British Journal of Political Science (2025), **55**, e4, 1–19
doi:10.1017/S0007123424000267

British Journal of Political Science

ARTICLE

How Does Shaming Human Rights Violators Abroad Shape Attitudes at Home?

Lotem Bassan-Nygate

Assistant Professor of Public Policy, John F. Kennedy School of Government, Harvard University, Cambridge, MA, USA
Email: lbassan@hks.harvard.edu

(Received 13 March 2023; revised 5 January 2024; accepted 28 June 2024)

Abstract

Does shaming human rights violators shape attitudes at home? A growing literature studies the effect of shaming on public attitudes in the target state, but far less is known about its effect in countries initiating the criticism – that is the *shamers*. In this article, I theorize that when governments shame human rights violators they shape both government approval and human rights attitudes at home. Utilizing two US-based survey experiments, I demonstrate that by shaming foreign countries, governments can improve their image at home and virtue signal their dedication to human rights. At the same time, shaming can modestly shape tolerance towards certain domestic human rights violations. I consider the generalizability of my results through comprehensive supplementary analyses, where experimental insights are corroborated with cross-national observational data. Overall, my findings can provide valuable insight into governments' incentives to engage in foreign criticism.

Keywords: naming and shaming; human rights; public opinion

In October 2022, 142 countries supported a UN General Assembly resolution calling Russia to reverse its attempts to annex regions in Ukraine, condemning 'violations of international law and human rights' (Lederer 2022). That same month, fifty countries issued a joint statement condemning the Chinese government's human rights violations and prosecutions of Uyghurs and other Turkic Muslims in Xinjiang (Lazaroff 2022), and six countries, including the US, Germany, and the UK, condemned human rights abuses in Ethiopia (US State Department 2022). These are just a few examples of one form of naming and shaming – governments' attempts to single out and publicly criticize foreign governments for their policies.

For the most part, research on shaming has focused on its effects on those *targeted by shame*, particularly in the context of human rights organizations and transnational advocacy networks (Hafner-Burton 2008; Hendrix and Wong 2013; Keck and Sikkink 1999; Krain 2012; Murdie and Davis 2012). Recent studies draw attention to the actions taken by states to name and shame foreign governments, and how those actions impact policies (Becker et al. 2024; Myrick and Weinstein 2022; Terman and Voeten 2018) and public opinion (Gruffydd-Jones 2019; Tingley and Tomz, 2022) in the target state. However, far less is known about the extent to which shaming affects public attitudes in the countries initiating the criticism, that is the *shamers*.

Drawing attention to shamers is important for several reasons. First, existing research provides only a partial picture by focusing on audiences in the target state, ignoring one of the most important attributes of shaming – that it is a public event. Indeed, shaming is a foreign policy

behaviour that is observed by and likely intended for a variety of different actors – both *foreign and domestic*. As such, focusing on the effects of shaming on one actor – the target – de-emphasizes its global impact and masks the incentives to shame, which precede the target's reactions.

Second, focusing on shamers can provide behavioural foundations for the prevalence of naming and shaming even when it is ineffective. A growing body of work suggests that government shaming, especially in the context of human rights, can create a backlash effect (Adler-Nissen 2014; Gruffydd-Jones 2019; Snyder 2020; Terman 2017). It is thus puzzling that governments continue to shame even when shaming backfires. One way to unpack this puzzle is by shifting our focus from targets of shame to its senders, focusing on the domestic political benefits of shaming that accrue even when it does not change state behaviour.

Building on these insights, I develop a theory that puts the shamers front and centre, explaining how and why shaming affects domestic approval and attitudes towards human rights at home. I argue that by shaming other countries on international platforms, governments 'virtue signal' their dedication to human rights and appease domestic audiences by appearing more engaged in foreign policy – improving perceptions of domestic respect for human rights. Additionally, I posit that shaming can shape individual support for human rights. On the one hand, shaming might constrain leaders domestically by fostering increased support for human rights to maintain consistency. On the other hand, when governments criticize human rights violators, they may bolster their moral standing, potentially legitimizing controversial policies and fostering greater acceptance of domestic violations of human rights.

To test this theory, I conducted two survey experiments in the US. In the first experiment, respondents were presented with a hypothetical scenario about a foreign country that violates human rights, manipulating whether the US government criticized the foreign country. The results indicated that respondents informed of the US government's shaming of the human rights violator were more inclined to endorse the government's actions and view the US government as committed to human rights. However, paradoxically, they also displayed modestly increased support for potential human rights violations within the US. In a subsequent experiment, I show that respondents reward their government for shaming a foreign government even when doing so bears considerable costs, namely when the target state is a US ally.

Acknowledging the limitations of survey experiments, I further discuss a series of supplementary analyses evaluating the external validity of my findings. Specifically, causally identified findings are corroborated with cross-national trends, demonstrating that governments' likelihood to engage in shaming, proxied by governments' tendency to shame other countries in the UN Universal Periodic Review (UPR), is correlated with an increase in incumbent voting and favourable perceptions of domestic respect for human rights. Additionally, I estimate the external robustness of my findings, demonstrating that my estimates are robust to populations that are highly different from my experimental samples.

I make three central contributions to the existing literature. First, by theorizing about the effects of shaming on attitudes at home, I shift the spotlight from targets to shamers and consider the broader implications of shaming. By shedding light on domestic audiences, we can begin to unpack the incentives of countries to engage in global criticism in the first place. Doing so may provide micro-foundations (Kertzer 2017) for puzzling phenomena such as the pervasiveness of shaming, even when it fails to change the targets' behaviours.

Second, I extend the literature on shaming by integrating works from public opinion (Guisinger and Saunders 2017; Tomz and Weeks 2020) with recent work from social and moral psychology (Kouchaki 2011; Lasarov and Hoffmann 2020; Merritt, Effron, and Monin 2010) to theorize about individual-level motivations. In doing so, I explain how and why shaming other countries shapes domestic public opinion and provide suggestive evidence for two central mechanisms – perceptions of morality and power – driving the effects I identify.

Third, by corroborating experimental evidence with observational cross-national trends, I marry two prominent methodological approaches in the literature. I join the recent movement in

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

the shaming literature towards experimental designs to identify the effects of shaming on individual attitudes (Greenhill and Reiter 2022; Koliev, Page, and Tallberg 2022; Tingley and Tomz 2022) and complement it with supplementary analyses that probe the generalizability of the results and take seriously valid concerns of external validity (Egami and Hartman 2023).

My findings underscore governments' incentives to engage in public shaming. I show that, by shaming human rights violators, governments can be rewarded domestically without making any practical improvement in human rights. Under certain circumstances, citizens may even become more tolerant of domestic violations of human rights in light of the shaming. This suggests that, in the long run, governments might enjoy domestic support while receiving domestic leeway to violate international norms. I further discuss these implications in the concluding section of this article.

Shifting the Spotlight From Targets to Shamers

Governments are often publicly criticized for violating human rights. Past research has focused, for the most part, on the effects of shaming on *targets of shame*. This body of work dates back more than two decades, with the seminal work of Keck and Sikkink (1999) and Risse et al. (1999) who drew attention to the ways in which transnational advocacy networks can pressure states to adopt norm-conforming policies. Over the years, a rich literature on human rights shaming has sprung up to study empirically how shaming shapes domestic policies (Allendoerfer, Murdie, and Welch 2020; Hafner-Burton 2008; Hendrix and Wong 2013; Krain 2012; Murdie and Davis 2012) and public opinion (Ausderan 2014; Bracic and Murdie 2020; Greenhill and Reiter 2022) in targeted states.

Much attention has been given in this literature to the effects of actions taken by human rights organizations (HROs) and their partners in civil society to pressure transgressive governments (Allendoerfer, Murdie, and Welch 2020; Hendrix and Wong 2013; Keck and Sikkink 1999; Murdie and Davis 2012). This article focuses on another form of shaming – *government shaming* – which refers to the actions taken by state actors to name and shame foreign governments, often labelled as a class of human rights diplomacy (Myrick and Weinstein 2022).

