

**The Effects of Naming and Shaming on Public Support for Compliance
with International Agreements: An Experimental Analysis of the Paris Agreement¹**

Research Note

Dustin Tingley
Michael Tomz

April 2019

Not for circulation or citation without permission of the authors

Under the Paris Climate Agreement, each country selects and announces its own goals for mitigating climate change, without any legal penalties for failing to follow through. How could such an agreement affect policy? Building on recent work in political science, psychology, and other disciplines, we examine the role that “naming and shaming” plays on public support for compliance with costly climate policies. Using survey experiments fielded to nationally representative samples in the United States, we show that naming and shaming can impact support for government policies. However, this impact depended on how much effort the target exerted towards meeting the announced goal. Furthermore, while governments can use counter-rhetoric to minimize the effect of shaming on domestic opinion, this counter-rhetoric does not completely erase the impact of naming and shaming.

¹ Our thanks to Jonas Tallberg and participants in the “Legitimacy in Global Governance” seminar, University of Toronto, March 2019, for feedback on an earlier draft.

Introduction

International agreements are a core element in the study of international relations. Topics covered by international agreements range from nuclear weapons to trade and investment to the environment. In the environmental arena, there has been an explosion of international environmental agreements over the last 40 years (Mitchell 2017). Scholars working in this literature have engaged with questions about the design of international agreements (e.g., Rosendorff and Milner 2001, Johns 2014), incentives to join (e.g., Von Stein 2005, Lupu 2013), and whether and why countries comply with agreement commitments (e.g., Tallberg 2002, Simmons and Hopkins 2005, Mitchell and Hensel 2007).

We focus on the last question: why do countries comply with international agreements? In her review of the topic, Simmons (2010) draws out four main schools of thought. In the “realist” tradition, compliance with international agreements is epiphenomenal. Countries always follow their interests, and agreements are relevant only when balance of power considerations are binding. In the “rational-functionalist” tradition, countries perceive benefits that come from rule-based behavior. Monitoring helps to ensure that countries meet their commitments by engaging reputational costs for non-performers. A third tradition focuses on the “nature of the domestic regime.” Here, the institutional and legal constraints in democracies make them more likely to comply with international commitments. The fourth tradition centers around “normative” approaches to compliance. In this account, “normative standards of appropriate conduct are socially constructed reference points against which state behavior can be gauged” (Simmons 2010, pg. 85).

In this paper, we focus on a recent, important, international climate agreement known as the Paris Agreement. Drafted at the end of 2015 and signed by 195 countries in 2016, the Paris Agreement draws on several of these traditions to help foster compliance. The Paris Agreement was a departure from previous efforts in international climate agreements in that it did not include costly punishments for violating its terms. Instead, countries voluntarily chose “Nationally Determined Contributions” and agreed to intensive reporting and monitoring requirements on how well they were meeting their goals. Reporting and monitoring functions are a core element of the rational-functionalist tradition of compliance.

The structure of the Paris Agreement also draws heavily on the normative tradition. The agreement outlines no external enforcement mechanisms. Instead, countries publicly declare their own carbon reduction targets. The hope is that the publicity of this pledge combined with processes to enhance transparency mean that countries that do not meet their targets will be called out by other countries as well as social groups such as NGOs. This public naming and shaming acts as a cost of non-compliance.² This “naming and shaming” approach has received

²In some senses, these voluntary features are what made it politically feasible for countries to join the Paris Agreement and avoid the challenges faced by previous agreements like the Kyoto

both fanfare and optimism (Brown 2015) as well as skepticism (Bawden 2016). Interestingly, there has been surprisingly little scholarship directly probing whether and how “naming and shaming” is effective at generating compliance with international agreements (examples of relevant work include Jacquet et al. 2011, Jacquet and Jamieson 2016, Taebi and Safari 2017, Clark 2018).

In this paper, we examine the effect of naming and shaming on support for government policies on climate change. Our research design lets us distinguish between preferences individual citizens have over a policy from the effect of naming and shaming. This enables us to investigate the conditions under which naming and shaming will decrease support for a policy, increase support for a policy, and have no effect on support for a policy.

Our experiments yield three key findings. First, shaming by foreign countries shifted domestic public opinion in favor of compliance, substantially increasing the political incentive to honor the Paris Agreement. This conclusion held not only in the population as a whole but for each of the three main partisan groupings (Democrats, Independents, and Republicans) that future administrations might want to court.

Second, the effects of shaming depended on how much effort the target government exerted. In our experiments, shaming had no impact on U.S. public opinion when the U.S. government made no attempt to comply or when it met its obligations in full. Shaming was, however, highly effective in undermining public support for modest efforts that reduced emissions but fell short of America’s Paris commitments. More generally, our research suggests that the marginal effect of shaming should be largest when directed against “intermediate” policies that are not *prima facie* shameful or laudable.

Third, governments can use counter-rhetoric to minimize the effect of shaming on domestic opinion. We tested reactions to two counter-narratives by the U.S. government: defiance and regret. Both reduced the impact of shaming, especially among Republicans and Independents. Nevertheless, in our experiments, counter-rhetoric could not completely erase the effects of shaming. To our knowledge, this research provides the first micro-level evidence about the power and limits of shaming with respect to international compliance.

In what follows we motivate our study by engaging with several literatures, lay out our hypotheses and research design, and then present the results from a nationally representative survey experiment in the United States.

Protocol (Falkner 2016). "The old treaty model with binding emission targets is gone in favor of voluntary commitments that depend on an effectual Transparency Mechanisms to track nations' progress. Going along with international norms, or being “Named and Shamed” for deviating from those norms, resides as the strongest pathway the global community has to drive compliance from parties” (Martini 2016).

Motivation

I. Does Naming and Shaming Contribute to Compliance?

Naming and shaming efforts are designed to target a specific actor for a particular behavior/policy, and cast that actor in a negative light due to the behavior/policy. That is, naming and shaming is when an individual, group, country, etc., is publicly identified as having acted in an immoral, illegal, or socially undesirable way. It is a strategy that attempts to compel better behavior by that actor in the future or even compel better behavior by others who would consider engaging in the immoral behavior in future. Naming and shaming can take a variety of forms. The standard form is through public announcements that identify the actor(s) and their unseemly behavior. Additionally, one recent line of work examines how international *rankings* can compel behavior (Kelley and Simmons 2015, Cooley and Snyder 2015). Getting a low rank is a form of shaming.

A broad range of scholarship examines the role of naming and shaming in changing behavior, including psychology (Van Der Pligt and Vliek 2016), law (Skeel 2001, Gershowitz 2008) and criminology (e.g., Astbury and Leeuw 2010). One area in which naming and shaming has been extensively studied by political scientists is around human rights violations (e.g., Krain 2012). This literature has attempted to identify the mechanisms through which naming and shaming compels compliant behavior--in this case, fewer human rights violations--by target countries.

Some scholars argue that naming and shaming can work through a *domestic pressure mechanism* (Risse-Kappen, Risse, Ropp and Sikkink 1999, Finnemore and Sikkink 1998, Keck and Sikkink 1998). If domestic audiences are mobilized by external criticism, this can put pressure on a variety of domestic political actors to change their state's behavior. If these arguments are correct, then would expect members of the public to change their opinion of appropriate behavior for their country after their country is shamed. One way that this can work is via greater awareness of a problem in a shamed country (Ausderan 2014). Our focus on public opinion provides an opportunity to look at how members of the public will respond to being shamed for their country's climate policy.³

Others argue that naming and shaming works indirectly by mobilizing external actors. For example, Murdie and Peksin (2013) find that human rights organizations help to make external interventions more likely. If these interventions are successful then naming and shaming would play a positive role in changing state behavior. Furthermore, in principle, such interventions might also impact domestic audiences in the target state, perhaps mobilizing them in a way discussed above.

