interest groups will use the information
provided in the review to punish low contributions. In the control condition, interest
group leaders cannot distinguish between government leaders who contributed the
maximum amount and those that contributed nothing when the threshold was not
met. Hence, government leaders in the control condition should have an equal risk
of removal, regardless of the amount they contribute. By contrast, in the treatment
condition, government leaders who contribute greater amounts to the fund and there-
fore appear on the “A” list should experience fewer attempts at removal compared to
those who contribute nothing. In other words, interest group leaders in the treatment
condition should become more selective about when they pay the costs to try and re-
move a government leader and condition these decisions on government contributions.
This expectation is expressed in the following hypothesis:

\[ H_3: \text{When the review procedure is in place, the likelihood of interest group punishment will decrease as government contributions increase.} \]

3.3 Results

On average, subjects designated as government leaders contributed greater amounts to
the fund in the treatment condition (9.53 ECUs) compared to the control condition
(5.28 ECUs), and this difference is statistically significant at the 0.01 level (t = -4.57, df = 190). Figure 3.1 depicts the mean levels of contributions broken down
by period and session type, with 95% confidence intervals. Recall that subjects in
each session played five rounds in each of three periods, alternating for each period
whether the treatment was included. In the ABA sessions, subjects played one period
in the control condition followed by one period in the treatment condition and a final
period in the control condition. In the BAB sessions, subjects first played one period
in the treatment condition followed by one period in the control condition and a final
In both session types, government contributions were substantially higher when the review procedure was in place, lending support to $H_1$. In the ABA sessions, government leaders contributed approximately 6.19 ECUs per round during the initial period, when the review procedure was absent. Contributions rose slightly after the review procedure was introduced in the second period to 7.47 ECUs, but this increase was not statistically significant ($p = 0.42$). However, contributions then dropped precipitously in the final period to 2.38 ECUs, a highly significant difference ($p < .01$). This decrease suggests that removing the review procedure after subjects have experienced playing the game with this information reduces contributions. This pattern appears again in the BAB sessions, where contributions dropped from 10.47 ECUs in the first period to 7.20 ECUs in the second period when the review procedure was removed. Contributions bounced back to 10.68 ECUs in the final period as the review procedure was reinstated. Although the confidence intervals overlap, the difference
between contributions in the first and second periods is statistically significant at the 0.1 level ($p = 0.06$), and the difference between the second and third periods is statistically significant at the 0.05 level ($p = 0.04$). The test statistics from all pairwise comparisons are summarized in the appendix.

Overall, the picture that emerges from these results is broadly supportive of $H_1$. Government contributions were indeed higher when the review procedure is in place. Furthermore, government leaders were substantially more likely to contribute at least some ECUs to the fund in the treatment condition compared to the control condition. When the review procedure was absent, government leaders refused to contribute any ECUs to the fund in 53 out of 100 instances. By contrast, when the review procedure was in place, there were only 15 cases of government leaders contributing no ECUs out of 92 leader-period observations. Because of the higher government contributions in the treatment condition, the threshold was also reached far more often during the treatment periods. Whereas the threshold was met in only 4 of 75 rounds in the control condition (5.33%), the threshold was met in 19 of 75 rounds in the treatment condition (25.33%).

Looking to the rates of interest group punishment, the frequency of punishment in the treatment condition is roughly the same as that observed in the control condition. Figure 3.2 displays the mean rates of punishment by period and session type. In both ABA and BAB sessions, interest group leaders punished at nearly identical rates in the first two periods, and rates of punishment declined in the final period. This pattern makes sense, since punishment is less valuable as the game draws to a close and the end to repeated interaction nears. Indeed, the only difference of means test that reaches conventional levels of statistical significance is the pairwise comparison between the second and third periods in the BAB session. The differences between all other pairs of periods do not reach conventional levels of statistical significance.¹

¹See appendix for all pairwise comparisons.
Although the treatment did not increase overall rates of punishment as predicted in $H_2$, it does seem that the review procedure increased the selectivity of interest group punishments, as predicted in $H_3$. Recall that interest group leaders received no information about their government leader’s contributions in the control condition. Thus, they could not distinguish between leaders who contributed the maximum amount and those that contributed nothing when the threshold was not met. Out of 375 observations in the control condition, there were 21 cases in which interest group leaders elected to punish leaders who had contributed all 20 possible ECUs and 187 cases in which interest group leaders refrained from punishing government leaders who contributed nothing. In the treatment condition, however, government leaders that contributed all 20 possible ECUs appeared on the “A” list, and those that contributed nothing appeared on the “B” list, which provided interest group leaders with some information about the quality of their contributions. As a result, in the treatment condition, there was only one instance of an interest group leader punishing...
a government leader who had contributed the maximum amount and just 75 cases in which interest group leaders did not punish government leaders who refused to contribute.

To further explore the nuances of interest group punishments in the treatment and control conditions, I estimate multilevel logistic regressions where the unit of analysis is interest group leader-round. The dependent variable is a binary indicator coded 1 for each round in which the interest group leader attempts to remove its leader and 0 otherwise. The main independent variables of interest are the treatment, a binary variable coded 1 for each round in which the review procedure was in place and 0 otherwise, and contribution, a measure of the number of ECUs contributed by the government leader paired with the relevant interest group leader in each round. Recall that this variable can take on three values: 0, 10, or 20. In a separate model, I also include an interaction between the treatment and contribution to assess whether the relationship between contributions and punishment is different in the treatment and control conditions, as predicted by $H_3$. In addition to these independent variables, I include fixed effects for session, period, and round, as well as random intercepts by interest group leader. These mixed effects should account for stochastic dependence between observations related to the same interest group leader, round of play, period, or session.

The results from this analysis appear in Table 3.1. The coefficients for the session, period, and round fixed effects are omitted due to space constraints. The first model corroborates the results from the difference of means tests above—the correlation between the treatment and interest group punishments is not statistically significant at conventional levels. Model 1 also suggests a negative correlation between contributions and punishment, which implies interest group leaders are less likely to punish government leaders who contribute greater amounts to the collective fund. However, once the interaction is included in Model 2, we see that this relationship is driven
Table 3.1: Multilevel Logistic Analysis of Interest Group Punishment

<table>
<thead>
<tr>
<th></th>
<th>Punishment (1)</th>
<th>Punishment (2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>0.261</td>
<td>1.352***</td>
</tr>
<tr>
<td></td>
<td>(0.197)</td>
<td>(0.257)</td>
</tr>
<tr>
<td>Contribution</td>
<td>-0.087***</td>
<td>-0.005</td>
</tr>
<tr>
<td></td>
<td>(0.014)</td>
<td>(0.017)</td>
</tr>
<tr>
<td>Treatment * Contribution</td>
<td>-0.186***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.028)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>-1.062*</td>
<td>-1.567***</td>
</tr>
<tr>
<td></td>
<td>(0.553)</td>
<td>(0.559)</td>
</tr>
</tbody>
</table>

Observations: 750
Log Likelihood: -374.377, -349.419
Akaike Inf. Crit.: 786.753, 738.837
Bayesian Inf. Crit.: 874.534, 831.239

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

by interest group leaders in the treatment condition. The positive coefficient for the treatment variable in Model 2 suggests interest group leaders are more likely to punish government leaders who contribute nothing in the treatment condition compared to those that contribute 10 or 20 ECUs. By contrast, in the control condition, there is no statistical relationship between contributions and punishment.

The substantive meaning of the interaction between the treatment and government contributions is depicted in Figure 3.3. The predicted probabilities are calculated using the observed value approach recommended by Hanmer and Kalkan (2013). The lines denote 95% confidence intervals, computed through simulation. The left graph shows the null relationship between government contributions and the likelihood of interest group punishment described above. However, the right graph depicts a steep negative relationship between contributions and punishment. Interest group leaders in the treatment condition whose government leaders contribute nothing to the fund are nearly 70% more likely to attempt removing their leaders compared to those in the control condition paired with leaders who refuse to contribute. Moreover, interest
group leaders in the treatment condition paired with government leaders who contribute all 20 ECUs to the fund punish at less than a quarter of the rate at which interest group leaders in the control condition assigned to government leaders who contribute the maximum amount punish. Interestingly, when government leaders contribute just 10 ECUs in the treatment condition, and their list placement is dependent upon whether the threshold is met, interest group leaders punish at roughly the same rate as those in the control condition at all levels of contributions. Together, these results strongly suggest interest groups use the information from the review procedure to decide when to pay costs to remove a leader, which offers compelling support for $H_3$.

All in all, the results provide support for the argument that information provision facilitates domestic mobilization that ultimately leads to more cooperative outcomes. Contributions were substantially higher when the review procedure was in place, as
predicted by $H_1$. Contrary to $H_2$, interest group punishments occurred at similar rates in both the treatment and control conditions, but further analysis reveals that interest group leaders in the treatment condition used information from the review procedure to selectively punish government leaders that did not contribute to the collective fund, as predicted by $H_3$. These findings highlight the value of information in motivating cooperative behavior.

3.4 Discussion

This paper shows that subjects in a controlled laboratory environment behave in ways consistent with the expectations derived from a theory about the effects of information from international institutions. I chose a laboratory experiment to test these expectations in order to sidestep the complex challenges to causal inference inherent in analyses that rely on observational data. Nevertheless, I recognize that this approach has its own limitations. Scholars often suggest that selecting between experimental and observational research designs entails a trade-off between internal and external validity. While an experiment may offer clear evidence of a causal relationship between two variables, the researcher may not be able to generalize conclusions derived from an experiment to other populations or situations. In the context of the present paper, this concern manifests as the question of whether one can draw conclusions about the relationship between information from international institutions and cooperation based on the results from this study.

Indeed, the experiment conducted in this study differs considerably from the sequence of events that would normally occur outside of the laboratory. For example, governments and interest groups rarely consider a single policy dimension in isolation. Rather, environmental protection competes with other pressing political issues, such as government spending, taxes, and regulation, for the attention of political actors during elections and policymaking. Aside from these contextual issues, the
experiment also differs in terms of the characteristics of the actors involved. College students are typically younger, more highly educated, and more liberal than the general population. Moreover, college students are not steeped in politics and strategy in the same way as actual political elites.

Yet while there is no doubt that this experiment is indeed a unique environment involving a limited sample, it still has value due to its capacity to produce insights into the causal relationships of interest. External validity, or the extent to which a finding can be generalized to other populations or contexts, has little meaning without internal validity, or confidence that there is indeed a causal relationship between the independent and dependent variables (McDermott, 2011). Importantly, the experiment in this study was designed to maximize experimental realism, or the extent to which subjects believe and engage with the situation they are confronting (McDermott, 2002). In other words, the experiment was designed to capture the main aspects of the phenomenon of interest in a way that interests and engages the subjects. Meanwhile, other features that could make the scenario appear more realistic but would ultimately distract subjects and undermine internal validity were omitted. Additionally, while the student sample was selected out of convenience, Druckman and Kam (2011) show that relying on student samples does not threaten external validity in cases where the theory predicts a homogeneous treatment effect.