Government shaming may include press releases, official statements, speeches, or reports – released independently or in international forums like the UN General Assembly or the Universal Periodic Review – to draw attention to the human rights abuses of a foreign government. While these efforts proved to be successful when paired with secret diplomacy in a recent US Campaign (Myrick and Weinstein 2022), a growing body of work suggests that government shaming is often counterproductive (Snyder 2020) – decreasing advocated policies (Becker et al. 2024; Terman 2017) and public attitudes towards human rights (Gruffydd-Jones 2019). Reportedly, policy circles hold similar positions, widely accepting the ineffectiveness of government shaming in promoting rights (Neier 2018).

In light of this growing evidence, it is puzzling that governments engage in human rights shaming despite its tendency to backfire. In fact, shaming appears to be a prevalent tactic used by a variety of states worldwide. This is evident in Fig. 1, which demonstrates the global distribution of shamers in the UPR between 2008 and 2020, where any government can shame any country after reviewing its human rights record. France is the most frequent shamer, with an average of 190 shaming instances per year, followed by Spain (173) and Canada (158). Most countries appear to engage in human rights shaming to varying degrees.

We can begin to address the puzzling prevalence of government shaming, despite its tendency to backfire, by diverting our attention from its targets to the countries that shame. Studying the *domestic* effects of shaming in the shamers' countries may shed light on the incentives that motivate governments to engage in shaming in the first place, independent of its ability to change state behaviour or public attitudes in shamed societies. Learning that shaming is effective in

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

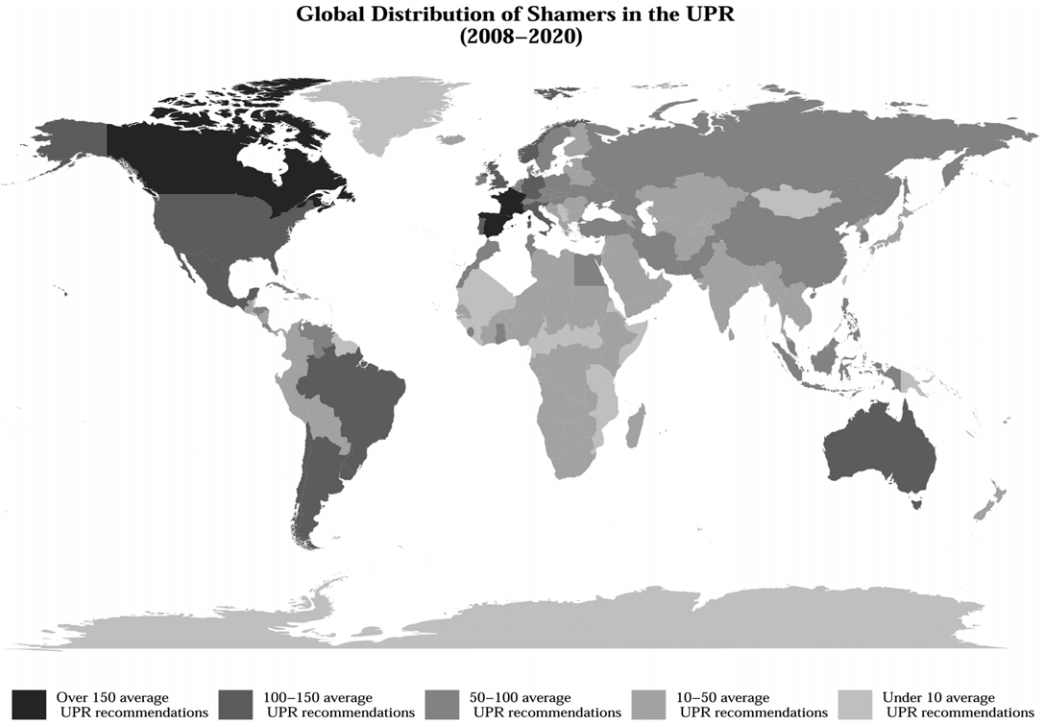

Figure 1. The average number of times each country has shamed other countries in a given year on the Universal Periodic Review (UPR) between 2008 and 2020.

domestic realms rather than international ones may provide behavioural foundations that explain its pervasiveness.

Indeed, while government shaming is often understood to be a tool used abroad, like many diplomatic efforts, it is most likely a two-level game (Putnam 1988) simultaneously observed by different audiences. Much like the overt nature of threats and troop deployment underscored by the audience costs theory (Fearon 1994), shaming functions primarily as a *public* event assessed by domestic audiences. Condemnations, whether carried out on international platforms like the UN (Lazaroff 2021), conducted during bilateral meetings (McCurry 2021), or stated in official government reports (US State Department 2020) receive media coverage and public attention.

Hence, various actors including the public in the target country, global human rights activists, and the general public back home are exposed to shaming and may perceive it and react to it in different ways. Thus, limiting our research agenda to target countries de-emphasizes the global impact of shaming, which may have reached well beyond shamed societies, ultimately shaping public attitudes towards human rights internationally. Such an impact may contribute to the overall (in)effectiveness of the human rights regime and shed light on the incentives of governments to shame in the first place.

While little attention has been allocated to shamers in the literature, some notable exceptions include recent works by Terman and Voeten (2018) and Terman and Byun (2022), who focus on shamers by analyzing shaming patterns in the UPR. These works provide rich theoretical and empirical insights into the politicization of the human rights regime, demonstrating that states shame selectively. They are thus primarily focused on the causes of shaming rather than their domestic consequences. In what follows, I argue that shaming entails certain domestic rewards for shamers. In the long run, these benefits may make international criticism a particularly appealing foreign policy behaviour for governments.

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

Shaming and Government Approval

What are the rewards associated with shaming? Terman and Byun (2022) theorize that countries shame selectively because they face a dilemma – on the one hand, they wish to maintain relationships with their allies; on the other hand, they want to collect the social rewards associated with shaming. These rewards, they argue, include appearing to be defenders of human rights and 'good citizens' of the international regime.

Indeed, in shaming human rights violators, governments likely appease third parties like global activists, NGOs, and other countries who are dedicated to the human rights project. However, shaming other countries may also include domestic *material* rewards such as government approval and increased vote share. I argue that by shaming other countries, governments can collect domestic rewards because the general public evaluates shaming in positive ways for several reasons.

First, domestic actors within the shamer's country who advocate human rights issues are particularly likely to support governments that vocalize human rights concerns. Indeed, leaders are often criticized for sidelining issues related to human rights and failing to hold other countries accountable for violations. Consider, for instance, criticism aimed at US President Joe Biden's administration for its reluctance to pressure US allies such as Egypt, Turkey, and Saudi Arabia in the face of repression (Wootson Jr 2021) or criticism of European governments for not taking a harder approach towards Israel (Haddad 2021). Domestic human rights activists and other concerned citizens often expect their government to hold human rights violators accountable. When their government publicly condemns violators, they view their leadership as embracing human rights issues.

Naturally, one may expect shaming to play a more significant role in appeasing leftists and doves, who are often considered more supportive of human rights (McFarland and Mathews 2005; Rathbun et al. 2016). However, when considering the politicized nature of human rights shaming (Krasner 1999), there are reasons to expect shaming to increase government approval across the board. Hawks, for example, may take particular pleasure in stigmatizing adversaries (Ash and Dolan 2021), viewing shaming as a form of global competition. Consider Republican criticism of US President Joe Biden for arguably taking a soft approach towards China's human rights violations (DeMarche 2021). It is possible that these criticisms were motivated by a sense of a missed opportunity to confront a global rival and demonstrate resolve rather than a sincere concern for human rights in China. Thus, while hawks may not be driven by concerns about human rights issues per se, they are also likely to support condemnations, particularly when the target is an adversary.

Indeed, recent experimental work suggests that the public across the political map, at least in the US, is equally sensitive to human rights issues when forming foreign policy attitudes (Tomz and Weeks 2020). Hence, although hawks and doves may be ultimately driven by different concerns (Brutger and Kertzer 2018), I argue that they are both likely to evaluate administrations positively in light of shaming. While it is possible that hawks may be driven by concerns about *power* or 'acting tough' on the international stage, doves may be driven by concerns regarding *morality*; there is no reason to expect either group to be particularly unsupportive *of shaming human rights violators. Hence, I expect the general public to evaluate human rights shaming in positive terms, perceiving their government as more proactive and engaged in foreign policy. It follows that:*

Hypothesis 1. *Shaming human rights violators increases domestic government approval.*

Shaming and Human Rights Perceptions and Attitudes

The benefits associated with shaming may surpass government approval. By shaming other countries, governments 'virtue signal' their dedication to human rights, which may have different

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

implications for public attitudes and perceptions. Shaming may serve as an elite cue, emphasizing the importance of human rights to domestic audiences. By criticizing other countries' human rights abuses, leaders may signal to their public that they are dedicated to human rights issues and that these topics are on the political agenda.