³ Some of this work focuses on the regime type of countries being shamed (e.g., Hendriz and Wong 2013). In this paper we do not speak to this literature given that our analysis is within a single country.

Conversely, others argue that naming and shaming is often not effective. Broadly speaking, one concern about relying on naming and shaming is that it will not be effective if there are no explicit costs (e.g., fines, sanctions) that are imposed. If there is no material incentive to follow through with an agreement then non-compliance would be expected when there are benefits to deviating. It is exactly this line of reasoning that has driven the design of many previous international environmental agreements. In practice, the ineffectiveness of naming and shaming has been documented by a number of scholars, including Hafner-Burton (2008), who finds that reports released by the prominent NGO Amnesty International have no impact on levels of torture in countries that are described as frequently practicing torture.

Another line of work raises the possibility that public shaming can even backfire. For example, Terman (2017) argues that naming and shaming can be counterproductive if leaders in a targeted state can portray it as being *an attack on the country's sovereignty and identity*. This dynamic is supported by some work in psychology that focuses on how naming and shaming operates when the entity doing the naming and shaming is considered to be an "outgroup." Here, the motivations of the outgroup are deemed suspect, leading to a backfire effect (Hornsey et al. 2002, Hornsey 2005). Further, as Carnegie and Carson (2018) note, norm erosion--or more broadly speaking the breaking of promises--has often received less attention than the creation of norms (Bailey 2008; Panke and Petersohn 2011). If shaming motivates those committing the bad act to continue their behavior, then the strength of the norm could diminish. Naming and shaming could ultimately lead to decreases in compliance.

II. Under What Conditions Will Naming and Shaming Be Effective?

While the preceding discussion gives some sense of why shaming might or might not work in generating compliance, it is helpful to dig a bit more specifically into the conditions under which shaming might or might not work. Some work considers institutional answers to these questions, such as the type of government in place in a country that gets shamed (e.g., Hendriz and Wong 2013). Terman and Voeten (2018) argue that shaming is effective only if it comes from a strategic partner.⁴ Kelley and Simmons (2015) show that rankings can impact behavior/policy as long as the entity doing the ranking has sufficient status and credibility. Franklin (2008) shows that naming and shaming can reduce political repression but only when the target state depends on capital from states doing the naming and shaming.

Other work that clarifies the conditions under which naming and shaming might have an impact focuses on the accuracy of naming and shaming. For example, O'Donnell (2006) looks at United Nations Security Council Resolution 1530, which falsely blamed the Euskadi Ta Askatasuna (ETA) for the 2003 Madrid bombings. The inaccurate claims and the ensuing response

⁴They argue that states want to maintain good relationships with countries that they depend on for other material benefits. So if a state is in a military alliance, shares a common geo-political ideology, or has a high level of trade with another state, then criticisms from this partner will be especially salient.

questioning the efficacy of the UNSC highlights the importance of reliability in public naming and shaming campaigns (see also Cabus and De Witte 2012). Individual perceptions of the accuracy of a naming and shaming event could thus be important. For example, if a country were named and shamed after taking what could be broadly construed as an acceptable policy choice (for example, complying with a commitment), the effect of naming and shaming might be diminished.

In this paper, we focus more on the characteristics of the policy itself than the institutions that generate it and we bracket the identity of who is doing the naming and shaming. Given our substantive area of interest we also focus more on compliance with prior agreements. We distinguish between three types of policies: no compliance, partial compliance, and full compliance.

When shame is cast on behavior/policy that is clearly deplorable--for example when there is no compliance at all--the additional impact of shame might not change how people evaluate a behavior or policy. All reasonable people would already agree that the act was not acceptable and so shame would do little additional work. No new information, perspective, etc., is brought to bear even though the act of shaming was accurate. In this case, naming and shaming might have little to no marginal impact on compliance because it w mechanisms that could lead to compliance, like through the "domestic pressure" mechanism discussed above.

Conversely, shaming on its own might have very little impact when the behavior/policy that is being shamed is not at all considered wrong, such as when there is complete compliance. Shaming desirable behavior might have little impact. Or it could even increase support further for the behavior.

This suggests that shaming might be more likely to have an effect within an intermediate range between clearly shame-worthy behaviors and those that are clearly not shameful. Behavior that is partially compliant lies within this region. The independent effect of shaming could change opinions over policy/behavior in this intermediate region. To our knowledge, existing scholarship has not dealt with this distinction between policies that are at the extremes versus those that are intermediate, perhaps because existing work primarily focuses on observational data. Indeed, observationally we might see most instances of shaming at the extreme of the behavior/policy being squarely opposed. But that does not mean that shaming in itself has any impact on support for the behavior/policy.⁵

⁵ Similarly, another condition where shaming might be important is when it helps citizens understand whether or not substantial reform--rather than more minor incremental reform--has actually taken place. That is, citizens might be unsure whether an effort by the government is praiseworthy. Shame could help citizens see they have been duped into supporting a small, less praiseworthy change. This dynamic is present in studies of "greenwashing." While Malhotra, Monin, and Tomz (2018) show that corporations can take minor pro-environmental steps to gain goodwill with the public ("greenwashing"), Parguel et al. (2011) find that external rating systems

III. Alternatives to Shaming: Positive Encouragement

A final consideration is to flip shaming around and consider the role of positive encouragement. Beyond being a staple strategy of enlightened parents, this perspective was mentioned by then UN Assistant Secretary-General on Climate Change János Pásztor, who indicated that he hoped there would be a “name and encourage” process (Falk 2015).⁶ Furthermore, Hornsey and Imani (2004) show that if the criticism is taken to be constructive it might decrease the probability of backfire. In our empirical analyses, we are able to see what happens when there is encouragement for a certain behavior and we compare its effect to other rhetorical strategies such as shaming. We are also able to triangulate these rhetorical strategies against a country’s policy choices.

IV. Responses to Shaming and Two-Sided Rhetoric: Defiance and Regret

In practice, shaming need not happen in isolation of how the target responds to shaming. In other words, rhetoric is two-sided in competitive political environments with competing actors and values (Druckman and Lupia 2016). Indeed, part of the motivation of the work by Terman (2017) is to point out that targets can respond to attempts by others to shame them. There are of course a variety of strategies available in this context, including questioning the motivations of the shamer or an attempt to distract attention away from the issue. We focus our attention on two responses a target might adopt in responding to naming and shaming: defiance and regret.

Defiance occurs when the shamed actor rebuts the accuracy of a claim or the motivation of the actor(s) doing the naming and shaming. Defiance in the human rights literature (e.g., Wachman 2001) considers Chinese responses to shaming of its human rights practices and is central to Terman’s analysis (2017). When might defiance work? In the criminology literature, Sherman (1993) argues that defiance will be especially salient and effective when shaming is considered to be unfair or excessive.

Regret plays a similarly important role in the criminology literature (Tangney and Dearing 2003, Saffrey et al. 2008). Relevant to the present context, Rees et al. (2015) show that pro-environmental behavior can be induced by triggering negative moral emotions like regret. Using a laboratory experiment with 114 college students, they manipulated whether climate change was induced by humans or naturally caused. Those in the human-caused condition expressed more negative emotions (e.g., guilt) and these emotions mediated an impact on

can prevent corporations from claiming environmental progress when in fact their impact is minimal. In other words, shame might be able to counter attempts at greenwashing.