Thus, I am confident that this study, despite its limitations, can still deliver useful insights into the world outside of the laboratory. Ultimately, external validity can only be established through replications across different settings and populations (McDermott, 2011), but establishing internal validity, as this study has done, is the critical first step in this scientific endeavor.
3.5 Conclusion

The question of how states can be motivated to sustain large-scale multilateral cooperation is a subject of intense interest for scholars and practitioners of international diplomacy and cooperation alike. Testing theories about the utility of different institutional configurations, however, remains difficult, as relatively few designs come into being and both institutional design and membership are determined by self-interested actors considering their likely future behavior. This paper provides supporting evidence for the argument that information provision facilitates domestic mobilization, which in turn leads to greater cooperative behavior. The experiment used here is modeled after the Paris Agreement, which includes a review procedure in which technical experts pass judgment on states’ emissions reductions at regular intervals. The review procedure is a critical source of information for domestic audiences, who can use this data to hold governments accountable to their stated commitments to slow the rise of global temperature. Results from the experiment indicate that government leaders contribute greater amounts to a collective fund when the review procedure is in place, and interest group leaders condition their behavior on government contributions when this information is available.

The implications of this research, therefore, are most directly applicable to the case of the Paris Agreement. While the Paris Agreement was initially hailed for its universal support among countries around the world, skepticism soon arose over its ability to produce meaningful emissions reductions. Observers wondered aloud whether the agreement was too weak to induce behavioral change, given that the language was vague and allowed states to design their own mitigation commitments subject only to a periodic review process that did not impose any punishment on states that failed to meet their pledges (e.g., Linos and Pegram 2016a). This study suggests the review procedure can empower domestic constituencies to pressure their
governments to fulfill their commitments. These results offer empirical backing for
supporters of the agreement’s design, who envisioned the institution working in part
through information provision to domestic audiences (Keohane and Oppenheimer
2016).

This paper also has broader implications for the study and practice of international
cooporation. As noted, scholars face complicated threats to inference when trying to
understand the mechanisms that facilitate or hinder cooperative enterprises on the
international stage. The experiment in this study provides additional support for
arguments that rely on domestic mobilization as the main explanation for compliance
with international agreements (Dai 2007; Simmons 2009). Moreover, this paper helps
clarify the conditions under which such mobilization is likely to occur and be effective.
Institutions that provide useful information enable domestic actors to identify the
appropriate timing for the application of political pressure and effectively tie political
support to policy output. For practitioners of international cooperation, then, this
paper suggests negotiators ought to focus on ensuring institutions are designed with
mechanisms for information production and dissemination. Whereas states may balk
at the prospect of stringent enforcement, information-producing provisions may be
less controversial and lead to greater participation without loss of effectiveness.

Finally, this paper points to multiple directions for future research. The experi-
mental design presented here can be easily modified to allow comparison of infor-
mation provision with other mechanisms of enforcement, such as fines and dispute
settlement procedures. Researchers could also incorporate elements of endogenous
institutional design by allowing leaders to select from different institutional config-
urations. These modifications would more closely emulate the process by which in-
ternational institutions arise and states interact with institutional procedures after
their formation. The laboratory offers an ideal setting in which to study these phe-
nomena and supplement observational analyses with results from a more controlled
environment. These studies would substantially contribute to our understanding of the domestic sources of international cooperation and the conditions under which these domestic mechanisms are activated.
### 3.6 Appendix 1: Pairwise Comparisons for Leader Contributions

<table>
<thead>
<tr>
<th>Comparison</th>
<th>First Group Mean</th>
<th>Second Group Mean</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA - ABA</td>
<td>6.19 (6.37)</td>
<td>7.47 (6.29)</td>
<td>t = 0.81, df = 62, p = 0.42</td>
</tr>
<tr>
<td></td>
<td>n = 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AB - ABA</td>
<td>2.38 (4.93)</td>
<td>10.47 (6.38)</td>
<td>t = -3.58, df = 62, p &lt; 0.01</td>
</tr>
<tr>
<td></td>
<td>n = 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAB - BAB</td>
<td>10.47 (6.29)</td>
<td>7.20 (7.52)</td>
<td>t = -1.91, df = 60, p = 0.06</td>
</tr>
<tr>
<td></td>
<td>n = 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BAB - BAB</td>
<td>7.20 (7.52)</td>
<td>10.68 (5.69)</td>
<td>t = 2.09, df = 60, p &lt; 0.05</td>
</tr>
<tr>
<td></td>
<td>n = 34</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABA - BAB</td>
<td>6.19 (6.37)</td>
<td>10.47 (6.38)</td>
<td>t = 2.71, df = 63, p &lt; 0.01</td>
</tr>
<tr>
<td></td>
<td>n = 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABA - BAB</td>
<td>2.38 (4.93)</td>
<td>10.68 (5.69)</td>
<td>t = 6.09, df = 56, p &lt; 0.01</td>
</tr>
<tr>
<td></td>
<td>n = 33</td>
<td></td>
<td></td>
</tr>
<tr>
<td>ABA - BAB</td>
<td>7.47 (6.29)</td>
<td>7.20 (7.52)</td>
<td>t = -0.16, df = 63, p = 0.87</td>
</tr>
<tr>
<td></td>
<td>n = 31</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3.2: Summary of pairwise comparisons for leader contributions. The groups compared in each row are marked in bold in the first column. Standard deviations are provided in parentheses. The numbers of subjects vary because yellow players that spend all five rounds in a period in the candidate pool are excluded from the calculations.
3.7 Appendix 2: Pairwise Comparisons for Interest Group Punishments

<table>
<thead>
<tr>
<th>Comparison</th>
<th>First Group Mean</th>
<th>Second Group Mean</th>
<th>t-test</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABA - ABA</td>
<td>0.36 (0.26)</td>
<td>0.36 (0.29)</td>
<td>t = 0.00, df = 47, p = 1.00</td>
</tr>
<tr>
<td>ABA - ABA</td>
<td>0.36 (0.29)</td>
<td>0.24 (0.30)</td>
<td>t = -1.63, df = 48, p = 0.11</td>
</tr>
<tr>
<td>BAB - BAB</td>
<td>0.28 (0.22)</td>
<td>0.29 (0.25)</td>
<td>t = 0.12, df = 48, p = 0.90</td>
</tr>
<tr>
<td>BAB - BAB</td>
<td>0.29 (0.25)</td>
<td>0.15 (0.19)</td>
<td>t = -2.21, df = 45, p &lt; 0.05</td>
</tr>
<tr>
<td>ABA - BAB</td>
<td>0.36 (0.26)</td>
<td>0.28 (0.22)</td>
<td>t = -1.17, df = 47, p = 0.25</td>
</tr>
<tr>
<td>ABA - BAB</td>
<td>0.22 (0.30)</td>
<td>0.15 (0.19)</td>
<td>t = -1.03, df = 40, p = 0.31</td>
</tr>
<tr>
<td>ABA - BAB</td>
<td>0.36 (0.29)</td>
<td>0.29 (0.25)</td>
<td>t = -0.94, df = 46, p = 0.35</td>
</tr>
</tbody>
</table>

Table 3.3: Summary of pairwise comparisons for interest group punishments. The groups compared in each row are marked in bold in the first column. Standard deviations are provided in parentheses.
Chapter 4

Foreign-Funded NGOs and Domestic Enforcement of Treaty Commitments

In issue areas such as human rights and environmental politics, non-governmental organizations (NGOs) play a critical role in the functioning of international treaties by enabling domestic enforcement of treaty commitments. In particular, NGOs deliver information about government actions to the public in order to alert citizens when their government is in violation of treaty obligations and motivate a response. Previous studies show that individuals generally disapprove of government actions that violate international law, provided they are made aware of the violations (Wallace, 2013; Chilton, 2014; Kreps and Wallace, 2016; Chaudoin, 2014). As messengers of this critical information, NGOs stimulate and coordinate the domestic mobilization that scholars recognize as critical to making global cooperation work (Simmons, 2009; Dai, 2007; Davis, Murdie and Steinmetz, 2012).

In recent years, however, governments have imposed stricter restrictions on NGOs, especially those that rely on foreign funding to finance their operations (Christensen and Weinstein, 2013; Dupuy, Ron and Prakash, 2016; Bakke, Mitchell and Smidt, 2019). For example, Russia notoriously passed a law in 2012 that requires any organizations that receive foreign funding to register with the government as “foreign agents” (Flikke, 2016). In 2017, Hungary began requiring organizations that receive
substantial foreign funding to register with the government and explicitly label themselves as “civic organizations funded from abroad” in all their publications (Amnesty International, 2019). A similar law took effect in Israel in 2018 (Krebs and Ron, 2018a). Meanwhile, scholars and activists have noted a growth in the vilification of NGOs that receive foreign funding. Across continents, NGOs report being accused of betraying their countries and working on behalf of foreign interests (e.g., Ron and Crow, 2015; Kiai, 2013; Doshi, 2016). These trends raise questions about the continued ability of NGOs to facilitate domestic enforcement of international obligations.

This paper investigates how labelling NGOs as foreign-funded influences their ability to alert citizens to violations of international treaties. Laws that require NGOs to specify in their publications whether they receive foreign funding and broad-based criticism of foreign-funded NGOs are particularly worrisome to the NGO community because many believe they stigmatize and delegitimize foreign-funded NGOs in the eyes of the public. Yet aside from anecdotal reports of such effects, there is little systematic research on the topic. I argue that identifying an NGO as foreign-funded reduces the organization’s credibility as a judge of the consistency between government policy and international law. Thus, individuals are less likely to believe their government is violating international law when the accusation comes from a foreign-funded NGO compared to an NGO that relies exclusively on domestic funding or one whose funding sources remain ambiguous. Because fewer individuals believe the message, the expected negative effect of informing citizens that a government policy violates international law on attitudes toward that policy is lessened.

I use a set of two survey experiments to test these expectations. Based on the logic outlined above, I predict that while accusations that a policy violates international law will generally depress support for that policy, this effect will be smaller when the NGO making the accusation is identified as foreign-funded compared to when the NGO is identified as domestically-funded or its funding sources are not specified.
I test these hypotheses in two identical experiments that vary the issue area of the policy in question. I focus on policies in environmental politics and human rights, since the NGOs that are targeted by anti-foreign funding laws and rhetoric tend to work in these areas.