Indeed, research suggests that public opinion is often susceptible to elite cues, especially around foreign affairs, which are distant from most voters' everyday concerns (Guisinger and Saunders 2017; Holsti 2004). Hence, although shaming may be cheap talk in the sense that shamers do not necessarily respect human rights back home, I argue that citizens are inclined to believe that governments who take foreign policy stances against violators are also more likely to respect human rights themselves, in part because they call attention to their own policies and behaviour. It thus follows that:

Hypothesis 2. *Shaming human rights violators increases favourable perceptions of domestic respect for human rights.*

At the same time, this elite cue may also shape domestic attitudes towards human rights, thereby increasing domestic support. Citizens who are driven by concerns about international image may become more supportive of human rights to avoid appearing hypocritical abroad. Recent experimental evidence suggests that individuals are aversive to hypocritical behaviour (von Grundherr, Jauernig, and Uhl 2021) and care about their country's international status (Powers and Renshon 2023) and reputation (Brutger and Kertzer 2018). Taken together, these works suggest that the public may become more supportive of maintaining a good human rights record at home to avoid criticism abroad for applying double standards.

This expectation bears resemblance to the logic of the audience costs theory, whereby individuals are averted by their leaders' inconsistent behaviour (Kertzer and Brutger 2016; Tomz 2007). Indeed, if individuals interpret human rights violations in the aftermath of shaming as a form of 'backing down' from a public statement, they might increasingly advocate for policies aligning with international law and human rights standards when faced with information on human rights shaming.

In addition, by criticizing other countries, and especially global rivals, governments likely draw attention to their own human rights practices and invite counter-shaming (Búzás and Bassan-Nygate 2024). Take for instance Beijing's criticism of US double standards (Fearnow 2021); following criticism over Uyghur human rights violations, the Chinese communist party argued the US has 'mountains of human rights problems in its own country', referencing racial inequality and the murder of George Floyd (Moritsugu 2021). Since such criticism is perceived as damaging to a country's image (Simmons 2009), the public may become more supportive of norm-confirming behaviour that abides by human rights after their government has taken a public stance on these issues. It follows that:

Hypothesis 3a. *Shaming human rights violators increases domestic support for human rights.*

Alternatively, shaming human rights violators may *decrease* domestic support for human rights. The literature on 'moral licensing' in psychology is a robust body of work that suggests that past moral behaviour can increase immoral, unethical, or otherwise problematic behaviours and attitudes (Merritt, Effron, and Monin 2010). Individuals' past good deeds allow them to engage in otherwise unethical behaviours without fearing to appear immoral. Individuals thus build 'moral credentials' when engaging in moral behaviour and draw on these past behaviours when they engage in immorality. Thus, to the extent that shaming human rights violators is perceived as moral, individuals may be *more likely* to support domestic violations of human rights.

Indeed, numerous studies demonstrate how moral behaviour, such as supporting a Black political candidate or buying green products, can license problematic and immoral behaviour such as engaging in prejudice (Effron, Cameron, and Monin 2009), cheating, and stealing

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

(Mazar and Zhong 2010). Some studies find that the moral licensing effect is so robust that even the prospect of morality, or simply planning to engage in moral behaviour in the future, can license current immoral attitudes and behaviours (Cascio and Plant 2015).

One may wonder whether moral licensing, a theory that primarily focuses on individuals rather than groups, applies to international politics. After all, in the moral licensing theory, individuals engage in moral behaviour themselves, whereas, in my application to IR, individuals observe their leaders' moral actions. However, recent findings from social psychology suggest that learning of one's in-group's past moral actions shapes one's moral attitudes and behaviours (Lasarov and Hoffmann 2020). For example, when learning of their group members' prior non-discriminatory behaviour in selecting Hispanic applicants, respondents gave more discriminatory ratings to Hispanic applicants (Kouchaki 2011). This suggests that individuals, who derive a sense of identity from their social associations (Tajfel 2010), build moral credentials based on their in-group's moral behaviour.

Applying this theoretical framework to naming and shaming, one might expect that criticizing human rights violators serves as a moral credential, allowing respondents to support otherwise immoral policies – such as violating human rights. Importantly, moral licensing is a non-conscious effect that operates by boosting perceptions of *morality*, increasing preferences to engage in relatively immoral behaviour by dampening the negative attributions associated with it (Simbrunner and Schlegelmilch 2017). In this sense, H2 serves as a complementary hypothesis to moral licensing, as the belief that their country is generally more respectful of human rights may boost respondents' perception of morality. These perceptions may, in turn, decrease support for domestic respect for human rights.

Hypothesis 3b. *Shaming human rights violators decreases domestic support for human rights.*

The Role of the Targeted Country

My theory suggests that shaming may be an effective tool for governments, which can increase their approval and perceptions of domestic respect for human rights while simultaneously obtaining domestic leeway to violate international norms. I assess the scope condition of this theory by considering the role of the targeted country in attenuating these effects. Although it is possible that shaming is effective regardless of the target's identity, I consider whether citizens support their government even when it shames a strategic and economic ally whose shaming has higher costs.

Indeed, shaming may be less appealing when the target is an ally of the shamer. Citizens may favor criticizing an adversary with whom their country has no security alliances or economic ties. Such criticism does not impose any threats to existing relationships and may thus be perceived as costless. Alternatively, shaming allies may be perceived as more genuine and ultimately more effective. Recent work suggests that, while criticizing allies is less common, such criticism is ultimately more likely to be accepted by targets (Terman and Voeten 2018).

Finally, it is possible that citizens hold both of these perceptions simultaneously and thus support shaming regardless of the identity of the shamed country. In the following section, I evaluate these hypotheses in two US-based survey experiments. Subsequently, I assess the generalizability of my findings by corroborating the causal evidence with observational trends and a series of supplementary analyses estimating the external robustness of my findings.

Shaming and Domestic Attitudes: Experimental Evidence

My theory suggests that publicly shaming other countries can shape government approval as well as human rights attitudes and perceptions. I tested this theory in two survey experiments in the US. In the first experiment, I manipulated information regarding the US government's criticism of

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

A country violates human rights; it imprisons and tortures some of its citizens because of their beliefs and silences human rights defenders and activists. The U.S. government [called the country out and publicly criticized it for its violation of human rights in the United Nations./ did not criticize the country in the United Nations or make any public statement about the country's human rights abuses.]The country continued to violate human rights.

Figure 2. Experiment I Experimental vignette. Respondents were randomly assigned to one of the two conditions. Red text signifies treatment condition and blue text signifies control condition.

a country that violates human rights on the UN platform. The second experiment included an additional treatment arm randomizing information regarding the relationship of the shamed country with the US. Doing so allows me to examine whether my theory holds when the stakes of shaming are higher. In the following section, I present the two experiments' research designs and discuss their results.

Research Design

The first and second survey experiments were fielded in the US during March 2022 and June 2022, respectively, using Lucid's survey platform. In the first experiment, 1,200 attentive respondents were recruited; in the second experiment, 3,000 attentive respondents were recruited. Respondents were matched (through Lucid) to the general US population, based on sex, age, race, and region of residence. Descriptive statistics for both samples are reported in Tables A1 and B1.

The surveys were programmed online using Qualtrics. In both experiments, respondents were exposed to the experimental vignette following a screening attention check and a party ID question (see sections A.9 and B.7). In the initial experiment, respondents were randomized into the conditions depicted in Fig. 2. All respondents were told about a hypothetical scenario in which a foreign country engages in human rights violations. Information about whether the US government publicly condemned the country in the UN for its behaviour was manipulated with a probability of 0.5. I follow Kertzer and Brutger (2016), and control for policy outcomes by fixing the final result of the scenario across both conditions: the other country continued to engage in violations. Doing so allows me to ensure treated respondents do not assume shaming leads to fewer violations in the target state, creating a harder test for my theory – testing the domestic benefits of shaming even when it does not change the target's behaviour.