⁶ This distinction is important in the international relations literature. For example, some have argued that positive inducements are unlikely to work (e.g., Drezner 1999) and Baldwin (1971) argues that there has been too much attention to negative rather than positive sanctions. While these authors are concerned more with material inducements (carrots versus sticks) the general point still applies to positive versus negative rhetoric.

support for pro-environmental behavior. Beyond the climate space, scholars of international relations have also been interested in the role of expressions of regret. For example, Lind (2009) studies the role of apologies between countries with a history of conflict (see also, for example, Olick 2013).

V. Discussion and Hypotheses

The ample literature on naming and shaming suggests that exploring its microfoundations will contribute to a better understanding of the conditions under which naming and shaming might change state behavior. Our work tries to do this in several respects. First, many studies use observational data, which offers many insights but makes it difficult to draw clear causal inferences. Second, many studies assume that public opinion is relevant to the mechanisms underlying naming and shaming, but this is rarely directly studied. Finally, despite the fact that one of the signature international agreements that may mitigate global climate disaster relies on naming and shaming, we are unaware of studies that interrogate whether--and under what conditions--it might work.

While our focus is on whether the impact of naming and shaming depends on the behavior of a country, it is also important to take into account the fact that different groups of citizens might respond differently. As a first approximation, we distinguish respondents according to their party affiliation. In the US, party affiliation is closely coupled with views about climate change and climate change regulations. The Republican and Democratic parties have articulated different platforms on the topic, and on average citizens have diverging views depending on their party affiliation (Dunlap and McCright 2011).

Hypotheses

In what follows we introduce an experimental design and results that probe the following hypotheses:

Main effect of naming and shaming: In general, naming and shaming will decrease support for a country's behavior.

Conditional effect of naming and shaming: Naming and shaming will be less effective when the country's effort is either clearly bad or clearly good. The impact of naming and shaming will be greatest at intermediate points of effort.

Counter rhetoric: Governments can minimize the effects of naming and shaming by countering a naming and shaming act.

Positive Encouragement: Positive encouragement will increase support for a level of effort that meets the country's commitments.

Design of Experiment 1

To study the consequences of shaming, we conducted four experiments, which we embedded in public opinion surveys that were administered to representative samples of adults in the United States in September-October 2018.

In all four experiments, respondents considered a hypothetical future in which the U.S. had joined the Paris Agreement and pledged to reduce emissions.⁷ We focused on the U.S. for several reasons. First, the U.S. emits more carbon than any country other than China, and U.S. carbon consumption per capita is among the highest in the world. To address global climate change, then, it is important to understand whether shaming could increase American public approval of emissions control laws. Second, in June 2017, President Donald Trump announced that the U.S. would withdraw from the Paris Agreement. Trump's decision opened a unique research opportunity: to present a future scenario in which the U.S. had just entered the agreement and was making decisions about compliance. Nevertheless, future research could test the effects of shaming in other contexts.

Our first experiment (N=2,884) contained two randomized elements: the amount of effort the U.S. exerted after joining the Paris Agreement, and whether foreign countries shamed the U.S. Experiment 1 served as a template for three follow-up experiments, described later in the memo.

In Experiment 1, all participants read the following preamble:

The Paris Agreement is an international agreement about climate change. Every country that joins the agreement promises to contribute to the worldwide goal of fighting climate change, by developing and carrying out a plan to reduce its emissions of carbon dioxide as quickly as possible.

In the future, the U.S. government must decide whether to join the Paris Agreement, and whether to pass new laws to reduce U.S. emissions of carbon dioxide. On the following screens, we will describe one approach the U.S. government could take in the future, and ask whether you approve or disapprove.⁸

All participants then considered a scenario in which a future U.S. administration joined the Paris Agreement.

⁷ In separate work we analyze the effect of joining the Paris Agreement on support for different levels of effort (Tingley and Tomz 2019).

⁸ Throughout the survey, we included comprehension questions to make sure participants understood the scenario they were reading. When analyzing the data, we restricted the sample to respondents who correctly answered at least 80% of the comprehension questions. For the text of the questionnaire, including comprehension questions, see the online appendix.

In 2021, the U.S. government announced that it would join the Paris Agreement. When it officially joined later that year, the U.S. said: “As a member of the Paris Agreement, we pledge to reduce U.S. emissions of carbon dioxide by 25%.”

Having established this context, we randomized what steps, if any, the U.S. government took to comply. Some participants read that the U.S. did not pass any new laws to reduce carbon emissions. Others read that the U.S. passed laws to reduce emissions by either 5% or 25%. We randomized the costs (in square brackets) of these emission control policies, but because costs were not our focus in this paper, we averaged over costs when analyzing the data. The text of the three conditions appears below.

No Action: Over the next few years, the government did not pass any new laws to reduce carbon dioxide emissions. Because it did not pass any new laws, the government did not affect U.S. energy prices or U.S. carbon emissions.

Cut 5%: Over the next few years, the government passed new laws to reduce carbon dioxide emissions. Experts agreed that the new laws would increase U.S. energy prices by [4 or 10]% and reduce U.S. carbon emissions by 5%.

Cut 25%: Over the next few years, the government passed new laws to reduce carbon dioxide emissions. Experts agreed that the new laws would increase U.S. energy prices by [4 or 10]% and reduce U.S. carbon emissions by 25%.

These policies differed not only in their effects on carbon emissions, but also in their compliance with the Paris Agreement. The first two policies, no action and cut 5%, violated America’s Paris commitment, whereas the third policy honored the pledge the U.S. had made. One can, therefore, think of cut 25% as synonymous with compliance, while viewing no action and cut 5% as instances of noncompliance.

We independently randomized whether foreign countries shamed the U.S. Half the the participants saw no mention of foreign shaming; the other half received a passage in which foreign countries characterized U.S. behavior as shameful. The content of the shaming varied, depending on what the U.S. had done.

Shaming if No Action: Many countries said the U.S. should be ashamed of itself. They criticized the U.S. for doing nothing to reduce U.S. emissions, and for violating the promises it made when it joined the Paris Agreement.

Shaming if Cut 5%: Many countries said the U.S. should be ashamed of itself. They criticized the U.S. for doing so little to reduce U.S. emissions, and for violating the promises it made when it joined the Paris Agreement.

Shaming if Cut 25%: Many countries said the U.S. should be ashamed of itself. They criticized the U.S. for doing so little to reduce U.S. emissions.

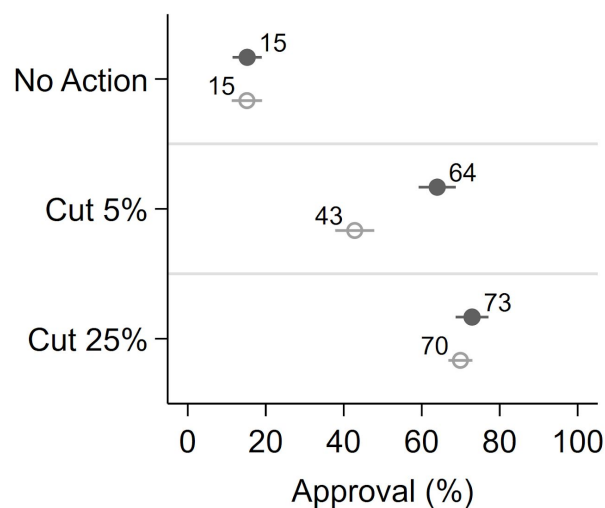
In all three passages, shamers asserted that the U.S. was doing little or nothing to reduce its carbon emissions. When the U.S. cut emissions by 25%, thereby honoring its Paris commitment, however, shamers could not lodge the additional complaint that the U.S. had failed to comply.