The results from the experiments provide partial support for this argument. When the policy in question is an environmental issue, the results are commensurate with my expectations about the effect of information about NGO funding sources. In general, NGO accusations that the policy violates international law reduces public support for that policy, and this effect remains when the NGO is identified as domestically-funded. When the NGO is identified as foreign-funded, however, the NGO’s accusations have no effect on respondents’ attitudes toward the policy. Yet when the policy in question is a human rights issue, there is no statistical difference in levels of support across the experimental conditions, which suggests information about international law, regardless of the source, has little impact on attitudes toward that policy. A brief exploratory analysis suggests the relationship between information about NGO funding sources and the effects of information about international law on public attitudes depends heavily on an individual’s ideological disposition. These mixed results highlight the importance of testing this argument in different contexts and suggest further research should be done to clarify the scope conditions of the theory.

This paper makes several important contributions to the study and practice of NGO advocacy and international law. As government restrictions on NGOs have spread around the world, scholars are just beginning to understand the roots of this trend and assess its effects (Christensen and Weinstein 2013, Dupuy, Ron and Prakash 2016, 2015, Dupuy and Prakash 2018, Bakke, Mitchell and Smidt 2019). Previous work has examined the impact of foreign funding restrictions on foreign aid and the activities of NGOs, but little has been done to investigate how these laws
and the discourse surrounding them impact the relationship between NGOs and the citizens they seek to reach in their campaigns. This study contributes valuable insight into the likely consequences of these developments along with useful advice for NGOs adapting to new restrictions.

More broadly, this paper also contributes to the study of international law by highlighting the strategies governments use to undermine mobilization that could otherwise force them to make policy changes (Bakke, Mitchell and Smidt, 2019). Recent work points to a variety of ways in which government repression and domestic contestation can attenuate the impact of international law on state behavior (Bakke, Mitchell and Smidt, 2019; Chaudoin, 2016; Bracic and Murdie, 2019). By focusing on government targeting of foreign-funded NGOs, this study illuminates an additional contextual feature that may condition the effects of international treaties.

This paper proceeds as follows. First, I discuss the role of NGOs in the domestic enforcement of international treaties and consider how the recent spread of restrictions on NGOs and criticism of foreign funding threaten to weaken this dynamic. I build on theories of source credibility in political communication to explain how labelling NGOs as foreign-funded can reduce their ability to influence public opinion about treaty violations. Next, I describe the experimental approach to test the hypotheses derived from the theory. Finally, I discuss the results of the experiments and their implications.

4.1 NGO Advocacy, Compliance, and Foreign Funding

NGOs play an important role in the functioning of international treaties because they engage domestic audiences with treaty obligations and motivate public pressure for compliance. Many scholars believe domestic political dynamics are a crucial explanatory factor in patterns of compliance with treaty commitments, especially in the context of environmental politics and human rights, where traditional mechanisms of
enforcement are less applicable (Simmons, 2009; Mansfield, Milner and Rosendorff, 2002; Dai, 2007). For example, Simmons (2009) argues that mobilization among citizens raises the costs of noncompliance with human rights commitments. In a similar vein, Dai (2007) argues that pro-compliance groups leverage information from international institutions to hold governments accountable to environmental and human rights treaty commitments.

NGOs represent a crucial link in the causal chain connecting international treaties to domestic audiences and finally to government behavior because they are often the central mechanism through which domestic audiences become aware of treaty commitments and government deficiencies in meeting these obligations. Generally speaking, an NGO is “any non-state, non-profit, and non-criminal organization” (Davies, 2019, 2). For the purposes of this paper, I narrow this definition further to include only those groups whose main focus is advocacy. This excludes organizations such as business associations, development aid organizations, and professional groups. While some NGOs, such as Greenpeace and Amnesty International, are large international organizations with offices around the world, others are entirely domestic organizations operating within a single state. For instance, Gorakhpur Environmental Action Group (GEAG) is an Indian NGO that campaigns for farmers’ rights and climate change resilience (Gorakhpur Environmental Action Group, 2019).

NGOs connect treaties to domestic audiences by collecting, framing, and disseminating information about government commitments. For example, NGOs helped galvanize support for policies to address discrimination and violence against women in part by adapting international ideas to local contexts (Merry, 2006; Htun and Weldon, 2012). In environmental politics, NGOs stimulated public support for government compliance with the International Convention for the Regulation of Whaling by re-framing whales as sentient animals deserving of protection (Wapner, 1995; Eilstrup-Sangiovanni and Bondaroff, 2014). Through the process of interpreting and
framing international obligations, NGOs explain to citizens how international treaties are relevant to them and why they should support compliance.

As messengers of treaty-related information, NGOs also alert citizens when mobilization may be necessary to pressure the government to better align its policies with existing international obligations. A vibrant literature on NGO “naming and shaming” suggests highlighting gaps between treaty commitments and government policy can be a fruitful strategy for instigating policy change (Murdie and Davis 2012; Murdie and Peksen 2015; Hendrix and Wong 2013; Ausderan 2014). Oftentimes, NGOs conduct original research to substantiate their accusations of treaty violations. For example, two Israeli NGOs released a report detailing alleged human rights violations at a prison in southern Israel in 2015 based on extensive interviews and analyses of medical records (Kadman 2015). NGOs are also a major source of global performance indicators such as the Press Freedom Index and the Climate Analysis Indicators Tool, which shift public debate towards comparisons with other countries to invoke concerns over state reputations (Kelley and Simmons 2019). Overall, these tactics create desires for mobilization that lower citizens’ costs of collective action. NGOs then drive these costs down even further by providing the organizational infrastructure necessary to coordinate action (Boulding 2010).

In sum, NGO advocacy can lead to greater treaty compliance through its effects on public opinion and broader domestic mobilization. NGOs educate domestic audiences about treaty content and build public support for compliance. When governments fail to fulfill their international obligations, NGOs alert citizens to these gaps in compliance and motivate them to take action. These actions can take a variety of forms, such as petitions, protests, and shifting electoral support toward opposition candidates. Each of these moves increases the government’s costs of noncompliance, which increases the probability that the government will change policy to better comport with its treaty commitments.
Importantly, however, the effects of NGO advocacy on treaty compliance are highly conditional. Regime type, the strength of civil society, and the distribution of preferences among citizens, to name a few, have all been identified as crucial moderators of this relationship (Simmons, 2009; Dai, 2007; Hendrix and Wong, 2013; Chaudoin, 2014b). Along with these contextual factors, governments also engage in regulation and sometimes violent repression of NGOs to curb their influence.

In particular, scholars and activists have recently documented a marked increase in legal restrictions on the activities of NGOs, especially those that rely on foreign funding sources (Dupuy, Ron and Prakash, 2016; Christensen and Weinstein, 2013; Carothers and Brechenmacher, 2014; Amnesty International, 2019; Musila, 2019; Bakke, Mitchell and Smidt, 2019). Many advocacy NGOs, who often struggle to raise sufficient funds from domestic sources, fall into this category. These laws include measures that create new barriers to entry for foreign-funded NGOs, impose burdensome reporting requirements, and explicitly limit the scope of activities in which they can engage. Recent work by Bakke, Mitchell and Smidt (2019) suggests these types of restrictions are directly tied to NGOs’ efforts to hold governments accountable to international commitments; they show that states impose more restrictions on NGOs when they are both party to human rights treaties and routinely commit human rights abuses. Data collected by Dupuy, Ron and Prakash (2015) shows that while fewer than 10 percent of countries passed restrictive laws related to foreign-funded NGOs before the end of the Cold War, 44 percent had done so as of 2012.

Along with new restrictions, foreign-funded NGOs are increasingly facing fierce criticism from government and media sources on account of their foreign donors. Governments that seek to restrict foreign funding of NGOs often accuse such organizations of improperly meddling in their domestic affairs and justify restrictive policies in terms of defense of national sovereignty (Poppe and Wolff, 2017; Matejova,
Parker and Dauvergne 2018; Carothers and Brechenmacher 2014). For example, the Venezuelan government repeatedly claimed that a 2010 law prohibiting foreign funding of pro-rights NGOs was intended to undercut US attempts to destabilize the country (Gill 2017). Before unveiling new legislation that required NGOs to identify their foreign benefactors, the Hungarian government mailed citizens a questionnaire including a section that stated “organizations...funded from abroad are working in Hungary with the goal of meddling in domestic affairs in a untransparent way. Those operations can be a danger to our independence,” and asked for respondents’ opinions about what the government should do about their “hazardous operations” (Bienvenu and Karasz 2017). In India, Prime Minister Narendra Modi has accused NGOs of acting as “fronts for foreign lobby organizations” (Doshi 2016). Similarly, the Israeli Prime Minister Benjamin Netanyahu defended his governments reporting requirements for foreign NGO funding as a sensible attempt to prevent foreign meddling in the country’s internal affairs (Beaumont 2016).

NGOs also appear to be well aware that accepting foreign funding can expose themselves to vociferous attacks. In interviews with human rights workers, Ron and Crow (2015) found some respondents reported being accused of spying for foreign governments or betraying their home governments. In Kenya, critics of foreign-funded NGOs commonly refer to civil society as “evil society” and demonize human rights activists as “traitors” acting at the behest of foreigners (Kiai 2013). The 2017 USAID Civil Society Organization Sustainability Index for Eastern Europe and Eurasia describes how campaigns that vilify foreign-funded NGOs have increased in recent years and led to a deterioration in the public image of civil society in several countries across Central and Eastern Europe (USAID 2018).

How does the shifting information environment surrounding advocacy by foreign-funded NGOs condition the relationship between NGO advocacy and public opinion that is critical to the enforcement of treaty commitments? Does labelling an NGO
as foreign-funded reduce its ability to shape public attitudes toward policies that violate international law? To date, no existing work has examined these questions. Previous studies using survey experiments show citizens generally respond negatively to hearing that their government has violated an international obligation (Wallace, 2013; Chilton, 2014; Kreps and Wallace, 2016; Chaudoin, 2014b), but these works typically hold constant the source of this information. One exception is [Kreps and Wallace (2016)], who investigate whether responses to information about compliance with international law vary depending on whether the source is the government, an international organization (IO), or an NGO. They find that messages from the United Nations (UN) and Human Rights Watch (HRW), a large international NGO, have greater effects on public opinion compared to the government. Messages from the UN were slightly more effective than those from HRW, but the difference between the two was not statistically significant. This work suggests individuals are sensitive to the source of information related to compliance with international law, but it leaves open the question of how citizens respond to NGOs with different funding sources.