The second experiment included an additional treatment arm, which provided subsequent information about the shamed target, as depicted in Fig. 3. Respondents in the *ally* condition were told that the target country has security and economic ties with the US while respondents in the *non-ally* condition were told that the target country has neither security nor economic ties with the US. This results in a fully crossed 2×2 design, where respondents were randomized into one of the four conditions with a probability of 0.25.

After reading the vignette, respondents in both studies were presented with several questions that I employ as outcomes. To test my primary hypothesis, H1, regarding government approval, participants were asked how much they approved or disapproved of the way the US government handled the situation in the scenario. Possible responses ranged from 'strongly approve' to 'strongly disapprove' on a 7-point scale, and have been re-coded such that higher numbers indicate more support.

To test my second hypothesis regarding the effect of shaming on perceptions of domestic respect for human rights, I measure human rights perceptions. Respondents were asked how much they believe the US government in the passage values and respects human rights, on a scale of 1 to 7, where higher numbers indicate more respect for human rights.

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

- A country violates human rights; it imprisons and tortures some of its citizens because of their beliefs and silences human rights defenders and activists.
- The country is [a U.S. ally. It has signed a military alliance with the U.S. and has high levels of trade with the U.S. / not a U.S. ally. It has not signed a military alliance with the U.S. and does not have high levels of trade with the U.S.]
- The U.S. government [called the country out and publicly criticized it for its violation of human rights in the United Nations./ did not criticize the country in the United Nations or make any public statement about the country's human rights abuses.]
- The country continued to violate human rights.

Figure 3. Experiment II Experimental vignette. Respondents were randomly assigned to one of the four conditions. Teal text signifies the ally condition, violet signifies the non-ally condition, red text signifies the shaming condition, and blue text signifies the no-shaming condition.

To test my third hypothesis regarding the effect of shaming on support for human rights, I follow previous work (Wallace 2013) and measure support for a specific human rights policy, which was mentioned in the vignette – the use of torture. Respondents report how much they support or oppose the use of practices like torture that violate human rights in the US; the responses range from 'strongly approve' to 'strongly disapprove' on a 7-point scale, and have been re-coded such that higher numbers indicate stronger opposition to torture.

I further include questions regarding two potential mechanisms – perceptions of US morality and perceptions of US power, which range from very moral/powerful to very immoral/weak, on a scale of 1 to 7, a manipulation check, and a question that gauges potential confounders by asking whether respondents thought of a specific country when reading the vignette (see Sections A.5, A.6, B.4, and B.6 for an analysis). The surveys concluded with a host of demographic questions.

I estimated OLS models identifying the effect of the shaming treatment on each outcome of interest (relating to government approval, perceptions of respect for human rights, and support for human rights). To enhance the precision of my estimates, I employ additional models including several demographic controls (sex, age, race, education, state of residence, and partisanship). In the second experiment, I also estimated the effect of the shaming treatment, controlling for the identity of the ally to enhance the estimates' precision. I further estimate the interaction effect of the shaming and ally treatments, considering whether the effects are attenuated by perceptions of the human rights violator.

Shaming Shapes Attitudes at Home

In Fig. 4, I report the main results of the initial experiment. I look at the average treatment effects (ATEs) of the shaming treatment on government approval, human rights perceptions, and opposition to the use of torture in the US. My results show a large statistically significant effect when testing my main hypothesis regarding government approval: learning that the US government shamed the country that violated human rights increases government approval by approximately 2.3 points, from 3 to 5.3 on a 7-point scale ($p < 0.0001$). This sizable effect is equivalent to over one standard deviation.

Next, I consider the effectiveness of shaming human rights violators on perceptions of the US's respect for human rights. The treatment effect has a similarly sizable effect on human rights

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

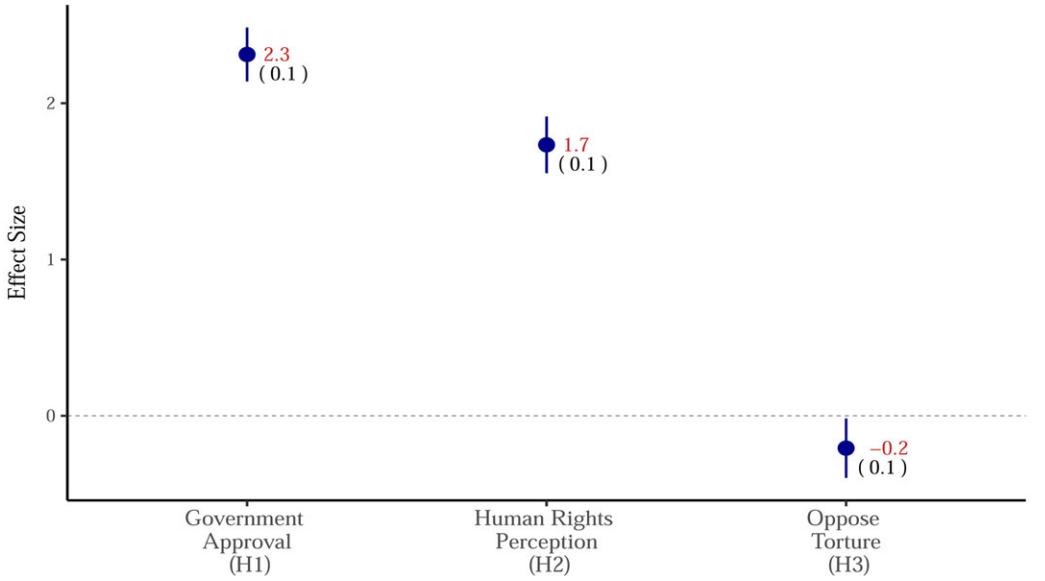

Figure 4. Treatment effect on main outcomes of interest (Experiment I). The *X*-axis represents three outcomes of interest – government approval, perceptions of respect for human rights, and opposition to torture, ranging from 1–7. Regression estimates are marked in red and robust standard errors are in parentheses.

perceptions, increasing the belief that the US government respects human rights by approximately 1.7 points, from 3.7 to 5.4 on a 7-point scale ($p < 0.0001$), which is equivalent to approximately 91 per cent of a standard deviation.

At the same time, shaming appears to modestly license problematic behaviour in the form of decreased opposition to the use of torture in the US. I find that respondents in the shaming condition were less likely to oppose torture by 0.2 points, from 5.8 to 5.6 on a 7-point scale ($p = 0.03$). Notably, this effect is small (equivalent in size to approximately 11 per cent of a standard deviation), which may be contingent on two facets – one theoretical and the other methodological. Theoretically, moral licensing manifests as a subconscious mechanism, which may explain why psychology literature reports modest effects (Simbrunner and Schlegelmilch 2017). Methodologically, it's plausible that my survey design introduced social desirability bias, where respondents felt compelled to oppose the use of torture, thus posing a harder test of the hypothesis (see Fig. B9).

Overall, all of the effects presented in Fig. 4 remain substantively similar when controlling for various pre-treatment demographic controls provided by Lucid, including sex, age, race, education, state of residence, and partisanship (see Table A2 in the Supplementary Material).