Having presented the scenario, we asked: “Taking into account all the decisions the U.S. government made in the passage you read, would you approve or disapprove of what the U.S. government did overall?” The response options were approve strongly, approve somewhat, neither approve nor disapprove, disapprove somewhat, or disapprove strongly. For simplicity we focused on a natural and easily interpretable quantity of interest, the percentage of respondents who approved, but the appendix shows that our conclusions held when we analyzed the full five-point scale, as well.

Main Effects in Experiment 1

We now summarize the findings from our first experiment. Figure 1 displays the percentage of Americans who approved of how the U.S. government behaved. The figure contains three rows, corresponding to the three approaches the government took: not passing any new laws to reduce carbon emissions (No Action), passing new laws to reduce emissions by 5% (Cut 5%), or passing new laws to reduce emissions by 25% (Cut 25%), thereby meeting its Paris commitments. The solid dots represent levels of public approval when there was no foreign shaming; the hollow dots represent approval when foreign countries shamed the U.S.; and the thin lines bisecting the dots are 95% confidence intervals.

Figure 1: Public Approval of U.S. Policy without Shaming (Solid Dots) and with Shaming (Hollow Dots)



In the vignettes without shaming (the solid dots), three findings emerged. First, only 15% of Americans approved when the government took no action to reduce carbon emissions. This finding corroborates other opinion polls at the time, which showed overwhelming public support for doing something about climate change (e.g., Gustafson et al. 2019).

Second, a large majority (64%) approved when the government slashed emissions by 5%. Compared to doing nothing, this modest action increased the government's popularity by a remarkable $64 - 15 = 49$ percentage points. Evidently, the government earned substantial domestic credit for cutting emissions, even though the cuts were minor and fell short of the Paris pledge. Previous research has found that corporations can gain goodwill by taking small steps to help the environment (Malhotra, Monin, and Tomz 2018). Our experiment revealed that governments have similar potential; they can boost approval through minor environmental measures, including ones that fail to meet their international environmental commitments.

Finally, nearly three-quarters of participants (73%) approved when the government took the more ambitious step of cutting emissions by 25%. On the one hand, this qualifies as further evidence of an American desire to do something about climate change. On the other hand, the marginal effect of additional effort on public opinion was relatively small. When the government quintupled its cuts from 5% to 25%, thereby living up to its Paris pledge, approval rose by only $73 - 64 = 9$ percentage points. We conclude that, absent shaming, the government could substantially increase its public image by passing modest legislation, but additional environmental effort—including compliance with international commitments—would not bring commensurate gains in popularity.

We next consider how shaming affected these conclusions. The hollow dots in Figure 1 show levels of approval when foreign countries characterized each of the three U.S. approaches as shameful. When the U.S. took no action, respondents who received the shaming treatment read that foreign countries “criticized the U.S. for doing nothing to reduce U.S. emissions, and for violating the promises it made when it joined the Paris Agreement.” This foreign criticism had no effect. Approval of inaction stood at 15%, regardless of whether foreigners shamed the U.S. or not.

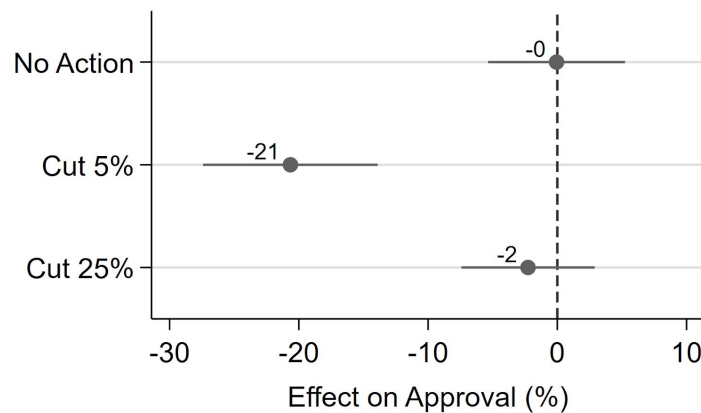
Shaming also proved inconsequential when the U.S. cut emissions by 25%. As noted earlier, 73% of Americans approved of this behavior when foreign countries abstained from shaming. Approval remained nearly same level, 70%, when foreign countries denounced the government for “doing so little to reduce U.S. emissions.” These patterns are consistent with our prediction that shaming would have little effect on behavior that was already *prima facie* shameful or compliant.

Shaming did, however, alter Americans perceptions of intermediate action. Recall that 64% of people in the no shaming condition approved when the government cut emissions by 5%. When foreigners denounced the U.S. for “doing so little to reduce U.S. emissions, and for violating the

promises it made when it joined the Paris Agreement,” approval fell to 43%. Thus, shaming deprived the government of credit it would have reaped for taking modest action that did not live up to its international commitments.

Figure 2 summarizes how shaming affected public approval, given each of the three U.S. policies in our vignettes. Each dot represents a treatment effect, and the thin lines are 95% confidence intervals. The figure shows a striking curvilinear pattern: shaming made no difference when the government remained idle or cut emissions by 25%, but it proved transformative when the government cut emissions by 5%. In that situation, shaming reduced approval by 21 percentage points, compared to an otherwise identical scenario in which foreign countries abstained from shaming. The confidence interval around this estimate ranged from 14 to 27 points, so we can be quite sure the effect did not arise by chance alone.

Figure 2: Effects of Shaming on Public Approval



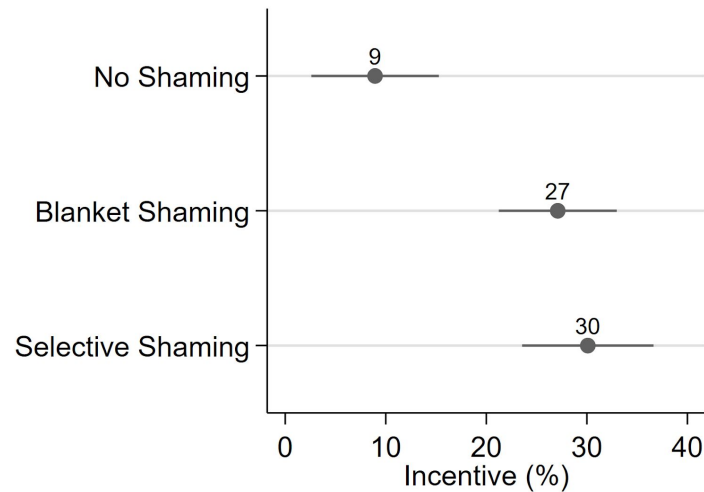
Incentives to Comply in Experiment 1

Given these effects, how might shaming affect the government’s incentive to comply with its Paris commitments? To find out, we used data from our first experiment to estimate how public approval would change if the government made deep cuts (25%) instead of shallow ones (5%), and how foreign shaming would affect the size of the gain. Below, we use the term “incentive” as shorthand for the estimated gain in approval. In reality, of course, many factors in addition to public opinion would affect the government’s willingness to comply with prior commitments. Nonetheless, our basic conclusions should hold as long as the political incentive to adopt a policy is increasing in the domestic popularity a government could gain by adopting the policy.

Figure 3 shows the government’s incentive to comply, conditional on three shaming strategies foreign countries could play. First, foreign countries could abstain from shaming altogether. Recall that, in vignettes without shaming (solid dots in Figure 1), 64% approved when the government cut emissions by 5%, whereas 73% approved when the government cut emissions

by 25%. Thus, even in the absence of shaming, the government could gain $73 - 64 = 9$ approval points by legislating deep, compliant cuts instead of shallow, noncompliant ones.