Some recent work has also begun to examine the effectiveness of NGO “shaming” in the context of government attempts to deflect and deny allegations of human rights abuses. [Bracic and Murdie (2019)] use a survey experiment to show that individuals are less supportive of an NGO-led campaign to free a political prisoner if the government labels the prisoner a terrorist. This study illustrates how governments can undermine NGO efforts to shape public opinion. In this example, however, the government attacks the content of the message rather than the organization delivering it. It is unclear how attempts to discredit the source of the information would influence its effects.

Therefore, in light of recent restrictions and criticism of foreign-funded NGOs, it is important to examine to what extent information about funding sources conditions the relationship between NGO advocacy and public opinion in the context of
compliance with treaty obligations. Because NGOs act as critical messengers of information about treaty commitments and potential violations, tactics that undermine NGO credibility pose a threat to this mechanism of enforcement. Understanding this dynamic will therefore deepen our understanding of the relationship between international law and state behavior and provide useful guidance for NGOs active in this area.

4.2 NGO Funding and Attitudes toward Treaty Violations

I argue that identifying an NGO as foreign-funded reduces its ability to communicate with domestic audiences about government compliance with international treaties because individuals view foreign-funded organizations as more biased and therefore less credible sources of information. Thus, the expected effect of information about treaty violations on attitudes toward government policy does not materialize because the recipients do not believe the message.

Scholars of political communication have long recognized that the effect of political messaging on public attitudes depends heavily on the source of the message (Druckman, 2001; Pornpitakpan, 2004; Lupia, 2013). Opinion change can only occur if the individual decides to accept new information (Zaller, 1992), and acceptance depends on the individual’s perception of the degree to which the information is relevant. To evaluate the relevance of political messages, individuals assess not only the content of the message but also the characteristics of the messenger. Individuals are more likely to accept a message if they perceive its source as trustworthy and credible. In a wide variety of contexts, studies consistently show that individuals respond differently to similar information attributed to different sources (e.g., Alt, Lassen and Marshall, 2016; Schatz and Levine, 2010; Botero et al., 2015; Boettcher and Cobb, 2009).

In the context of accusations of treaty violations, NGOs derive credibility from being perceived as relatively unbiased sources of information. A biased source is one
that is predisposed to favor a particular point of view. Therefore, it has an incentive to withhold or alter information that would undermine support for its preferred viewpoint. For example, a ruling government will be more likely to present information that paints itself in a positive light and bar the release of unflattering information in order to increase its chances of staying in power. Likewise, a business is predisposed to release only the information that will lead consumers to buy more of its products.

By contrast, an unbiased source is one that has no incentive to deliver untruthful information or censor its messaging in order to promote a particular viewpoint. An independent NGO whose mission is to identify human rights abuses, for example, has an incentive to state that abuses are occurring only when that is indeed the case.

It is important to note that bias is a relative, continuous measure. Few, if any, entities are completely unbiased. Even in the context of NGOs, scholars have documented cases in which human rights NGOs ignore some abusers to focus on other targets or exaggerate allegations of abuse (Meernik et al., 2012; Ron, Ramos and Rodgers, 2005; Hill Jr., Moore and Mukherjee, 2013). Like other political actors, NGOs are strategic actors that deliberately curate their messaging to achieve their policy objectives. The critical point, however, is that NGOs are typically perceived as relatively unbiased compared to other sources of information about treaty commitments, such as government entities. The 2019 Edelman Trust Barometer shows that individuals across the globe consistently rate NGOs as more trustworthy compared to the government and the media, and work on media trust shows that perceptions of trust are intimately related to perceptions of bias and accuracy (Edelman Intelligence, 2019; Knight Foundation, 2018).

Yet because NGOs are highly dependent on donors for the financial support that keeps them running, they also have incentives to behave in ways that “keep the donors happy” to prevent them from withdrawing support. Hence, whether an NGO is perceived as unbiased depends heavily on the characteristics of its donors. For example,
consumers of ethical products view industry-funded NGOs that claim to monitor whether a product is indeed produced ethically as less credible than independently-funded NGO monitors (Nooruddin and Sokhey 2012). In this case, the industry-funded NGO is perceived to be biased toward certifying products whose sale will profit the organization’s donors.

Translating this logic to the context of donors’ national origins, I argue that citizens will be less affected by messages attributed to foreign-funded NGOs because they view these organizations as more biased and therefore less credible compared to other NGOs. Although some argue that foreign benefactors provide a salutary service in that they help amplify the voices of groups who may be underrepresented in the policymaking process (Krebs and Ron 2018), citizens are unlikely to view this process as wholly benevolent. Rather, citizens are likely to recognize that foreign actors would only fund groups that advocate for policies they support. Hence, funding is more likely to go to groups that promote policies in line with the interests of their foreign benefactors. Consequently, citizens are more likely to see foreign funding of NGOs as a mechanism through which states or other international actors seek to influence the policies of sovereign states rather than as a benign attempt to remedy inequities in the distribution of political resources.

There is ample anecdotal evidence in line with this view in the examples discussed earlier. The frequency with which foreign-funded NGOs are referred to as “spies,” “traitors,” or “foreign lobbyists” suggests citizens are amenable to the view that foreign donors direct advocacy NGOs toward activities that promote foreign interests over those of their home government. For example, the Venezuelan government’s claim that pro-rights NGOs are working on behalf of the US to destabilize the country implies these organizations are biased toward relaying messages that will serve US interests to Venezuela’s detriment.

Of course, which foreign donors are viewed as nefarious will vary by country.
In Venezuela, the government explicitly linked foreign funding of NGOs to the US. In Israel, criticism of foreign-funded NGOs often centers around European Union (EU) funding for pro-Palestinian organizations, especially those that support the Boycott, Divestment, and Sanctions (BDS) movement [Ahren 2019]. In Hungary, the government zeroed in on US billionaire George Soros and his support for liberal NGOs that aid migrants [Nelson 2018]. Yet even when critics have specific donors in mind, restrictive legislation targets all foreign funding sources and the discourse paints all foreign-funded organizations as problematic. Hence, examining the effects of labelling an NGO as “foreign” is an important first step toward understanding the broader impacts of this phenomenon.

It is also likely that the effect of information about NGO funding sources on public attitudes will vary depending on the policy issue in question. Some of the most vociferous attacks on foreign-funded NGOs are directed towards those that advocate for human rights. In many countries, there is comparatively little discussion of foreign-funded NGOs advocating for environmental policy changes. One potential explanation for this disparity is that human rights issues such as torture and minority rights tend to be far more salient to the general public compared to environmental policy issues such as greenhouse gas emissions and overfishing. Because of this higher salience, criticism from NGOs poses a greater threat to regime viability when it relates to human rights issues. Yet while it may be the case that traditionally human rights issues have been more salient to the public, there are cases where environmental issues take center stage and NGOs critical of the government find themselves targeted by foreign funding restrictions. In India, for example, the government has waged a multi-year “war on Greenpeace” and similar NGOs in response to these groups’ efforts to block coal mining projects and other policies that could harm the environment [Lakshmi 2015]. Following a series of restrictions and increased scrutiny of groups that receive foreign funding, Greenpeace announced in early 2019 that it would shut
down two regional offices and substantially cut its staff in India [Reuters, 2019]. In the research design that follows, I deliberately select issues that are of relatively low salience to lessen issue-driven confounding and isolate the effect of information about NGO funding sources.

In sum, I expect that messages attributed to foreign-funded NGOs will have smaller impacts on attitudes toward policies that allegedly violate international law compared to messages attributed to domestically-funded NGOs or NGOs whose funding sources are left unspecified. These expectations presume, of course, that informing citizens that a policy violates international law depresses support for that policy, a contention with broad empirical backing in the literature [Chilton, 2014; Wallace, 2013; Kreps and Wallace, 2016; Chaudoin, 2014b]. When an NGO is labelled as foreign-funded, however, citizens are prompted to consider how foreign donors may be directing NGO advocacy to serve foreign interests, introducing bias into their activities. Hence, this label undermines the NGO’s credibility as a judge of the consistency between government policy and international law, which reduces the effectiveness of its messages calling attention to treaty violations. By contrast, citizens have less reason to doubt the credibility of NGOs that rely exclusively on domestic funding sources or those whose funding sources remain ambiguous. These expectations translate into the following hypotheses:

\( H_1: \) Accusations that a government policy violates international law will reduce citizen support for that policy.

\( H_2: \) Accusations that a government policy violates international law will lead to a smaller reduction in citizen support for that policy when the accusation is attributed to a foreign-funded NGO compared to a domestically-funded NGO.

\( H_3: \) Accusations that a government policy violates international law will lead to a smaller reduction in citizen support for that policy when the accusation is attributed to
a foreign-funded NGO compared to an NGO whose funding sources are left unspecified.

4.3 Research Design

To evaluate my hypotheses, I utilize an experimental design in which I directly manipulate whether a policy is said to violate international law and the source of that information. An experimental approach is suitable for this study because it allows me to identify the causal impact of labelling an NGO as foreign-funded on attitudes toward the relevant policy while holding constant the extraneous features that could confound this relationship. By randomly assigning respondents to treatment conditions, I can be confident that the treatment groups are, in expectation, balanced across all observed and, more importantly, unobserved confounders. This is critical in the context of attitudes toward government policy, where various factors such as prior exposure to information about NGO advocacy or the policy in question could mask the true relationship between the treatment and the outcome.

I conduct the study on a sample of US respondents because the US context offers a setting with relatively minimal pre-treatment effects. Unlike countries where legislation targeting foreign-funded NGOs is widely debated, the topic of foreign funding to NGOs is not commonly discussed in media reports. Therefore, respondents will be less likely to associate the “foreign-funded” label with a specific NGO, injecting their prior opinions into the experimental setting and potentially confounding the results. There is a concern, however, that US respondents will associate the “foreign-funded” label with Russian funding, given the ongoing discussion of Russian meddling in US elections. I attempt to minimize this possibility by describing the vignette as a future scenario. Even with this risk, though, I believe the US sample is the best place to begin this research program.

The research design involves a vignette-based survey experiment. I randomly assign subjects to one of four experimental conditions. In each condition, respondents
read a vignette that describes a policy advocated by the current US president. In one experimental condition, respondents only read about the policy in question and are given no information about international law. This condition allows me to estimate the baseline level of support for the policy. It will then be used to assess the effect of accusations that the policy violates US treaty commitments. In the remaining conditions, respondents read that an NGO believes the policy will violate international treaties that the US has signed. In one condition, respondents are given no information about the NGO’s funding sources. Comparing outcomes from this treatment condition to the baseline will allow me to test $H_1$, which states support for the policy in question should be lower when citizens hear the policy could violate international law. The final two conditions vary the funding source of the NGO in the vignette. In one condition, respondents are told that the NGO is a foreign-funded organization. In another condition, respondents are told that the NGO is completely domestically-funded. Including these two conditions will allow me to evaluate $H_2$ and $H_3$, which state that the reduction in support from hearing allegations of treaty violations should be smaller when the source is a foreign-funded NGO compared to a domestically-funded NGO or an NGO whose funding sources are not specified. Immediately following the vignette, respondents are asked to state their level of support for the policy in question.