In the Supplementary Material (section A.2), I further report results from a pre-registered analysis testing the effect of shaming on a broader measure of support for international law (developed by Bayram (2017)). Interestingly, I identify a small but statistically significant effect in the opposing direction, whereby shaming *increased* respondents' obligations to international law, which is in contrast to H3b but per H3a. This finding is puzzling as I did not anticipate the shaming treatment to affect these two measures in opposing directions. One (post hoc) explanation may relate to the (in)effectiveness of moral licensing across different domains. Indeed, while there has been some evidence of cross-domain moral licensing (Mazar and Zhong 2010), recent follow-ups were unable to fully replicate this finding (Urban, Bahník, and Kohlová 2019). It is thus possible that the moral licensing effect simply does not extend to the domain of international law. Given the weak effect size, and since I was not able to replicate this finding in my

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

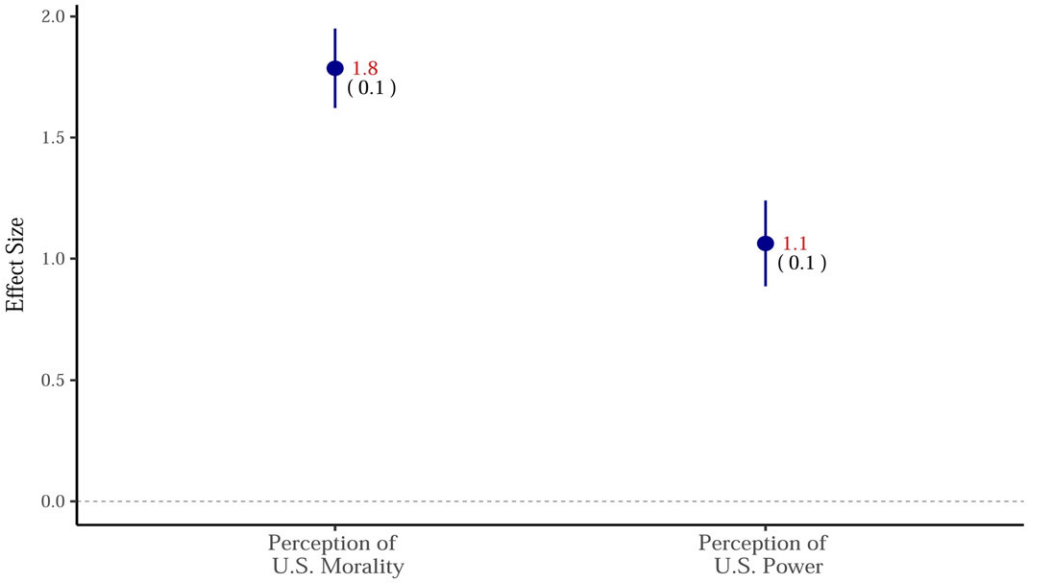

Figure 5. Treatment effect on potential mechanisms (Experiment I). The *X*-axis represents two outcomes of interest – perceptions of US morality and power, ranging from 1–7. The *Y*-axis represents the effect size for each model. Regression estimates are marked in red and robust standard errors are in parentheses.

second experiment (see table B10), future research should examine more carefully whether shaming can simultaneously shape favourable attitudes towards international law or adjacent domains.

Perceptions of Morality and Power as Mediators

Thus far, I have shown that shaming human rights violators increases government approval and perceptions of respect for human rights while modestly decreasing opposition to torture. However, one may wonder why these patterns emerge. I explore two central mechanisms that may drive some of the effects identified in the initial experiment – perceptions of power and morality.

As explicated in the theoretical section, treated respondents may be more likely to approve of their government simply because they believe shaming human rights violators is the ethical thing to do, and thus find their government to be more moral. An additional mechanism may relate to the US position vis-à-vis other countries, suggesting that governments can assert power and dominance by holding violators accountable and attempting to influence their domestic policies.

To explore whether perceptions of morality and power are potential mechanisms, I employed two post-treatment questions about US morality and power (see Section A.9 of the Supplementary Material) and analyzed whether responses were shaped by my shaming treatment. As demonstrated in Fig. 5, I find that treated respondents were more likely to view the US government in the story as more moral and more powerful. Learning that the US shamed the human rights violator increases perceptions of US morality by 1.8 points, from 3.4 to 5.3 on a 7-point scale ($p < 0.0001$) and increases perceptions of US power by approximately 1.1 points, from 3.7 to 4.9 on a 7-point scale ($p < 0.0001$). These effects are equivalent in size to about 0.64 and one standard deviation, respectively. In the Supplementary Material, I conduct a mediation analysis utilizing Imai et al.'s 2010 mediation package to consider whether morality and power mediate the effect identified in H1. The results of the analyses are depicted in Fig. A2. In both models, the average causal mediation effect (ACME), the direct effect, and the total effect are positive and statistically significant at $p < 0.001$.

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

I further consider whether perceptions of morality drive the effect of my shaming treatment on decreased opposition to torture. Since I hypothesize that shaming human rights violators operates as a moral credential, I expect the morality indicator to mediate the effect between shaming, and decrease support for human rights. In other words, I expect that shaming will license immoral behaviour if it is perceived to be moral. Thus, I further conducted a mediation analysis where the perception of morality was the mediator and opposition to torture was the outcome. The result, reported in Fig. A1, suggests that the average causal mediation effect and the total effect are negative and statistically significant at $p < 0.05$.

Importantly, all mediators described above were not randomly assigned and, thus, there was a concern for omitted variable bias that accounts for both the mechanism and outcome, violating the 'sequential ignorability' assumption necessary for mediation analysis (Imai, Keele, and Tingley 2010). I discuss these concerns in Section A.4 and conduct a sensitivity analysis, calculating the mediation effect for different magnitudes of a potential confounder's effect on the mediator and the outcome. I also tested whether findings in this first experiment may have been affected by potential issues relating to confounding and information leakage (Dafoe, Zhang, and Caughey 2018) (Section A.6) or attrition (Section A.8). I did not find evidence supporting this concern.

The Effect of Shaming by Allies and Non-Allies

Although there seems to be little concern relating to information leakage in the first survey experiment, meaning that the treatment did not differentially cause respondents to think of a specific country (see Table A4), the findings from the first experiment leave open questions regarding the scope of the theory. Figure A6 plots the names of countries respondents revealed may be the hypothetical human rights violator in the vignette. Respondents in both conditions mentioned countries that are not allies of the US, with adversaries like Russia and China being the most common responses.

While this finding does not pose a threat to internal validity, it raises questions regarding the scope of the theory. Namely, *will these findings replicate when the stakes of shaming are higher and the country is a US ally?* I address this issue by varying the target of shaming in a second survey experiment. As evident from Fig. 6, the shaming treatment had large and statistically significant effects on government approval and perceptions of domestic respect for human rights across both ally and non-ally conditions. Table B7, which tests the effect of the interaction term of both treatments, reveals that respondents were slightly *more* supportive of the US government when it criticized an ally ($\beta = 0.2,\ p = 0.06$). This effect may be driven by relatively lower baseline support for the US government in the ally condition, where respondents may have punished the US government in the scenario for having allies that violate human rights to begin with.

The effects of the shaming treatment on government approval and perceptions of domestic respect for human rights remained substantively similar to the effects identified in the first experiment. The shaming treatment in the second experiment increased government approval by approximately 1.4 points on a 7-point scale ($p < 0.0001$), equivalent to about 76 per cent of a standard deviation (Tables B7). The shaming treatment increased perceptions of respect for human rights by 1 point on a 7-point scale ($p < 0.0001$), which is equivalent to about 55 per cent of a standard deviation (Tables B8).

However, this additional study reveals mixed findings regarding H3. The effect of shaming on opposition to torture is negative, as in the first experiment, but is only statistically significant when controlling for the target's identity and pre-treatment demographic covariates (see Table B9). Indeed, as is evident from Fig. 6, the moral licensing effect appears to be driven by respondents in the non-ally condition. This finding raises the question, why would shaming of non-allies license problematic behaviour whereas shaming of allies fails to do so? My theoretical framework suggests that shaming should increase tolerance for problematic attitudes when it establishes 'moral credentials' and reinforces beliefs about the morality of one's ingroup (Kouchaki 2011;