Figure 3: Political Incentive to Comply (Cut Emissions by 25%),
Given Three Foreign Shaming Strategies



Second, foreign countries could play a strategy of blanket shaming: criticizing the U.S. regardless of how much or little it did to address climate change. Recall that, when foreign countries shamed the U.S. indiscriminately (hollow dots in Figure 1), 43% of respondents approved of cutting emissions by 5%, whereas 70% approved of cutting emissions by 25%. Thus, if other countries engaged in blanket shaming, the U.S. government could gain $70 - 43 = 27$ approval points by making the cuts it promised when it joined the Paris Agreement.

Finally, foreign countries could use shaming selectively: criticizing the U.S. if and only if it failed to honor its pledge to slash emissions by 25%. As we have seen, 43% of Americans approved when the U.S. cut emissions by 5% and foreigners responded with shaming (hollow dot in the second row of Figure 1). In contrast, 73% approved when the U.S. cut emissions by 25% and foreigners refrained from shaming (solid dot in the third row of Figure 1). Thus, faced with the prospect of selective shaming, the U.S. government could gain $64 - 34 = 30$ approval points by complying with the Paris Agreement.

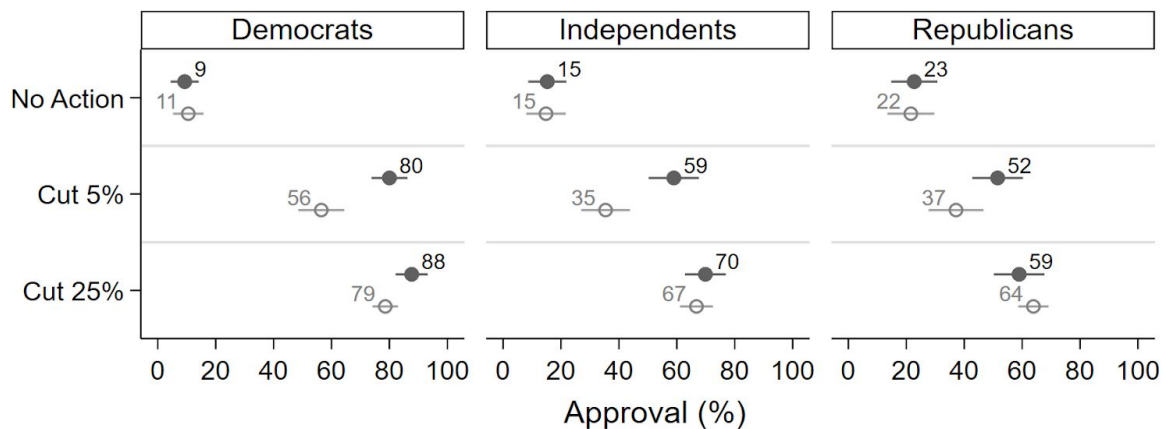
We draw two major conclusions from Figure 3. First, shaming can incentivize the U.S. government to honor its international commitments. In our experiments, the political incentive to comply was three times stronger when foreign countries shamed as when they did not. Second, blanket shaming was as effective as selective shaming. Blanket shaming could have been a double-edged sword, reducing support not only for modest cuts but also for ambitious measures. In our study, though, shaming proved ineffective when the government cut emissions by 25% as per its Paris pledge. Consequently, blanket shaming performed just as well as selective shaming.

Effects by Party in Experiment 1

Thus far, we have examined attitudes in the sample as a whole, but previous research about climate change has found substantial differences by political party. Belief in anthropogenic climate change, support for emissions control policies, and faith in international organizations tend to be stronger among Democrats than among Republicans. We therefore investigated whether our conclusions might change if we split the sample by political party. Remarkably, though, Figures 4–6 show that our conclusions held not only in the pooled sample, but also within each partisan camp.

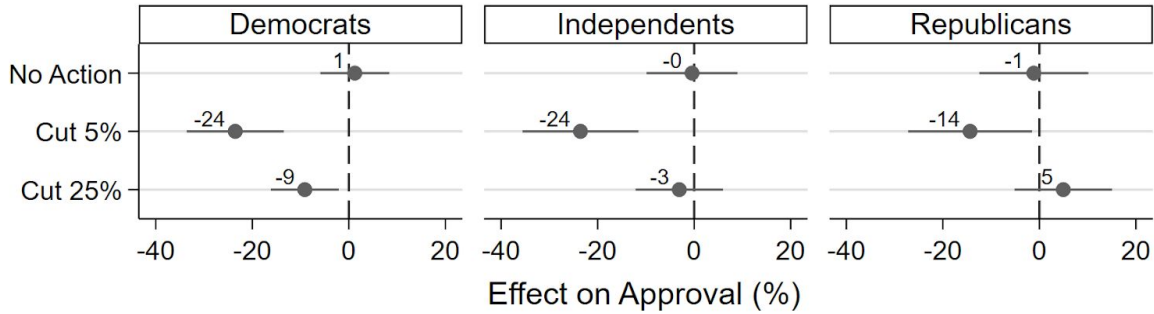
Figure 4 presents the percentage of Democrats, Independents, and Republicans who approved of how the U.S. government behaved in our vignettes. When foreign countries refrained from shaming (solid dots), several patterns stood out. First, approval of inaction was lower—and approval of action was higher—among Democrats than among Republicans, with independents usually falling in between. Nevertheless, majorities in all three groups preferred action over inaction. Second, within each partisan group, approval increased massively when the government shifted from inaction to moderate action, but the marginal benefits of taking more substantial action (thereby complying with Paris) were relatively small.

Figure 4: Public Approval of U.S. Policy without Shaming (Solid Dots) and with Shaming (Hollow Dots), By Party



Next, consider how members of each partisan group responded when foreign countries shamed the U.S. (hollow dots). Here, too, familiar patterns emerged. In all three groups, shaming greatly reduced public approval of intermediate action but had little effect at the two extremes. The one exception was Democrats, whose impressions of cut 25% soured when foreign countries engaged in shaming (see also Figure 5, which summarizes the effects of shaming). Overall, though, the patterns in Figures 4 and 5 resemble the patterns in Figures 1 and 2.

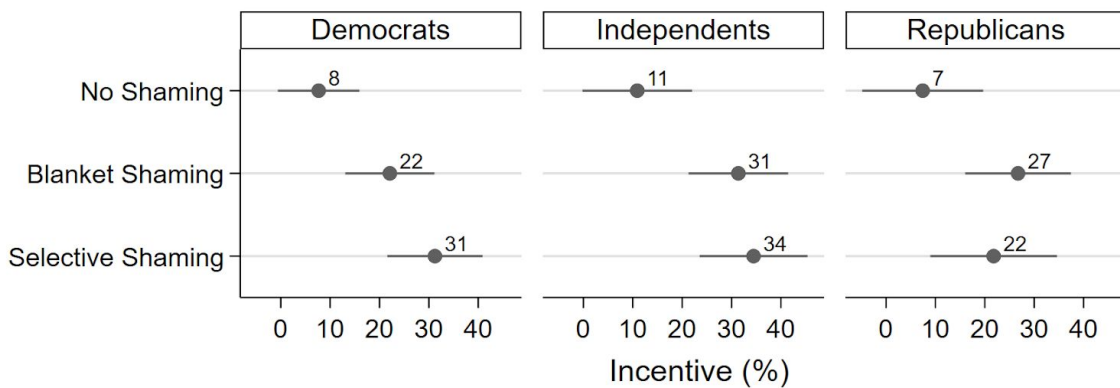
Figure 5: Effects of Shaming on Public Approval, By Party



Finally, for each party, we asked how shaming would affect the political incentive to comply with the Paris Agreement. Splitting the sample by party is informative because politicians might pander to their own party or the median voter, rather than the electorate as a whole. By dividing the data along partisan lines, we could test how shaming would affect the political calculations of a government that cared primarily, if not exclusively, about the opinions of Democrats, or Independents, or Republicans.