I run two versions of the experiment using two different policies as the subject of vignette. This allows me to detect any differences across issues areas and enhances the external validity of the results. I use one environmental policy and one human rights policy since environmental and human rights treaties are both areas in which scholars expect NGO advocacy to be decisive in pushing for compliance (see Dai, 2007; Simmons, 2009). Additionally, legislative restrictions and criticism of foreign-funded NGOs are often directed towards NGOs that advocate for environmental protection and human rights. Thus, discussing a policy from either issue area is a realistic
<table>
<thead>
<tr>
<th>Experimental Condition</th>
<th>NGO Accusation of Treaty Violation</th>
<th>NGO Funding Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline ($B$)</td>
<td>No</td>
<td>—</td>
</tr>
<tr>
<td>Treatment + No Funding Info. ($T_N$)</td>
<td>Yes</td>
<td>None</td>
</tr>
<tr>
<td>Treatment + Domestic Funding ($T_D$)</td>
<td>Yes</td>
<td>Domestic</td>
</tr>
<tr>
<td>Treatment + Foreign Funding ($T_F$)</td>
<td>Yes</td>
<td>Foreign</td>
</tr>
</tbody>
</table>

Table 4.1: Summary of experimental conditions.

reflection of the political situation individuals would be likely to face. In Experiment 1, respondents read that the US president would like to make it easier for companies to build new oil and gas pipelines. In Experiment 2, respondents read that the US president would like to expand the use of solitary confinement in US prisons.\footnote{The latter policy issue was used in a recent study by Chilton (2014) about the effects of international law on public opinion.} Hence, each experiment involves two treatments: whether an NGO alleges that the policy violates treaty obligations and whether the NGO is described as foreign-funded, domestically-funded, or neither. To denote the content of each of the treatment conditions, I use subscripts that relate to the NGO funding source. The condition with no funding information is denoted $T_N$, while the condition in which the NGO is described as domestically-funded is denoted $T_D$. Finally, the condition in which the NGO is described as foreign-funded is denoted $T_F$. Table 4.1 summarizes the four experimental conditions.

At the beginning of each experiment, respondents answer a series of demographic questions, including age, gender, race, political ideology, and education level. Then, following a design used in similar survey experiments (e.g., Mattes and Weeks [2018]), I inform subjects that they will read about “a situation the US could face in the future.” I then provide the following background information, which is a modified version of the text used in a recent study of leader reputation by Mattes and Weeks [2018]: “The year is 2031. The US President is John Richards. President Richards took office in

\footnote{See appendix for the full survey instrument.}
2029 after serving in the US Senate for six years.” Using a hypothetical leader in the future lessens the risk of pre-treatment effects. For instance, respondents’ attitudes toward the current US president could influence their responses to NGO criticism of government policy. Referring to a hypothetical leader that could conceivably arise from either of the dominant political parties reduces this type of bias.

After reading the background information, respondents receive some information about the policy in question. In Experiment 1, respondents in all four conditions read “President Richards recently announced he wants to make it easier for companies to build new oil and gas pipelines. With more pipelines, the US will be able to transport oil and gas more efficiently to meet its energy needs.” In Experiment 2, respondents read “President Richards recently announced he wants to expand the use of solitary confinement in US prisons. Supporters of the use of solitary confinement argue that its use is necessary to maintain prison discipline and ensure the safety of prisoners and guards alike.”

Next, respondents in conditions $T_N$, $T_D$, and $T_F$ read that an NGO opposes this policy and alleges that it violates international treaties that the US has signed. In $T_N$ of Experiment 1, respondents read “A non-governmental organization (NGO), Clean Air for All, opposes this policy, arguing that making it easier for companies to build new pipelines violates international environmental agreements that the US has signed.” Similarly, in $T_N$ of Experiment 2, respondents read “A non-governmental organization (NGO), Human Rights for All, opposes this policy, arguing that expanding solitary confinement violates international human rights agreements that the US has signed.” As in the background information, I describe a hypothetical NGO to avoid introducing bias based on respondents’ previous evaluations of existing NGOs.

In $T_D$, respondents read nearly identical language in which the NGO is referred to as “a US-funded NGO.” In $T_F$, the NGO is referred to as “a foreign-funded NGO.”

---

2 The second sentence of this passage comes from Chilton (2014).
To reinforce the NGO funding treatments, I also include a sentence at the end of the vignette that describes the organization’s funding sources as either US or foreign. For example, in $T_D$ of Experiment 1, the vignette concludes “Clean Air for All is an environmental NGO that relies on funding from US sources around the country.” In $T_F$ of Experiment 2, the vignette ends with “Human Rights for All is a human rights NGO that relies on funding from foreign sources around the world.”

Finally, respondents are asked to rate on a five-point scale the extent to which they agree or disagree with a statement that the president should implement the policy in question. In Experiment 1, respondents indicate their agreement with the following statement: “The President should make it easier for companies to build new oil and gas pipelines.” In Experiment 2, the statement reads “The President should expand the use of solitary confinement in US prisons.” The five options available to respondents are “strongly agree,” “somewhat agree,” “neither agree nor disagree,” “somewhat disagree,” and “strongly disagree.”

Following the question about support for the policy in question, I also include a series of checks to ensure respondents paid attention to the vignette and received the treatment. These questions appear on a separate page and respondents cannot return to the vignette once they move past it. Respondents in all four experimental conditions are asked to recall the policy considered in the scenario. Then, respondents in the three treatment conditions that reference international law are asked to recall the NGO’s main argument against the policy in question. Finally, respondents in the NGO funding treatment conditions are asked to recall the NGO’s primary source of funding. The interpretation of the results presented below remains the same regardless of whether I analyze the full sample or subsets of respondents who passed these different topic and manipulation checks.

Because the theory outlined above speaks to the effect of the foreign-funding label

---

3 The order of these questions is randomized to minimize ordering effects.
Table 4.2: Summary of tests following from the hypotheses.

<table>
<thead>
<tr>
<th>Hypothesis</th>
<th>Implied Test</th>
<th>Reduced Form</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_1$</td>
<td>$s</td>
<td>B − s</td>
</tr>
<tr>
<td>$H_2$</td>
<td>$s</td>
<td>B − s</td>
</tr>
<tr>
<td>$H_3$</td>
<td>$s</td>
<td>B − s</td>
</tr>
</tbody>
</table>

on the relationship between information about treaty violations and policy support, the true outcome of interest is not the objective level of support for the policy in each condition but rather the difference in the mean level of support ($s$) between the baseline ($B$) and each of the conditions that reference international law ($T_N, T_D, T_F$). Based on $H_1$, the difference between the baseline and the condition in which a reference is made to international law but no information about NGO funding sources is given should be positive ($s|B − s|T_N > 0$). Following $H_2$, the difference between the baseline and the condition in which the NGO is foreign-funded should be less than the difference between the baseline and the condition in which the NGO is domestically-funded ($s|B − s|T_F < s|B − s|T_D$). Finally, $H_3$ implies that the difference between the baseline and the condition in which the NGO is foreign-funded should be less than the difference between the baseline and the condition in which no information about NGO funding is provided ($s|B − s|T_F < s|B − s|T_N$). Table 4.2 summarizes these expectations in their full form as well as a mathematically-equivalent reduced form for ease of analysis.

4.4 Results

A total of 1,600 US-based adults took part in this study on January 9, 2020. Respondents were randomly assigned to participate in one of the two experiments. 786 individuals participated in Experiment 1 and 814 participated in Experiment 2. Respondents were recruited via Amazon’s Mechanical Turk (MTurk) and directed to take the survey on the Qualtrics online platform. While subject pools recruited from
MTurk are not nationally representative, previous work has shown that such convenience samples tend to produce estimates of average treatment effects similar to those obtained with nationally representative samples (Berinsky, Huber and Lenz, 2012). Thus, MTurk is an appropriate venue to field this study. Each respondent was paid $0.65 for completing the survey, which took an average of 140 seconds (2.3 minutes).

The outcome of interest is the reported level of support for the policy in question on a five-point scale. Following previous studies (Kreps and Wallace, 2016; Tomz and Weeks, 2019), I create a binary dependent variable that is coded 1 for all respondents who stated they “somewhat” or “strongly” agree that the government should implement the policy in question and 0 otherwise. This yields a variable whose mean is easily interpreted as the proportion of respondents expressing support for the given policy.

4.4.1 Experiment 1: Environmental Policy

Figure 4.1 plots the mean level of support for easing the development of oil and gas pipelines across the four conditions in Experiment 1 (left) and the average treatment effect by NGO funding source (right), with 95% confidence intervals. Looking to the graph on the left, the baseline level of support for this policy is just shy of 50%, which indicates the sample is generally split on the question of whether a US President should adopt the policy in question. Yet once respondents are informed that an NGO believes this policy violates international environmental agreements that the US has signed, support drops to just 32%. This decrease is statistically significant at the 0.05 level, which offers initial support for \( H_1 \). The effect of this new information is largely the same regardless of whether the NGO is identified as domestically-funded or its funding sources are left unspecified. If the NGO is identified as foreign-funded, however, the level of support for the policy is nearly identical to that of the baseline. In line with \( H_2 \) and \( H_3 \), support in the foreign-funding condition
Figure 4.1: Proportion of respondents expressing support for easing the development of oil and gas pipelines (left) and average treatment effect by NGO funding source (right), with 95% confidence intervals.

(T_F) is higher than support in the domestic funding condition (T_D) (p < 0.01) and the no funding condition (T_N) (p < 0.1). Restricting this analysis to those respondents who successfully recalled the topic of the vignette, the NGO’s argument, and the NGO’s primary funding source yields similar results.

To better visualize the connection between these results and the theory, the graph on the right shows how the effect of information that a policy violates international law on public support for that policy varies by the stated funding source of the NGO delivering the message. This plot shows that informing respondents that an NGO believes an environmental policy violates international law reduces support for that policy by 13 [3, 23] percentage points. When the NGO is identified as domestically-funded, the effect is slightly more pronounced at 16 [6, 26] percentage points. When respondents are told that the NGO is foreign-funded, however, there is no statistically distinguishable effect of this information on support for the policy in question.

Overall, these results are commensurate with my expectations regarding the effect
of information about NGO funding sources on the relationship between accusations that a policy violates international law and public attitudes toward that policy. While information attributed to domestically-funded NGOs or NGOs whose funding sources are left unspecified produces the expected negative effect on attitudes toward a policy said to violate international law, information from foreign-funded NGOs has no statistically distinguishable effect on the average level of support for that policy.