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

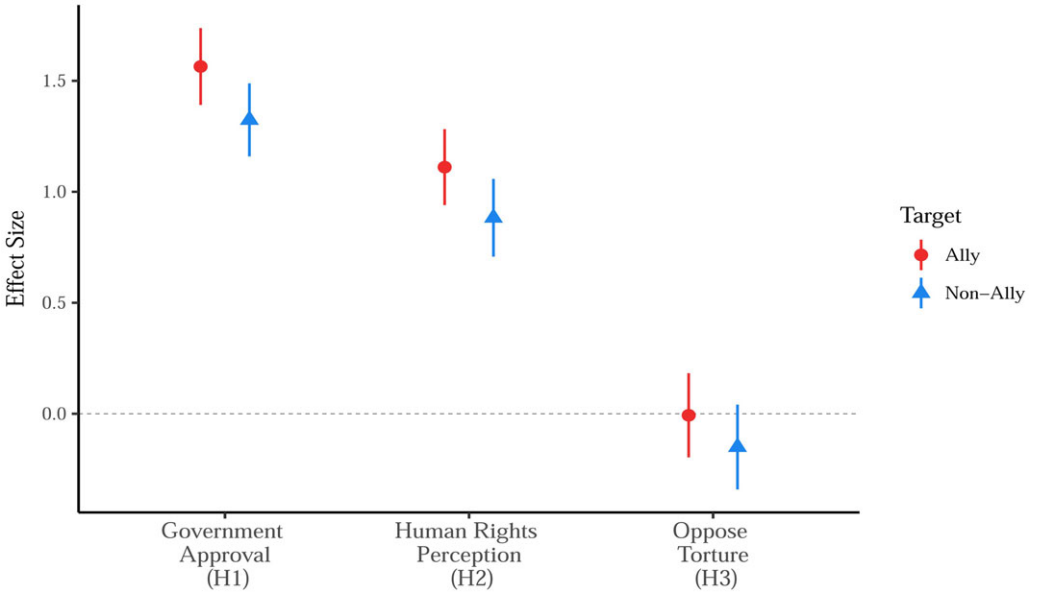

Figure 6. The effect of the shaming treatment by target conditions (Experiment II). The *X*-axis represents three outcomes of interest – government approval, perceptions of respect for human rights, and opposition to the use of torture, ranging between 1–7. The shaming treatment effect for respondents in the *ally* condition corresponds to the red circle. The shaming treatment effect for respondents in the *non-ally* condition corresponds to the blue triangle.

Lasarov and Hoffmann 2020). If shaming decreases opposition to torture by increasing moral credentials, we should expect respondents in the 'non-ally' condition to perceive shaming as more moral than respondents in the 'ally' condition.

Figure B8 strengthens this assumption and demonstrates that non-ally shaming is more likely to establish moral credentials by increasing the belief that the US government is moral. Indeed, respondents perceived the US government as less moral when it shamed an ally than when it shamed a non-ally ($\beta = -0.4,\ p < 0.001$). While I cannot definitively say, this effect may be driven by respondents' disapproval of their government for having allies that violate human rights in the scenario. In the Supplementary Material, I further report results from a mediation analysis where perceptions of morality are the mediator and opposition to torture is the outcome, showing that the average causal mediation effect is negative and statistically significant (see Fig. B10).

Overall, the results from this additional study increase our confidence that shaming human rights violators can shape individual attitudes but also shed light on the limitations of the moral licensing effect – a point that I return to in the concluding section of this paper.

Probing External Validity

While the experimental findings presented in the preceding sections provide causal evidence regarding the ability of shaming to improve government approval and shape human rights attitudes, they face certain limitations relating to external validity. Survey experiments, in particular, face scrutiny regarding the extent to which their findings can be generalized beyond the specific study (Coppock 2019). Egami and Hartman (2023), usefully decompose external validity into four components that relate to populations (X-validity), treatments (T-validity), outcomes (Y-validity), and contexts (C-validity). In this section, I consider the generalizability of my findings along some of these dimensions and discuss supplementary analyses and additional observable implications of my theory to mitigate some of these concerns.

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

A primary external validity concern relates to the extent to which effects identified in one sample generalize to a broader population of interest (Berinsky, Huber, and Lenz 2012). Since my survey experiments relied on US-based online samples, the estimated causal effects may not generalize to other target populations. Although recent studies demonstrate that these concerns may have been overblown (Bassan-Nygate et al. 2024; Coppock, Leeper, and Mullinix 2018), I assess this question in section C.1 of the Supplementary Material by estimating the external robustness of my findings. Specifically, I follow Devaux and Egami (2022), and estimate how different a population should be from my experimental sample to explain away the target population's average treatment effect. I report extremely high estimates of external robustness, ranging from 0.87–1, which are much higher than Devaux and Egami's proposed upper bound benchmark for moderate-high external robustness (0.57). This implies that my estimates are robust to populations that are highly different from my experimental sample.

Another critical concern pertains to what Egami and Hartman (2023), define as T-validity, also known as ecological validity or mundane realism, which encompasses the authenticity of the information presented in the survey experiment and whether respondents engage with it as they would in the real world. Indeed, it is possible that while respondents endorse governments that shame foreign countries in a hypothetical and highly controlled survey environment, the dynamics may differ in the real world, where information on human rights criticism may be less visible in the public domain.

To alleviate some of these concerns and examine whether similar patterns emerge outside the survey experiment, I report supplementary analyses of some observable implications of my theory in section C.3 of the Supplementary Material, utilizing shaming data covering over 70 countries between 2008 and 2020. Specifically, I analyze global trends of shaming and public opinion using cross-national data from the UN Universal Periodic Review (UPR) combined with relevant outcomes collected by the Comparative Study of Electoral Systems (CSES) survey and the World Values Survey (WVS).

I follow Terman and Byun (2022) and estimate government shaming using data from the UPR, a state-driven mechanism of the UN Human Rights Council where governments review and criticize each other's human rights records. I show that survey respondents whose government shames more countries in the UPR are more likely to vote for the incumbent and perceive their country as more respectful of human rights in the following year. This association is weak but statistically significant, even when including a host of country-level and individual-level controls.[1]

This supplementary analysis provides suggestive evidence that the patterns identified in the survey experiment–where domestic audiences reward their governments for engaging in foreign criticism–may persist in the real world. It does, however, have its limitations. Not only does the analysis face obvious endogeneity concerns, but it is also limited in its ability to capture variations in reactions to shaming between individuals within a country and by the target's identity – a point that I return to in the concluding section of this paper. Notably, the supplementary observational analysis also requires strong assumptions about how individuals access public criticism, particularly when using UPR as a proxy for shaming. Indeed, while government activity in the UPR is publicly available and occasionally publicized by human rights NGOs and media outlets (Amnesty International 2022; Gavilan 2022), most voters likely do not follow it closely. I thus interpret UPR shaming as a proxy measuring the extent to which a country, in general, is likely to engage in human rights shaming which may manifest in other channels beyond the UPR like press releases or official speeches.[2]

[1]All models include country and year-fixed effects to account for differences between countries over time. Furthermore, I show in section C.3.4 of the Supplementary Material that my results are relatively robust to several alternative modelling and present results from a placebo test, demonstrating that UPR shaming at time $t+1$ is not associated with my dependent variables at time t.

[2]The small β observed in the supplementary analysis may thus be construed as a lower-bound conservative point estimate, considering that individuals do not observe UPR shaming directly.

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

Still, the observational analysis and my theory more broadly require strong assumptions about the extent to which the public is aware of human rights criticism. To shed light on this assumption, I analyze responses from a follow-up survey (see Supplementary Material C.2), revealing that 86.3 per cent of respondents recalled instances in which US officials criticized a foreign country's human rights policies in the past *month*. This finding aligns with others who have shown that the general public is quite knowledgeable about human rights issues[3] as well as recent public opinion polls that demonstrate that the US public is increasingly supportive of holding human rights violators accountable, even when doing so comes with an economic cost (Silver 2021; YouGov 2023). Hence, while not every individual can be expected to be aware of every instance in which their government criticizes another country, there appears to be suggestive evidence that the public generally pays attention to shaming and values criticism directed toward foreign governments.

To conclude, in this section, I probe the external validity of findings concerning the impact of shaming on government approval and human rights attitudes. While acknowledging the limitations of my experimental results, I discuss results from comprehensive supplementary analyses that reinforce some of the causal evidence. Addressing generalizability concerns, I estimate the external robustness of my analyses and corroborate my experimental insights with cross-national observational data, revealing correlations between government shaming and subsequent public perceptions. I further provide suggestive evidence that my required assumptions regarding the salience and visibility of human rights criticism are sensible.

Conclusion

The primary goal of this paper has been to develop and test a theory that puts human rights shamers front and centre. While past work often makes reasonable assumptions about the domestic consequences of shaming human rights violators abroad, I delineate testable hypotheses that I evaluate in two US-based survey experiments. I argue that shaming is a public event that is assessed by domestic audiences and is rewarded at home. Overall, I find that criticizing other countries' human rights records: (1) increases government support; (2) increases perceptions of domestic respect for human rights; and (3) modestly decreases opposition to torture when shaming shapes perceptions of government morality.