Figure 6 confirms that shaming would raise the incentive to comply, regardless of which electoral segment the government was courting. Suppose, for example, that the government cared only the approval of Democrats. In that case, the incentive to comply would be nearly four times stronger if foreign countries shamed selectively, than if they refrained from shaming altogether. Similarly, selective shaming would triple the political incentive to comply among governments catering to Independents or Republicans. We conclude that foreign shaming can alter domestic political incentives not only on average, but also for governments with strong partisan biases.

Figure 6: Political Incentive to Comply (Cut Emissions by 25%), by Party



Design and Effects in Experiment 2

Our first experiment examined the consequences of shaming. We found that countries could use shaming to manipulate U.S. public opinion, creating political pressure to honor the Paris Agreement. We next investigated whether a hybrid strategy, blending shaming and praise, could be even more effective at changing the domestic political calculus. Could foreign countries maximize the incentive by praising the U.S. when it complied but shaming the U.S. when it did not?

To test this hypothesis, we introduced a new experimental condition in which foreign countries praised the U.S. for cutting emissions by 25%. People assigned to this condition read: “Many countries said the U.S. should be proud of itself. They praised the U.S. for doing so much to reduce U.S. emissions, and for honoring the promises it made when it joined the Paris Agreement.” We administered this new condition to 527 participants in October 2018.

We anticipated that foreign encouragement would make Americans even more positive about “cut 25%,” creating an even stronger incentive to comply. To our surprise, though, praise had no effect. Recall (from Figure 1) that 73% of participants approved when the U.S. cut emissions by 25% and the vignette contained no foreign commentary. Approval rose by only one point, to 74%, when foreign countries responded with praise instead of silence. Moreover, praise proved inconsequential not only on average, but also for all three partisan clusters: Democrats, Independents, and Republicans (see online appendix). Thus, in the situations we portrayed, mixtures of shaming and praise performed no better than simple shaming strategies.

Design and Effects in Experiment 3

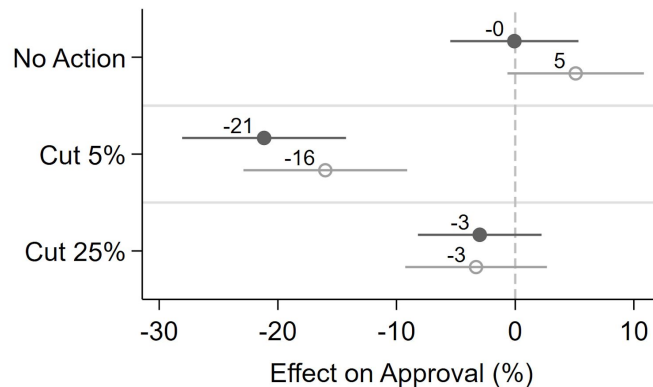
Earlier in the article, we hypothesized that the U.S. government could use rhetoric to negate the effects of foreign shaming. Our discussion centered on two rhetorical strategies the government could use: defiance and regret. We predicted that either strategy could work, but each would resonate with a different audience. Specifically, we expected that defiance would work better among Republicans than among Democrats, while regret would work better among Democrats than among Republicans.

To study the effects defiance, we introduced a new experimental condition in which the U.S. government responded defiantly to shaming. After learning that foreign governments had characterized U.S. behavior as shameful, participants read: “The U.S. government responded by saying that other countries had no right to meddle in our affairs. It said other countries were trying to hurt or control the United States.” We administered this condition to 1,214 participants in October 2018. By combining these data with measures from Experiment 1, we managed to estimate the effects of shaming with and without a defiant rebuttal.

Figure 7 presents the effects of shaming when the U.S. did not respond (solid dots), and when it responded with defiance (hollow dots). All treatment effects in Figure 7 were estimated with

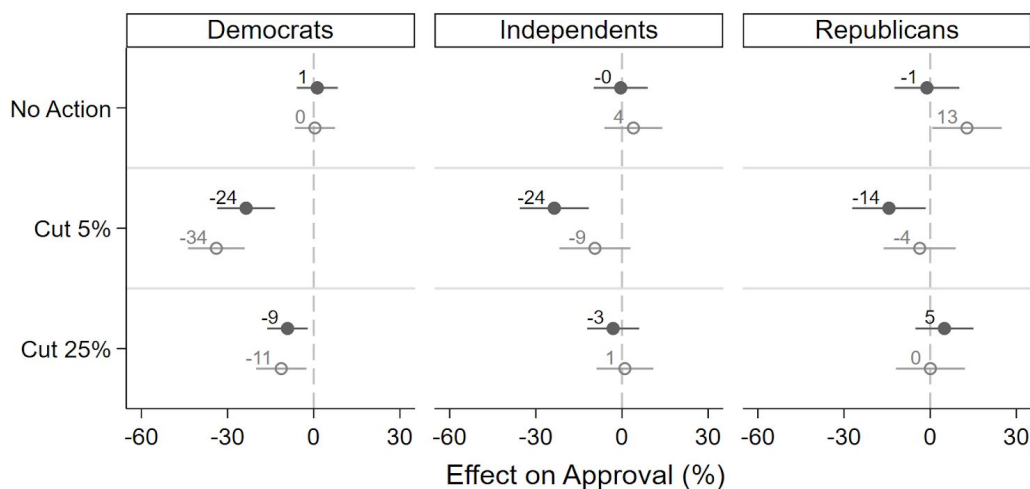
respect to a baseline in which foreign countries abstained from shaming. Figure 7 shows that defiance was not particularly effective at counteracting the effects of shaming. The treatment effects were roughly the same, with or without defiance.

Figure 7: Effects of shaming without a rebuttal (solid dots) and with a defiant rebuttal (hollow dots)



When we split the data by party affiliation, though, some interesting differences emerged (Figure 8). Consider the middle row, where shaming had the biggest effects on approval. Among Republicans, defiance counteracted most of the effects of shaming. The impact fell from 14 points to only 4 points, an effect that was statistically indistinguishable from zero. Defiance also persuaded Independents; the effect of shaming among that group shrank from 24 points to only 9 points. But defiance backfired among Democrats, who reacted more negatively to the combination of shaming and defiance than to shaming alone. Our experiments suggest that defiant rebuttals would be effective when pandering to Republicans and Independents, counterproductive when courting Democrats, and of little consequence on average.

Figure 8: Effects of shaming without a rebuttal (solid dots) and with a defiant rebuttal (hollow dots), by Party

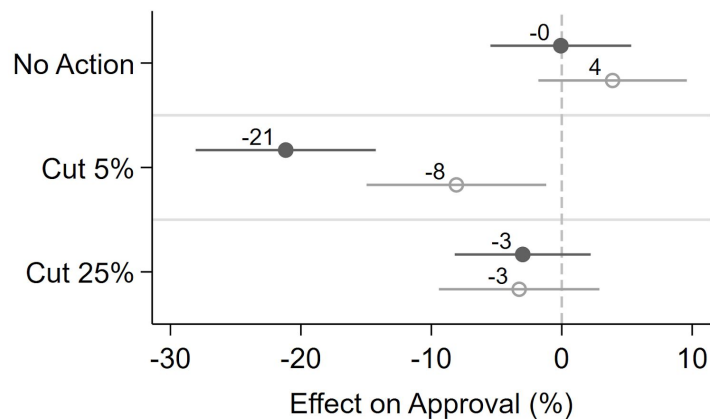


Design and Effects in Experiment 4

We introduced one more experimental condition to test whether and how regret might ameliorate the effects of shaming. After learning that foreign governments had shamed the U.S., participants read: “The U.S. government responded by saying that other countries were right. It apologized for not doing more and said it would work to reduce U.S. emissions in the future.” We administered this condition to 1,160 participants in October 2018. We combined the new data with information from Experiment 1 to estimate the effects of shaming with and without regret.