4.4.2 Experiment 2: Human Rights

Figure 4.2 displays the mean level of support for expanding the use of solitary confinement in US prisons across the four conditions in Experiment 2 (left) and the average treatment effect by NGO funding source (right), with 95% confidence intervals. The graph on the left shows that the baseline level of support for this policy is considerably lower than that of easing the development of oil and gas pipelines in Experiment 1. Just 34% of respondents in the baseline condition expressed some support for this policy. New information that the policy violates international human rights agreements does not appear to have much of an effect in this context. There does appear to be a slight dip in support when a domestically-funded NGO is identified as the messenger of this information, but there remains considerable overlap in the confidence intervals. Across all four treatment conditions, none are statistically distinguishable from one another at conventional levels of statistical significance. Unsurprisingly, then, none of the treatment effects are statistically distinguishable from zero, though the effect for the domestic-funding treatment is close to conventional levels of statistical significance and in the expected direction.

These results cast some doubt on the generalizability of the findings from Experiment 1. One potential explanation for the null results from Experiment 2 could be that it is difficult to move public opinion on issues that the public is already largely for or against. The fact that nearly two-thirds of respondents in the baseline con-
dition either opposed or expressed a neutral attitude toward expanding the use of solitary confinement suggests there is little room to change attitudes on this topic. This finding is surprising given that Chilton (2014) reports evidence that references to international law do move public attitudes toward policy in this area.

While these results are somewhat discouraging, they also underscore the importance of running multiple experiments to examine the same question. Further investigating the generalizability of the findings from Experiment 1 would be a fruitful direction for future research. One way to move forward would be to replace the solitary confinement issue with a policy that presents a more salient trade-off to respondents. In Experiment 1, the oil and gas pipelines are presented as a trade-off between economic interests and environmental concerns. Although the solitary confinement issue was framed as a trade-off between prison discipline, safety, and human rights concerns, respondents are less likely to be personally concerned about guards and prisoners, since very few expect to find themselves imprisoned or employed as a
Table 4.3: Sample sizes by experiment, condition, and ideology.

<table>
<thead>
<tr>
<th>Experimental Condition</th>
<th>Total</th>
<th>Lib.</th>
<th>Cons.</th>
<th>Mod.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline ($B$)</td>
<td>204</td>
<td>82</td>
<td>75</td>
<td>47</td>
</tr>
<tr>
<td>Treatment + No Funding ($T_N$)</td>
<td>204</td>
<td>110</td>
<td>63</td>
<td>31</td>
</tr>
<tr>
<td>Treatment + Domestic ($T_D$)</td>
<td>187</td>
<td>91</td>
<td>54</td>
<td>42</td>
</tr>
<tr>
<td>Treatment + Foreign ($T_F$)</td>
<td>191</td>
<td>90</td>
<td>64</td>
<td>37</td>
</tr>
<tr>
<td>Baseline ($B$)</td>
<td>212</td>
<td>103</td>
<td>74</td>
<td>35</td>
</tr>
<tr>
<td>Treatment + No Funding ($T_N$)</td>
<td>195</td>
<td>93</td>
<td>70</td>
<td>35</td>
</tr>
<tr>
<td>Treatment + Domestic ($T_D$)</td>
<td>189</td>
<td>94</td>
<td>65</td>
<td>30</td>
</tr>
<tr>
<td>Treatment + Foreign ($T_F$)</td>
<td>218</td>
<td>110</td>
<td>77</td>
<td>31</td>
</tr>
</tbody>
</table>

4.4.3 Heterogeneous Treatment Effects by Ideology

While the results from Experiment 1 are broadly supportive of my hypotheses, examining only the average treatment effects tells us little about the mechanism at work. One way to indirectly examine the plausibility of my argument is to investigate the potential for heterogeneous treatment effects across the ideological spectrum. If in fact individuals are responding differently to foreign-funded and domestically-funded NGOs due to perceptions of bias, then these effects should be largely homogeneous across ideological dispositions.

To perform this analysis, I coded for each respondent whether they consider themselves to be conservative, liberal, or moderate. Before the vignette, respondents were asked to place themselves on a seven-point ideological scale ranging from extremely liberal to extremely conservative. I coded individuals as liberal if they described themselves as “extremely liberal,” “liberal,” or “slightly liberal.” Similarly, I coded individuals as conservative if they described themselves as “extremely conservative,” “conservative,” or “slightly conservative.” I only coded individuals as moderate if they chose “moderate,” the middle category. Table 4.3 displays the sample sizes...
Figure 4.3: Proportion of respondents expressing support for easing the development of oil and gas pipelines by ideological disposition (left) and average treatment effect by NGO funding source and ideological disposition (right), with 95% confidence intervals.

for each condition broken down by ideology. Because of the smaller sample sizes and lower statistical power, the analyses that follow are strictly exploratory and the results should be treated with caution.

Figure 4.3 displays the same information from Figure 4.1 for respondents in each ideological category in Experiment 1. The first thing to note is that initial support for easing the development of oil and gas pipelines is substantially higher among conservatives compared to liberals, which is unsurprising given this policy’s association with the Republican party. Support among conservatives is more than twice as large as support among liberals, a difference of 38 percentage points. The baseline level of support among moderates is slightly lower than that of conservatives, though there is significant overlap in the confidence intervals.

Interestingly, there is no statistically distinguishable effect of information about international law among conservatives when this message is attributed to an NGO
whose funding sources are left ambiguous or an NGO identified as domestically-funded. When the NGO is identified as foreign-funded, however, there appears to be a slight increase in support for the policy. This suggests that while conservatives are generally unmoved by references to international agreements, they become even more supportive of policies that violate international law when calls for compliance originate from foreign-funded NGOs, a movement in the opposite direction of what previous research would lead us to expect.

Among liberals, information about international law appears to have the greatest effect when attributed to a domestically-funded NGO and this effect is significant at the 0.05 level. The effect of information from an NGO whose funding sources are left unspecified is also negative but not statistically distinguishable from zero. When the NGO is identified as foreign-funded, however, levels of support remain almost identical to the baseline. For moderates, the treatment effects are all negative and statistically significant, though the confidence intervals are much larger, making comparisons between the conditions difficult.

These results suggest the average treatment effects discussed earlier are driven by information about NGO funding sources pulling policy support in opposite directions among conservatives and liberals. Liberals act most consistently with the theory outlined above, since they are most influenced by messages from domestically-funded NGOs and do not respond to messages attributed to foreign-funded NGOs. Yet conservatives seem to display a backlash effect when the message is attributed to foreign-funded NGOs and increase their support in response to calls for compliance.

These findings hint at a potential alternative explanation—that labelling NGOs as foreign-funded matters because it makes NGO messaging less effective overall and sparks a backlash among conservatives. Thus, laws and rhetoric that target foreign-funded NGOs may allow governments to not only tamp down criticism of policies that violate international law but also increase support for these policies among segments
Figure 4.4: Proportion of respondents expressing support for expanding the use of solitary confinement in US prisons by ideological disposition (left) and average treatment effect by NGO funding source and ideological disposition (right), with 95% confidence intervals.

of the population. Future research could further explore this possibility by increasing the sample size to include greater numbers of respondents in each ideological category and including follow-up questions after the vignette to probe respondent attitudes toward the NGO in question. For example, asking respondents to rate how much they trust the NGO could help shed light on how much of the effect is driven by perceptions of source bias compared to prior policy attitudes.

Figure 4.4 breaks down the results from Figure 4.2 in Experiment 2 by ideological category. Again, we see a similar pattern in that conservatives are much more supportive of the policy compared to liberals. In this context, however, moderates are closer to liberals than conservatives. The overall pattern here is largely the same as in Experiment 1, though the treatments effects are not statistically significant at conventional levels. This analysis further suggests a deeper analysis of the interac-
tion between ideological disposition and information about NGO funding sources is warranted.

4.5 Conclusion

This paper examines how identifying NGOs as foreign-funded influences their ability to alert citizens to violations of international treaties. Particularly in the contexts of environmental politics and human rights, NGOs are widely recognized as critical actors in the domestic enforcement of treaty commitments in part because of their ability to mobilize domestic constituencies in support of greater compliance. Yet as awareness of the power of NGOs has grown, so too have government attempts to block NGOs from engaging in these activities. This paper focuses on the consequences of a particular manifestation of this trend: laws that require NGOs to specify in their publications whether they receive foreign funding and broad-based criticism of foreign-funded NGOs.

I argue that labelling an NGO as foreign-funded reduces its ability to shift public opinion about policies that are said to violate international law because individuals view foreign-funded NGOs as more biased and therefore less credible sources of information. Evidence from two survey experiments offers partial support for this argument. In the context of an environmental policy, individuals do appear to be less influenced by accusations that the policy violates international law when that information is attributed to a foreign-funded NGO compared to when the NGO is identified as domestically-funded or its funding sources are not specified. When the NGO is labelled as domestically-funded or its funding sources are left ambiguous, the claim that the policy violates international law substantially reduces support for that policy, a finding in line with previous studies (Chaudoin 2014b; Chilton 2014; Kreps and Wallace 2016; Wallace 2013). A nearly identical experiment in which the policy in question was a human rights issue, however, yielded null results that
indicate messages about compliance with international law have little effect on public attitudes toward that policy. This surprising results points to the need to continue this research program to clarify the scope conditions of the theory.

An exploratory analysis of heterogeneous treatment effects by ideology further underscores the need for future research in this area. Investigating differences in treatment effects across the ideological spectrum, I find suggestive evidence that while liberals and moderates generally respond negatively to hearing that a policy violates international law and this effect is attenuated when the source of that information is a foreign-funded NGO, conservatives actually become more supportive of the policy when messages originate from foreign-funded NGOs. These results suggest a potential backlash effect among conservatives that demands further exploration. It would also be worthwhile to examine whether this backlash effect is limited to the US context or persists in other countries. A recent study by Lupu and Wallace (2019) finds evidence of a backlash against international law among Israelis. A useful extension of this work would be to replicate the present study in the Israeli context and in other democratic countries.

Another direction for future research would be to investigate how the effects of information about NGO funding sources vary by institutional context. In authoritarian regimes, domestically-funded NGOs may be viewed as less credible sources because they are more susceptible to government influence. Thus, labelling an NGO as foreign-funded may prove to be advantageous in this context, since it signals greater independence from the government.