These findings suggest that shaming may be a beneficial tool for governments. Indeed, two different survey experiments in the US demonstrate that information on government shaming has a sizable effect on government approval and favourable perceptions of domestic respect for human rights. These patterns are consistent with suggestive evidence from an observational analysis reported in the Supplementary Material. From a policy perspective, political elites who want to appease domestic audiences can engage in human rights shaming and improve their image back home.

However, what is good news for governments may have negative implications for the human rights movement. By shaming other countries, leaders can shape domestic perceptions of respect for human rights without doing any of the groundwork necessary; neither adopting new policies that improve human rights domestically nor making official commitments to human rights. Shaming thus constitutes 'cheap talk'[4] that increases approval without tying governments' hands by committing to human rights domestically. These implications are even more alarming when considering the potentially negative externalities of shaming other countries toward support for human rights. Indeed, my findings suggest that shaming can have a negative, albeit modest, effect on domestic opposition to torture, which may provide leeway for governments to engage in transgressive behaviour at home.

[3]See: Tomz and Weeks (2020) Appendix VIII.

[4]At least domestically, as shaming may bear international costs by souring existing relationships, as highlighted by Terman and Byun.

https://doi.org/10.1017/S0007123424000267 Published online by Cambridge University Press

Despite these contributions, this article faces some limitations that motivate future research. In general, I provide stronger support for my hypotheses regarding the domestic benefits of shaming, in comparison to my third – 'moral licensing' – hypothesis. Specifically, results from my survey experiments yield large effect sizes for hypotheses 1 and 2, which are complemented with suggestive cross-national evidence. On the other hand, H3 was solely tested in US-based survey experiments and yielded modest effects. Future work should thus strive to probe the external validity of this finding by implementing similar designs beyond the US context.

The empirical analyses of H3 also reveal more nuanced findings, which warrant further research. The second experiment reveals that shaming can license opposition to torture only when the moral image of the US is not tainted by cooperating with human rights violators. This finding highlights the limits of moral licensing, which is only effective when it reinforces beliefs about the government's morality. If constituents hold highly immoral perceptions of their government, shaming may not effectively license support for violations. In addition, shaming did not seem to decrease support for international law more broadly, which raises questions regarding the effectiveness of moral licensing across domains. Future work should thus examine more carefully the scope conditions of the moral licensing finding.

Second, although I provide a theoretical discussion of mechanisms that may drive my identified effects as well as some suggestive empirical evidence, my empirical analysis falls short of providing direct evidence regarding causal mechanisms. As explicated in this paper, these mechanisms may relate to perceptions of morality and power. Future work should adapt rigorous designs that are capable of manipulating not only treatments but also potential mechanisms (Imai et al. 2011) to advance our understanding of the effects of shaming.

Third, in this article, I established that by shaming human rights violators abroad, governments are rewarded at home. However, it's important to note that the dynamics of shaming are intricate, influenced by repeated interactions, and vary significantly across political contexts and targeted countries. Exploratory analyses reported in section A.7 of the Supplementary Material suggest that reactions to shaming may vary as a function of the targeted state. Yet, further research is indispensable to develop comprehensive theories and conduct tests exploring how the impact of shaming varies across these diverse contexts and in relation to individual preferences. A related direction for future research would be to investigate whether shaming in certain contexts generates audience costs, wherein governments face consequences for subsequent policies that contradict the initially publicized shaming, thus considering foreign criticism as a part of a sequence of events.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/S0007123424000267.

Data availability statement. Replication data for this article can be found in Harvard Dataverse at https://doi.org/10.7910/DVN/10ECQZ.

Acknowledgements. I am grateful to Zoltan Buzas, Adeline Lo, Molly Melin, Yoav Nygate, Jon Pevehouse, Jonathan Renshon, Nadav Shelef, Jessica Weeks, and Chagai M. Weiss for their helpful feedback.

Financial support. This work was supported by the Political Science Department at the University of Wisconsin–Madison.

Competing interests. The author declares none.

Ethical statement. All survey experimental procedures received approval from the University of Wisconsin's Institutional Review Board. Informed consent was obtained from all participants before proceeding to the survey.

References

Adler-Nissen R (2014) Stigma management in international relations: Transgressive identities, norms, and order in international society. *International Organization* **68**(1), 143–176.

Allendoerfer MG, Murdie A and Welch RM (2020) The path of the boomerang: Human rights campaigns, third-party pressure, and human rights. *International Studies Quarterly* **64**(1), 111–119.

Amnesty International (2022) Thailand: Mock fashion show protester sentenced to two years for 'insulting the monarchy'. Available from https://www.amnesty.org/en/latest/news/2022/09/thailand-mock-fashion-show-protester-sentenced-to-two-years-for-insulting-the-monarchy/ (accessed 17 March 2024).

Ash K and Dolan TM (2021) Politics of interconfessional empathy and schadenfreude in Lebanon. *Conflict Management and Peace Science* **38**(6), 718–741.

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

Bassan-Nygate L (2025) "Replication Data for: How Does Shaming Human Rights Violators Abroad Shape Attitudes at Home?", https://doi.org/10.7910/DVN/10ECQZ, Harvard Dataverse, V1.

Bassan-Nygate L, Renshon J, Weeks J and Weiss CM (2024) The generalizability of IR experiments beyond the us. *American Political Science Review*.

Bayram AB (2017) Due deference: Cosmopolitan social identity and the psychology of legal obligation in international politics. *International Organization* **71**(S1), S137–S163.

Becker J, Kreps SE, Poast P and Terman R (2024) Transatlantic shakedown: presidential shaming and NATO burden sharing. *Journal of Conflict Resolution* **68**(2-3), 195–229.

Berinsky AJ, Huber GA and Lenz GS (2012) Evaluating online labor markets for experimental research: Amazon. com's Mechanical Turk. *Political Analysis* **20**(3), 351–368.

Bracic A and Murdie A (2020) Human rights abused? Terrorist labeling and individual reactions to call to action. *Political Research Quarterly* **73**(4), 878–892.

Brutger R and Kertzer JD (2018) A dispositional theory of reputation costs. *International Organization* **72**(3), 693–724.

Búzás ZI and Bassan-Nygate L (2024) Race, shaming, and international human rights. *American Journal of Political Science*.

Cascio J and Plant EA (2015) Prospective moral licensing: Does anticipating doing good later allow you to be bad now? *Journal of Experimental Social Psychology* **56**, 110–116.

Coppock A (2019) Generalizing from survey experiments conducted on Mechanical Turk: A replication approach. *Political Science Research and Methods* **7**(3), 613–628.

Coppock A, Leeper TJ and Mullinix KJ (2018) Generalizability of heterogeneous treatment effect estimates across samples. *Proceedings of the National Academy of Sciences* **115**(49), 12441–12446.

Dafoe A, Zhang B and Caughey D (2018) Information equivalence in survey experiments. *Political Analysis* **26**(4), 399–416.

DeMarche E (2021) Biden takes heat for different norms remark while discussing China, human rights. Available from https://www.foxnews.com/politics/biden-takes-heat-for-different-norms-remark-while-discussing-china-human-rights (accessed 17 March 2024).

Devaux M and Egami N (2022) Quantifying robustness to external validity bias. *Available at SSRN* 4213753 (*accessed* 17 *March* 2024).

Effron DA, Cameron JS and Monin B (2009) Endorsing Obama licenses favoring whites. *Journal of Experimental Social Psychology* **45**(3), 590–593.

Egami N and Hartman E (2023) Elements of external validity: Framework, design, and analysis. *American Political Science Review* **117**(3), 1070–1088.

Fearnow B (2021) China blasts us 'human rights violations double standard': 'Robber acting like a cop'. Available from https://www.newsweek.com/china-blasts-us-human-rights-violations-double-standard-robber-acting-like-cop-1590144 (accessed 17 March 2024).

Fearon JD (1994) Domestic political audiences and the escalation of international disputes. *American Political Science Review* **88**(3), 577–592.

Gavilan J (2022) Why the upcoming un review of human rights under Duterte matters. Available from https://www.rappler.com/nation/things-to-know-united-nations-review-philippines-human-rights-record-rodrigo-duterte/ (accessed 17 March 2024).

Greenhill B and Reiter D (2022) Naming and shaming, government messaging, and backlash effects: Experimental evidence from the convention against torture. *Journal of Human Rights* **21**(4), 399–418.