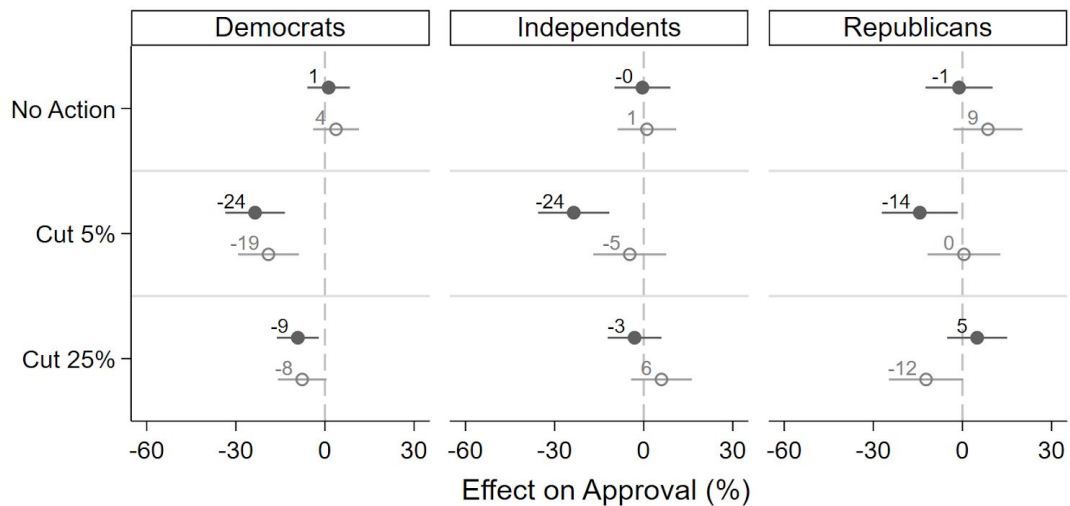
Figure 9 shows that regret was fairly effective in counteracting shaming. When the government passed laws to cut emissions by 5%, shaming alone caused approval to fall by 21 percentage points. Expressing regret reduced the effect to 8 percentage points, a much smaller though still substantial effect. Thus, a government that wanted to escape compliance could minimize the political consequences of shaming by acknowledging that the foreign critiques were valid, and by promising to make more cuts in the future.

Figure 9: Effects of shaming without a rebuttal (solid dots) and with regret (hollow dots)



Finally, we estimated the treatment effects within each partisan group (Figure 10). Focus again on the middle row, which shows how the public reacted when the government cut emissions by 5%. Shaming alone tended to reduce approval within each group. Contrary to expectations, though, regret was more effective in countering the effects of shaming on Republicans and Independents, than on Democrats. This again suggests that governments—especially ones catering to Republicans and Independents—could minimize the effects of shaming by apologizing and promising to do more in the future.

Figure 10: Effects of shaming without a rebuttal (solid dots) and with regret (hollow dots), by Party



Future Research/Conclusion

This paper provides some of the first micro-level evidence for how naming and shaming could contribute to treaty compliance. While this type of ‘normative tradition’ in the study of treaty compliance is just one of several traditions (Simmons 2010), it plays a crucial role in the Paris Agreement. We provide direct evidence of how members of the public might respond to their country’s climate policies being shamed.

Of course, there is a range of additional aspects of naming and shaming that we do not consider. For example, does the reaction to naming and shaming depend on Identify of the shamer (Hornsey et al. 2001)? Would responses differ depend on if the United States is shamed by an ally or not, or by a country that has met its commitments, or not? The effect of naming and shaming might also depend on whether or not other countries keep their commitments. While previous studies show that members of the public do not support lowering environmental effort if other countries reduce their effort (Tingley and Tomz 2014), it might be that the effect of shaming is reduced.

Another line of inquiry would investigate how to avoid backlash. For example, Hornsey et al. (2008) explore how the shaming party can combine their message with an indication that they face similar problems. Given the recent focus on defiance (e.g., Terman 2017) and norm erosion (e.g., Carnegie and Carson 2018) in international relations, this will an important area to investigate further. This may especially be the case in the environmental domain where incentives to pursue non-environmentally friendly policies are extremely strong.

Our experiment did not test the long-run sustainability of expressing regret. In Experiment 4, the U.S. government acknowledged that shamers were right, apologized for not doing more, and reiterated its commitment to reduce U.S. emissions in the future. This rhetorical strategy helped in the short run, but it is not clear how long the public would accept this excuse. If, as time passed, the U.S. failed to take additional action and bring itself into compliance, would shamers regain the rhetorical advantage? In this context, shamers might have an advantage because they can more readily continue their shaming whereas continually expressing regret and then not doing anything would be unsustainable.

Finally, in this research note we do not study the specific mechanisms that link shaming to the effects we observe. For example, does shaming generate concerns about the reputation of their country? Does it generate feelings of guilt within individuals? That is, do people internalize shame that is applied at the country level? And does shaming generate a greater motivation to be politically active? All of these are important questions in helping us understand how normative appeals can help to drive compliance.

References

Astbury, B., & Leeuw, F. L. (2010). Unpacking black boxes: mechanisms and theory building in evaluation. *American Journal of Evaluation*, 31(3), 363-381.

Ausderan, J. (2014). How naming and shaming affects human rights perceptions in the shamed country. *Journal of Peace Research*, 51(1), 81-95.

Bailey, J. L. (2008). Arrested development: the fight to end commercial whaling as a case of failed norm change. *European Journal of International Relations*, 14(2), 289-318.

Baldwin, D. A. (1971). The power of positive sanctions. *World Politics*, 24(1), 19-38.

Bawden, Tom. 'COP-21: Paris deal far too weak to prevent devastating climate change, academics warn', *Independent*, 8 Jan. 2016, <http://www.independent.co.uk/environment/climate-change/cop21-paris-deal-fartoo-weak-to-prevent-devastating-climate-change-academics-warn-a6803096.html>.

Brown, Donald (2015) " Why "Shaming" Is An Important Tool That Could Lead to Climate Change Action In Paris And Beyond"
<https://ethicsandclimate.org/2015/11/14/why-shaming-is-an-important-tool-that-could-lead-to-climate-change-action-in-paris-and-beyond/>