As government restrictions on NGOs continue to spread around the world, understanding the consequences of these developments is critical to assessing the future efficacy of NGO advocacy, especially as it pertains to compliance with international treaties. In line with anecdotal evidence from activists in the field, this study suggests government restrictions on foreign-funded NGOs and the discourse that surrounds
them substantially hinder the ability of NGOs to do the work that scholars believe is critical to making global cooperation on issues such as climate change and human rights work. For policymakers, then, this study underscores the need for international actors to continue to push back against undue restrictions on civil society around the world. For scholars of international relations, this study highlights the importance of recent efforts to understand government restrictions on NGOs such as that by [Bakke, Mitchell and Smidt (2019)] and suggests additional work should be done to better understand the causes and consequences of this trend. This study offers a framework for investigating these consequences in the realm of public opinion that can serve as the foundation for future research in this area.
4.6 Appendix 1: Survey Instrument

The survey included questions divided into five categories: consent information, demographics, vignette, topic/manipulation checks, and attention check. In each of the following sections, I give the main questions and text from the survey. The surveys used for Experiments 1 and 2 were identical except for the vignette and topic/manipulation checks, as noted below.

Demographics

1. Are you male or female?
   - Male
   - Female
   - Other

2. What is the highest level of school you have completed or the highest degree you have received?
   - Did not graduate high school
   - High school diploma or equivalent
   - Some post-High-School education, but no Bachelor’s degree
   - Bachelor’s degree
   - Graduate degree

3. What is your age?
   - 18-24 years old
   - 25-34 years old
   - 35-44 years old
4. What is your race?

- White
- Black or African-American
- American Indian or Alaska Native
- Asian
- Native Hawaiian or Pacific Islander
- Other

5. Where would you place yourself on this scale?

- Extremely liberal
- Liberal
- Slightly liberal
- Moderate
- Slightly conservative
- Conservative
- Extremely conservative

**Vignette (Experiment 1)**

We are going to describe a situation the US could face in the future, in 2031. Some parts of the description may seem important to you; other parts may seem unimportant.

Here is some background:
The year is 2031. The US President is John Richards, who took office in 2029 after serving in the US Senate for six years.

President Richards recently announced he wants to make it easier for companies to build new oil and gas pipelines. With more pipelines, the US will be able to transport oil and gas more efficiently to meet its energy needs.

[Treatment]

Please tell us whether you agree or disagree with the following statement:
The President should make it easier for companies to build new oil and gas pipelines.

- Strongly agree
- Somewhat agree
- Neither agree nor disagree
- Somewhat disagree
- Strongly disagree

Treatments

1. **Baseline (B):** (nothing)

2. **No funding treatment** ($T_N$): A non-governmental organization (NGO), Clean Air for All, opposes this policy, arguing that making it easier for companies to build new pipelines violates international environmental agreements that the US has signed.

3. **Domestic funding treatment** ($T_D$): A US-funded non-governmental organization (NGO), Clean Air for All, opposes this policy, arguing that making it easier for companies to build new pipelines violates international environmental
agreements that the US has signed. Clean Air for All is an environmental NGO that relies on funding from US sources around the country.

4. **Foreign funding treatment ($T_F$):** A foreign-funded non-governmental organization (NGO), Clean Air for All, opposes this policy, arguing that making it easier for companies to build new pipelines violates international environmental agreements that the US has signed. Clean Air for All is an environmental NGO that relies on funding from foreign sources around the world.

**Vignette (Experiment 2)**

We are going to describe a situation the US could face in the future, in 2031. Some parts of the description may seem important to you; other parts may seem unimportant.

Here is some background:

The year is 2031. The US President is John Richards, who took office in 2029 after serving in the US Senate for six years.

President Richards recently announced he wants to expand the use of solitary confinement in US prisons. Supporters of the use of solitary confinement argue that its use is necessary to maintain prison discipline and ensure the safety of prisoners and guards alike.

[Treatment]

Please tell us whether you agree or disagree with the following statement:
The President should expand the use of solitary confinement in US prisons.

- Strongly agree
- Somewhat agree
• Neither agree nor disagree
• Somewhat disagree
• Strongly disagree

Treatments

1. **Baseline (B):** (nothing)

2. **No funding treatment** ($T_N$): A non-governmental organization (NGO), Human Rights for All, opposes this policy, arguing that expanding solitary confinement violates international human rights agreements that the US has signed.

3. **Domestic funding treatment** ($T_D$): A US-funded non-governmental organization (NGO), Human Rights for All, opposes this policy, arguing that expanding solitary confinement violates international human rights agreements that the US has signed. Human Rights for All is a human rights NGO that relies on funding from US sources around the country.

4. **Foreign funding treatment** ($T_F$): A foreign-funded non-governmental organization (NGO), Human Rights for All, opposes this policy, arguing that expanding solitary confinement violates international human rights agreements that the US has signed. Human Rights for All is a human rights NGO that relies on funding from foreign sources around the world.

**Manipulation/Topic Checks (Experiment 1)**

On the previous screen, you read about a situation the US could face in the future, in which a US President was considering a certain policy. What was that policy?

• Lowering fuel efficiency standards

---

4The order of these questions and their answer options were randomized.
- Eliminating regulations that limit emissions from US power plants
- Making it easier to build new oil and gas pipelines
- Expanding offshore drilling

On the previous screen, you read about an environmental NGO. What was the NGO’s main argument against the policy described in the scenario?

- It will violate international environmental agreements that the US has signed.
- It will violate the US Constitution.
- It will violate indigenous people’s human rights.
- It will violate the Clean Air Act of 1963.

On the previous screen, you read about an environmental NGO. What was the NGO’s primary source of funding?

- Foreign
- Domestic
- I don’t remember

Manipulation/Topic Checks (Experiment 2)

On the previous screen, you read about a situation the US could face in the future, in which a US President was considering a certain policy. What was that policy?

- Eliminating habeas corpus
- Expanding the use of the death penalty
- Expanding the use of solitary confinement
• Eliminating due process for violent criminals

On the previous screen, you read about a human rights NGO. What was the NGO’s main argument against the policy described in the scenario?

• It will violate international human rights agreements that the US has signed.
• It will violate the US Constitution.
• It will violate the US Civil Rights Act of 1964.
• It will violate due process rights.

On the previous screen, you read about a human rights NGO. What was the NGO’s primary source of funding?

• Foreign
• Domestic
• I don’t remember

Attention Check[^1].

We would like to get a sense of your general preferences.

Most modern theories of decision making recognize that decisions do not take place in a vacuum. Individual preferences and knowledge, along with situational variables can greatly impact the decision process. To demonstrate that you’ve read this much, just go ahead and select both red and yellow among the alternatives below, no matter what your favorite color is. Yes, ignore the question below and select both of those options.

What is your favorite color?

[^1]: The text for this question comes from [Berinsky, Margolis and Sances (2014)](Berinsky2014).
• White
• Black
• Red
• Yellow
• Green
• Blue
Chapter 5

Conclusion

This dissertation addresses a fundamentally important question in international politics: can global international institutions change state behavior in directions that alleviate the problems those institutions were created to resolve? I focus my attention on global cooperation in environmental politics and human rights because these issue areas present complex challenges for designing effective institutions. In both environmental politics and human rights, negotiators have incentives to make treaties more flexible to ensure widespread participation. Yet flexibility makes it easier for governments to ratify and comply with treaties without making substantial policy changes. Compounding this challenge is the fact that widely recognized mechanisms of enforcing international obligations, such as reciprocity, retaliation, and reputation (Guzman 2008), are less applicable in these areas.

In this dissertation, I address this question with three stand-alone essays about the relationship between global cooperation, NGO advocacy, and domestic politics. Previous work suggests that, despite the challenges noted above, international institutions can still make a difference because domestic actors work to hold governments accountable to international commitments (e.g., Simmons 2009; Dai 2007). This domestic mobilization, often led by NGOs, raises the costs of violating treaty obligations and encourages governments to comply with their stated commitments. In other words, treaty enforcement can come from within a state rather than from outside.
The three essays of this dissertation build on this idea that domestic politics matters for treaty compliance and effectiveness, deriving expectations about variation in the effectiveness of international institutions and highlighting trends that may impact the continued strength of this mechanism. In the first essay, I focus explicitly on the perceived trade-off between depth, or the extent to which states must modify their behavior to comply with a treaty, and participation. This trade-off is often viewed as especially pervasive in the context of human rights and environmental politics because many expect the public to view treaties in these areas with broad international support as more legitimate (Victor, 2006). Indeed, studies show domestic audiences are more supportive of such treaties and the policy changes they require when more countries participate in the agreement (Bechtel and Scheve, 2013; Tingley and Tomz, 2012). The pressure to increase participation encourages negotiators to incorporate flexibility mechanisms such as imprecise language and allowing reservations, which make it less costly for governments to ratify treaties by requiring fewer initial policy changes. Yet while this flexibility may make it easier to expand participation, many believe it comes at the cost of undercutting treaty effectiveness by allowing states to legally commit to and comply with treaties without making substantial policy changes.

However, I argue that this trade-off is less severe than scholars have been led to believe because initially shallow commitments can and often do deepen over time in response to domestic mobilization. Specifically, I show that states are more likely to deepen initially shallow commitments under conditions conducive to effective mobilization. I focus on two factors that facilitate and encourage mobilization: democratic institutions and influxes of information and assistance from international actors. Democratic institutions enable aspiring activists to meet, organize, collect and disseminate information, and direct political pressure toward government elites. In nondemocracies, however, additional conditions must be present for effective mobiliza-
tion to emerge. When international attention is focused on a state’s implementation of the treaty, domestic groups receive much-needed information and direct assistance from outside actors, which helps them identify deficiencies in government policies and gather public support for their initiatives. In a statistical analysis of reservations to the Convention on the Elimination of all forms of Discrimination Against Women (CEDAW), I show that democracies are more likely to withdraw reservations compared to nondemocracies, and the likelihood of withdrawal among nondemocracies increases following consideration of the country’s periodic report.

Therefore, this first essay shows that flexible treaties can still be effective because ratifying a treaty, even with reservations, sparks domestic political changes that ultimately encourage governments to remove their reservations and become legally accountable to the entire treaty. Put differently, this essay challenges claims that flexibility in general and reservations in particular undermine the effectiveness of international treaties by allowing states to avoid commitments that entail substantial policy change. Previous studies of reservations focus almost exclusively on the initial decision to enter reservations, examining which countries enter reservations, how many reservations are entered, and which parts of the treaty are affected (Simmons 2009, Hill Jr. 2016, McKibben and Western 2018). Though some scholars do note that reservations are not permanent and some countries do withdraw them (Koremenos 2016, 186), this phenomenon has not been studied systematically. Hence, this paper contributes directly to the human rights literature by identifying the conditions under which reservations are withdrawn. Additionally, it provides empirical backing to claims that allowing reservations may in fact be beneficial to the human rights treaty regime, for it promotes broad participation (Swaine 2006). More broadly, this essay also contributes to the international cooperation literature by showing that initially shallow commitments can deepen over time under conditions that facilitate domestic mobilization. This suggests designing agreements flexibly to enhance participation is
not as damaging as critics allege.