Gruffydd-Jones JJ (2019) Citizens and condemnation: Strategic uses of international human rights pressure in authoritarian states. *Comparative Political Studies* **52**(4), 579–612.

Guisinger A and Saunders EN (2017) Mapping the boundaries of elite cues: How elites shape mass opinion across international issues. *International Studies Quarterly* **61**(2), 425–441.

Haddad B (2021) How Europe became pro-Israel. Available from https://foreignpolicy.com/2021/05/20/how-europe-became-pro-israel/ (accessed 17 March 2024).

Hafner-Burton EM (2008) Sticks and stones: Naming and shaming the human rights enforcement problem. *International organization* **62**(4), 689–716.

Hendrix CS and Wong WH (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.

Holsti OR (2004) *Public Opinion and American Foreign Policy*. Ann Arbor: University of Michigan Press.

Imai K, Keele L and Tingley D (2010) A general approach to causal mediation analysis. *Psychological Methods* **15**(4), 309.

Imai K, Keele L, Tingley D and Yamamoto T (2011) Unpacking the black box of causality: Learning about causal mechanisms from experimental and observational studies. *American Political Science Review* **105**(4), 765–789.
Keck ME and Sikkink K (1999) Transnational advocacy networks in international and regional politics. *International Social Science Journal* **51**(159), 89–101.
Kertzer JD (2017) Microfoundations in international relations. *Conflict Management and Peace Science* **34**(1), 81–97.
Kertzer JD and Brutger R (2016) Decomposing audience costs: Bringing the audience back into audience cost theory. *American Journal of Political Science* **60**(1), 234–249.
Koliev F, Page D and Tallberg J (2022) The domestic impact of international shaming: Evidence from climate change and human rights. *Public Opinion Quarterly* **86**(3), 748–761.
Kouchaki M (2011) Vicarious moral licensing: the influence of others' past moral actions on moral behavior. *Journal of Personality and Social Psychology* **101**(4), 702.
Krain M (2012) J'accuse! does naming and shaming perpetrators reduce the severity of genocides or politicides? *International Studies Quarterly* **56**(3), 574–589.
Krasner SD (1999) *Sovereignty: Organized Hypocrisy*. Princeton: Princeton University Press.
Lasarov W and Hoffmann S (2020) Social moral licensing. *Journal of Business Ethics* **165**(1), 45–66.
Lazaroff T (2021) 157 nations call on Israel to halt the exploitation of Palestinian resources. Available from https://www.jpost.com/arab-israeli-conflict/157-nations-call-on-israel-to-halt-exploitation-of-palestinian-resources-685487(accessed 17 March 2024).
Lazaroff T (2022) Israel joins 50 nations in condemning China's imprisonment of Uyghurs. Available from https://www.jpost.com/israel-elections/article-721178 (accessed 17 March 2024).
Lederer EM (2022) UN demands Russia reverse illegal annexations in Ukraine. Available from https://apnews.com/article/russia-ukraine-middle-east-syria-united-nations-fd7ede16dec560ebb49f9465644de1b1 (accessed 17 March 2024).
Mazar N and Zhong C-B (2010) Do green products make us better people? *Psychological Science* **21**(4), 494–8.
McCurry J (2021) US and China publicly rebuke each other in first major talks of Biden era. Available from https://www.theguardian.com/world/2021/mar/19/us-china-talks-alaska-biden-blinken-sullivan-wang (accessed 17 March 2024).
McFarland S and Mathews M (2005) Who cares about human rights? *Political Psychology* **26**(3), 365–385.
Merritt AC, Effron DA and Monin B (2010) Moral self-licensing: When being good frees us to be bad. *Social and Personality Psychology Compass* **4**(5), 344–357.
Moritsugu K (2021) China bashes us over racism, inequality, and pandemic response. Available from https://abcnews.go.com/US/wireStory/china-bashes-us-racism-inequality-pandemic-response-76648636 (accessed 17 March 2024).
Murdie AM and Davis DR (2012) Shaming and blaming: Using events data to assess the impact of human rights INGOS. *International Studies Quarterly* **56**(1), 1–16.
Myrick R and Weinstein JM (2022) Making sense of human rights diplomacy: Evidence from a US campaign to free political prisoners. *International Organization* **76**(2), 379–413.
Neier A (2018) "Naming and shaming": Still the human rights movement's best weapon. Available from https://www.openglobalrights.org/Naming-and-shaming-still-the-human-rights-movements-best-weapon/?lang = English (accessed 17 March 2024).
Powers R and Renshon J (2023) International status concerns and domestic support for political leaders. *American Journal of Political Science* **67**(3), 732–747.
Putnam RD (1988) Diplomacy and domestic politics: The logic of two-level games. *International Organization* **42**(3), 427–460.
Rathbun BC, Kertzer JD, Reifler J, Goren P and Scotto TJ (2016) Taking foreign policy personally: Personal values and foreign policy attitudes. *International Studies Quarterly* **60**(1), 124–137.
Risse T, Ropp SC and Sikkink K (1999) *The Power of Human Rights: International Norms and Domestic Change*, **volume 66**. Cambridge, UK: Cambridge University Press.
Silver L (2021) Pressing China on human rights – even if it hurts economic relations – has Americans' bipartisan support. Available from https://www.pewresearch.org/short-reads/2021/04/06/pressing-china-on-human-rights-even-if-it-hurts-economic-relations-has-americans-bipartisan-support/ (accessed 17 March 2024).
Simbrunner P and Schlegelmilch BB (2017) Moral licensing: A culture-moderated meta-analysis. *Management Review Quarterly* **67**(4), 201–225.
Simmons BA (2009) *Mobilizing for Human Rights: International law in Domestic Politics*. Cambridge, UK: Cambridge University Press.
Snyder J (2020) Backlash against naming and shaming: The politics of status and emotion. *The British Journal of Politics and International Relations* **22**(4), 644–653.
Tajfel H (2010) *Social Identity and Intergroup Relations*, **volume 7**. Cambridge, UK: Cambridge University Press.
Terman R (2017) Rewarding resistance: Theorizing defiance to international norms. Unpublished Manuscript.
Terman R and Byun J (2022) Punishment and politicization in the international human rights regime. *American Political Science Review* **116**(2), 385–402.
Terman R and 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 and Tomz M (2022) The effects of naming and shaming on public support for compliance with international agreements: an experimental analysis of the Paris Agreement. *International Organization* **76**(2), 445–468.
Tomz M (2007) Domestic audience costs in international relations: An experimental approach. *International Organization* **61**(4), 821–840.
Tomz MR and Weeks JL (2020) Human rights and public support for war. *The Journal of Politics* **82**(1), 182–194.
Urban J, Bahník Š and Kohlová MB (2019) Green consumption does not make people cheat: Three attempts to replicate moral licensing effect due to pro-environmental behavior. *Journal of Environmental Psychology* **63**, 139–147.
US State Department (2020) 2020 country reports on human rights practices. Available from https://www.state.gov/reports/2020-country-reports-on-human-rights-practices/ (accessed 17 March 2024).
US State Department (2022) Joint statement on resumption of hostilities in northern Ethiopia. Available from https://www.state.gov/joint-statement-on-resumption-of-hostilities-in-northern-ethiopia/ (accessed 17 March 2024).
von Grundherr M, Jauernig J and Uhl M (2021) To condemn is not to punish: An experiment on hypocrisy. *Games* **12**(2), 38.
Wallace GP (2013) International law and public attitudes toward torture: An experimental study. *International Organization* **67**(1), 105–140.
Wootson Jr CR (2021) Biden: U.S. leadership on human rights depends on our record at home. Available from https://www.washingtonpost.com/politics/biden-human-rights/2021/10/15/39b03048-2dc3-11ec-baf4-d7a4e075eb90_story.html (accessed 17 March 2024).
YouGov (2023) Foreign relations – human rights. Available from https://today.yougov.com/topics/society/trackers/foreign-relations-human-rights (accessed 17 March 2024).

Cite this article: Bassan-Nygate L (2025) How Does Shaming Human Rights Violators Abroad Shape Attitudes at Home? *British Journal of Political Science*. https://doi.org/10.1017/S0007123424000267