- Carnegie, A., & Carson, A. (2018). The Spotlight's Harsh Glare: Rethinking Publicity and International Order. *International Organization*, 72(3), 627-657.
- Cabus, S. J., & De Witte, K. (2012). Naming and shaming in a 'fair' way. On disentangling the influence of policy in observed outcomes. *Journal of Policy Modeling*, 34(5), 767-787
- Clark, K. (2018). The Paris Agreement: It's Role in International Law and American Jurisprudence. *Notre Dame Journal of International Comparative Law*, 8, 107.
- Cooley, A., & Snyder, J. (Eds.). (2015). *Ranking the world*. Cambridge University Press.
- Drezner, D. W. (1999). The trouble with carrots: Transaction costs, conflict expectations, and economic inducements. *Security Studies*, 9(1-2), 188-218.
- Druckman, J. N., & Lupia, A. (2016). Preference change in competitive political environments. *Annual Review of Political Science*, 19, 13-31.
- Dunlap, R. E., & McCright, A. M. (2011). Organized climate change denial. *The Oxford handbook of climate change and society*, 1, 144-160.
- Falk, Pamela, "Climate negotiators strike deal to slow global warming". CBS News. CBS Interactive Inc. 12 December 2015. Retrieved March 12 2019.
<https://www.cbsnews.com/news/cop21-climate-change-conference-final-draft-historic-plan/>
- Falkner, R. (2016). The Paris Agreement and the new logic of international climate politics. *International Affairs*, 92(5), 1107-1125.
- Finnemore, M., & Sikkink, K. (1998). International norm dynamics and political change. *International Organization*, 52(4), 887-917.
- Franklin, J. C. (2008). Shame on you: The impact of human rights criticism on political repression in Latin America. *International Studies Quarterly*, 52(1), 187-211.
- Gershowitz, A. M. (2008). Prosecutorial Shaming: Naming Attorneys to Reduce Prosecutorial Misconduct. *UC Davis Law Review*, 42, 1059.
- Gustafson, A., Bergquist, P., Leiserowitz, A., Maibach, E. (2019). *A growing majority of Americans think global warming is happening and are worried*. Yale University and George Mason University. New Haven, CT: Yale Program on Climate Change Communication.
- Hafner-Burton, E. M. (2008). Sticks and stones: Naming and shaming the human rights enforcement problem. *International Organization*, 62(4), 689-716.

- Hendrix, C. S., & Wong, W. H. (2013). When is the pen truly mighty? Regime type and the efficacy of naming and shaming in curbing human rights abuses. *British Journal of Political Science*, 43(3), 651-672.
- Hornsey, M. J., Oppes, T., & Svensson, A. (2002). "It's OK if we say it, but you can't": Responses to intergroup and intragroup criticism. *European Journal of Social Psychology*, 32(3), 293-307.
- Hornsey, M. J. (2005). Why being right is not enough: Predicting defensiveness in the face of group criticism. *European Review of Social Psychology*, 16(1), 301-334.
- Hornsey, M. J., & Imani, A. (2004). Criticizing groups from the inside and the outside: An identity perspective on the intergroup sensitivity effect. *Personality and Social Psychology Bulletin*, 30(3), 365-383.
- Jacquet, J., Hauert, C., Traulsen, A., & Milinski, M. (2011). Shame and Honour Drive Cooperation. *Biology Letters*, 7, 899–901. doi:10.1098/rsbl.2011.0367.
- Jacquet, J., & Jamieson, D. (2016). Soft but Significant Power in the Paris Agreement. *Nature Climate Change*, 6(7), 643–646. doi:10.1038/nclimate3006.
- Johns, L. (2014). Depth versus rigidity in the design of international trade agreements. *Journal of Theoretical Politics*, 26(3), 468-495.
- Keck, M. E., & Sikkink, K. (1999). Transnational advocacy networks in international and regional politics. *International Social Science Journal*, 51(159), 89-101.
- Kelley, J. G., & Simmons, B. A. (2015). Politics by number: Indicators as social pressure in international relations. *American Journal of Political Science*, 59(1), 55-70.
- Krain, M. (2012). J'accuse! Does naming and shaming perpetrators reduce the severity of genocides or politicides?. *International Studies Quarterly*, 56(3), 574-589.
- Lind, J. (2009). Apologies in international politics. *Security Studies*, 18(3), 517-556.
- Lupu, Y. (2013). The informative power of treaty commitment: using the spatial model to address selection effects. *American Journal of Political Science*, 57(4), 912-925.
- Malhotra, N., Monin, B., & Tomz, M. (2019). Does Private Regulation Preempt Public Regulation?. *American Political Science Review*, 1-19.
- Martini, Catherine. (2016). "Transparency: The Backbone of the Paris Agreement"
<https://envirocenter.yale.edu/transparency-the-backbone-of-the-Paris-Agreement>

- Mitchell, Ronald. (2017) "International Environmental Agreements Database Project", <https://iea.uoregon.edu/>
- Mitchell, S., & Hensel, P. R. (2007). International institutions and compliance with agreements. *American Journal of Political Science*, 51(4), 721-737.
- Murdie, A., & Peksen, D. (2013). The impact of human rights INGO shaming on humanitarian interventions. *The Journal of Politics*, 76(1), 215-228.
- O'Donnell, T. (2006). Naming and shaming: The sorry tale of Security Council Resolution 1530 (2004). *The European Journal of International Law*, 17(5), 945-968.
- Olick, J. K. (2013). *The politics of regret: On collective memory and historical responsibility*. Routledge.
- Panke, D., & Petersohn, U. (2012). Why international norms disappear sometimes. *European Journal of International Relations*, 18(4), 719-742.
- Parguel, B., Benoît-Moreau, F., & Larceneux, F. (2011). How sustainability ratings might deter 'greenwashing': A closer look at ethical corporate communication. *Journal of Business Ethics*, 102(1), 15.
- Rees, J. H., Klug, S., & Bamberg, S. (2015). Guilty conscience: motivating pro-environmental behavior by inducing negative moral emotions. *Climatic Change*, 130(3), 439-452.
- Risse-Kappen, T., Risse, T., Ropp, S. C., & Sikkink, K. (Eds.). (1999). *The power of human rights: International norms and domestic change* (Vol. 66). Cambridge University Press.
- Rosendorff, B. P., & Milner, H. V. (2001). The optimal design of international trade institutions: Uncertainty and escape. *International Organization*, 55(4), 829-857.
- Saffrey, C., Summerville, A., & Roese, N. J. (2008). Praise for regret: People value regret above other negative emotions. *Motivation and Emotion*, 32(1), 46-54.
- Skeel, D. A. (2001). Shaming in corporate law. *University of Pennsylvania Law Review*, 149(6), 1811-1868.
- Sherman, L. W. (1993). Defiance, deterrence, and irrelevance: A theory of the criminal sanction. *Journal of research in Crime and Delinquency*, 30(4), 445-473.
- Simmons, B. (2010). Treaty compliance and violation. *Annual Review of Political Science*, 13, 273-296.

- Simmons, B. A., & Hopkins, D. J. (2005). The constraining power of international treaties: Theory and methods. *American Political Science Review*, 99(4), 623-631.
- Taebi, B., & Safari, A. (2017). On Effectiveness and Legitimacy of 'Shaming' as a Strategy for Combatting Climate Change. *Science and Engineering Ethics*, 23(5), 1289-1306.
- Tallberg, J. (2002). Paths to compliance: Enforcement, management, and the European Union. *International organization*, 56(3), 609-643.
- Tangney, J. P., & Dearing, R. L. (2003). *Shame and Guilt*. Guilford Press.
- Terman, R. (2017). Rewarding resistance: Theorizing defiance to international norms. Unpublished Manuscript.
- Terman, R., & Voeten, E. (2018). The relational politics of shame: Evidence from the universal periodic review. *The Review of International Organizations*, 13(1), 1-23.
- Tingley, D., & Tomz, M. (2014). Conditional cooperation and climate change. *Comparative Political Studies*, 47(3), 344-368.
- Tingley, D., & Tomz, M. (2019) International Commitments and Domestic Opinion: The Effect of the Paris Agreement on Public Support for Policies to Address Climate Change. Working paper.
- Van Der Pligt, J., & Vliek, M. (2016). *The Psychology of Influence: Theory, research and practice*. Routledge.
- Von Stein, J. (2005). Do treaties constrain or screen? Selection bias and treaty compliance. *American Political Science Review*, 99(4), 611-622.
- Wachman, A. M. (2001). Does the diplomacy of shame promote human rights in China?. *Third World Quarterly*, 22(2), 257-281.