In the second essay, I argue that this domestic mobilization mechanism has implications for which types of institutions should be most likely to lead to their intended effect. I build on previous work that highlights the role of information in enabling domestic enforcement of treaty commitments (Dai, 2007; Chaudoin, 2014) to argue that institutions that provide useful information for domestic audiences should be more effective at eliciting cooperative behavior from states compared to those that do not provide such information. This is a logical conclusion to draw from previous research, but it is one that cannot be tested using observational data due to endogeneity and selection effects. It is likely that negotiators only include information-producing mechanisms when states are less likely to comply with treaty commitments but states would only ratify such agreements when they expect to comply. Therefore, the effects of information provision cannot be convincingly separated from expectations about compliance. These dynamics make causal inference especially difficult if not impossible in this context.

To address this problem, I design an original laboratory experiment modeled after the “pledge and review” procedure in the 2015 Paris Climate Agreement (Keohane and Oppenheimer, 2016). The experimental setting allows me to directly manipulate the extent to which the institution provides information. Analyzing the results from this experiment, I show that government leaders contribute greater amounts to a collective fund when the institution provides information about the quality of their contributions. I also show that interest group leaders condition their decisions to try and remove a government leader on the quality of government leaders’ contributions when such information is available. These findings support my claim that international institutions have greater effects on state behavior when they consistently provide information for domestic audiences.

This second essay also shows that domestic mobilization does not organically
emerge in response to any and all treaties. In order for domestic actors to enforce treaty obligations, they must have access to information that allows them to assess the extent to which government policies align with treaty goals. Because governments have incentives to suppress the release of this kind of information, international actors play a crucial role in collecting this information and making it accessible to domestic audiences. Hence, recognizing that domestic actors matter for treaty compliance does not mean the international community has no role to play in ensuring treaties are enforced. Rather, the international community can make treaties more effective by ensuring they are designed in ways that produce useful information for domestic groups.

In the third essay, however, I call attention to a troubling trend that is just beginning to catch the attention of scholars in this field. In recent years, activists and policymakers have decried the marked increase in government restrictions on NGOs and criticism of civil society actors, particularly those that rely on foreign funding to finance their operations. These developments have direct implications for the domestic mobilization mechanism discussed in the first two essays, because NGOs enable domestic enforcement of treaty commitments in part by alerting citizens when their government is in violation of treaty obligations. As messengers of this critical information, NGOs stimulate and coordinate the domestic mobilization that makes global cooperation work. If these new restrictions and criticism make it more difficult for NGOs to shape public opinion about policies that violate treaty commitments, then it is likely that global treaties will become less effective in the future.

As more scholars direct their attention toward investigating which countries impose restrictions and where civil societies face the greatest challenges, I turn to the question of what consequences these restrictions will have for NGO operations. More
specifically, I investigate how labelling NGOs as foreign-funded influences their ability to alert citizens to violations of treaty commitments. Despite anecdotal reports that the foreign-funded label stigmatizes and deligitimizes foreign-funded NGOs in the eyes of the public, there is little systematic research on the topic. I argue that identifying an NGO as foreign-funded reduces the organization’s credibility as a judge of the consistency between government policy and international law. Thus, individuals are less likely to believe their government is violating international law when the accusation comes from a foreign-funded NGO compared to an NGO that relies exclusively on domestic funding or one whose funding sources remain ambiguous. Because fewer individuals believe the message, the expected negative effect of informing citizens that a government policy violates international law on attitudes toward that policy is lessened.

I use survey experiments to show that NGO messages about compliance with international law are less effective at changing public attitudes when the NGO is identified as foreign-funded compared to when it is identified as domestically-funded or its funding sources are left unspecified. I also provide suggestive evidence that these effects are moderated by ideology, though investigating these dynamics will require substantial additional research. These findings suggest government criticism and restrictions targeting foreign-funded NGOs make it more difficult for these organizations to reach and influence citizens when governments violate their international obligations.

Therefore, this dissertation substantially deepens our understanding of how global institutions matter and how international and domestic actors can work together to maximize their effects. In addition to the contributions to the literature outlined above, this dissertation also offers useful recommendations for policymakers who wish to see global treaties achieve their goals. Most importantly, this dissertation highlights the importance of supporting civil society groups that advocate for greater
compliance with international obligations and greater contributions to global initiatives. The conclusions from the first two essays depend on the continued ability of NGOs to operate freely enough that they can perform the functions that make domestic enforcement of treaty commitments possible. If restrictions on NGOs and broad-based criticism of their activities continue to proliferate, then the effectiveness of international treaties, particularly in human rights and environment politics, is likely to suffer. The third essay shows how laws that require NGOs to specify in their publications whether they have received foreign-funded and criticism of foreign-funded NGOs make it more difficult for these organizations to influence public opinion about government policies that violate international law, a critical link in the causal chain connecting domestic mobilization to compliance with international treaties. International organizations and NGOs have already sounded the alarm about the “closing space for civil society” around the world (e.g., Carothers and Brechenmacher 2014; Amnesty International 2019). This dissertation underscores the importance of supporting these efforts to push back against these restrictions and offers a glimpse into the potential consequences if such efforts are unsuccessful.

Given the crucial role of NGO advocacy in treaty compliance and effectiveness, this dissertation also offers recommendations for negotiators designing new international agreements. More specifically, the first essay suggests negotiators should not be afraid to incorporate flexibility mechanisms as needed to enhance participation. Even if flexibility allows governments to ratify without making substantial policy changes, entering the institution and interacting with the treaty regime can spark domestic pressure for change that ultimately translates into deeper commitments. The second essay suggests negotiators should also ensure that treaties are designed with mechanisms for producing information, such as periodic review and other monitoring procedures. This information helps domestic audiences hold their governments accountable to treaty commitments and increases the likelihood that governments will
make substantial contributions to global efforts to deal with collective problems such as climate change.

This dissertation also points to a number of directions for future research. The first essay shows that initially shallow commitments can and often do deepen over time as domestic groups mobilize in favor of treaty goals. Ultimately, these groups wish to see their governments make policy changes that would bring them more in line with the standards outlined in the treaty. Withdrawing reservations is an important step in this process, since it enables domestic actors to hold governments legally accountable to provisions from which they were previously exempt. Future research should explore precisely how behavioral change fits into this picture. Do countries improve respect for human rights after withdrawing reservations? Why do some states make policy changes before removing reservations while others wait to act until after withdrawing reservations or simply remove reservations and never follow through on related policy changes? Addressing these questions will further deepen our understanding of how states interact with treaties that incorporate flexibility devices.

The second essay provides an experimental framework for assessing the causal impacts of different institutional designs that can serve as the foundation for much-needed work in this area. The challenges to causal inference in the study of international institutions are well-documented in the literature. To address these concerns, scholars often turn to complex statistical techniques that rely on potentially problematic assumptions (see Braumoeller et al. [2018]). Laboratory experiments certainly have their limitations as well, but combining these different methods allows researchers to draw more sound conclusions about these phenomena. Future work should compare information provision in international institutions with other mechanisms of enforcement, such as fines and dispute settlement procedures, to better understand which design features contribute the most to treaty effectiveness. Another fruitful direction would be to incorporate elements of endogenous institutional
design, allowing subjects to select into a specific institutional configuration. These modifications would enable scholars to make more credible claims of causality and deliver clearer recommendations as to how treaties can be made more effective.

Finally, the third essay suggests much more work is needed to understand the likely consequences of restrictions on NGOs for the relationship between NGO advocacy and public opinion. My exploratory analysis points to the potential for ideology to serve as a critical moderator of the effects of information about NGO funding sources on public responses to accusations that government policies violate international law. Future work should replicate my experiments with larger, more representative samples, to draw stronger conclusions about these effects. My work also offers a word of caution to researchers that design studies based on a single policy issue. My failure to replicate the results from the environmental politics experiment in the human rights context suggests future researchers should be careful not to make sweeping claims based on a study in a single policy area. Rather, researchers should think carefully about the characteristics of a given policy issue that could influence their results and test their hypotheses in multiple domains if possible.

Overall, such extensions would greatly enhance our understanding of the politics of international institutions. This dissertation tells us that treaties have great potential to shift how governments behave by engaging activists and citizens in the political process. It offers additional evidence that treaties matter and points specifically to ways in which future agreements can be designed to make them even more effective. It also highlights the critical importance of NGO advocacy in global governance and offers a glimpse into the consequences we are likely to observe if restrictions on and criticism toward NGOs continue to be a fixture of politics around the world. Therefore, this dissertation makes a crucial contribution to our understanding of the world and our ability to shape outcomes.
Bibliography


Coppedge, Michael, John Gerring, Carl Henrik Knutsen, Staffan I. Lindberg, Svend-Erik Skaaning, Jan Toerrell, David Altman, Michael Bernhard, Agnes Cornell, M. Steven Fish, Haakon Gjerlow, Adam Glynn, Allen Hicken, Joshua Krussell, Anna Lührmann, Kyle L. Marquardt, Kelly McMann, Valeriya Mechkova, Moa


Freeman, Marsha A. 2009. Reservations to CEDAW: An Analysis for UNICEF. New York: UNICEF.


Kadman, Noga. 2015. Backed by the System: Abuse and Torture at the Shikma Interrogation Facility. HaMoked and B’Tselem. URL: https://www.btselem.org/publications/summaries/201512_backed_by_the_system


URL: https://foreignpolicy.com/2018/06/07/democracies-need-a-little-help-from-their-friends/

Krebs, Ronald R. and James Ron. 2018b. “Why Israelis should welcome foreign funding of their NGOs.”.

URL: https://www.ynetnews.com/articles/0,7340,L-5268554,00.html


Lakshmi, Rama. 2015. “Why India is at war with Greenpeace.”.


Laub, Zachary. 2015. “International Sanctions on Iran.”.

URL: https://www.cfr.org/backgrounder/international-sanctions-iran

Linos, Katerina and Tom Pegram. 2016a. “Strong words make treaties more effective. So is the Paris climate accord worded too flexibly?”.


**URL:** https://unfccc.int/sites/default/files/english_paris_agreement.pdf

**URL:** https://treaties.un.org/Pages/ViewDetails.aspx?src=TREATY&mtdsg_no=IV-8&chapter=4&clang=en#44

United Nations. N.d. “Glossary.”.  
**URL:** https://treaties.un.org/pages/overview.aspx?path=overview/glossary/page1_en.xml

**URL:** https://www.usaid.gov/europe-eurasia-civil-society


