Does Issue Domain Matter? A Reassessment of Interstate Shaming at the UN Human Rights Council

Jiwon Kim[*]

- Please do not cite or circulate -

Abstract

Why do some human rights issues receive more attention at the UN than others? This study posits the importance of states' heterogeneous issue preferences in understanding their shaming behavior. First, allowing sender effects to vary across issues reveals meaningful heterogeneity in states' issue emphases after accounting for affinity and network structure. Second, clustering the posterior distribution of sender–issue effects uncovers normative communities—distinct but shallow patterns of residual issue prioritization that cut across geopolitical alignments. Third, small countries rise as linchpins of these normative clusters. These findings suggest issue preferences contribute a secondary but structured layer of variation that does not conform to a universal issue hierarchy.

Short Title: Norm Clusters at the UNHRC

Keywords: naming and shaming; human rights; Universal Periodic Review; text analysis

Funding Declaration: There is no funding support to declare.

---

[*]Ph.D. Candidate at Department of Political Science, Emory University, 327 Tarbutton Hall, 1555 Dickey Drive, Atlanta, Georgia, USA 30322, Email: jiwon.kim@emory.edu

1 Introduction

Why do some human rights issues receive greater attention at the United Nations than others, and what does this reveal about the politics of international accountability? A prominent recent argument holds that patterns of issue emphasis reflect a strategic calculation: shaming on some issues is politically costlier and more regime-threatening for the target, prompting states to reserve these issues for adversaries while deploying safer topics when criticizing allies (Terman and Byun 2022). Notably, this account rests on the theory that both states' issue preferences and targets' issue sensitivities follow a uniform pattern across countries.

Yet the empirical design underlying this argument makes it difficult to evaluate whether such cross-issue patterns hold. Estimating separate regressions for each issue area prevents direct comparison of issue effects and treats variation across issues as exogenous to the model. Moreover, conventional dyad–year fixed-effects OLS specifications remove the very features existing scholarship suggests are potentially meaningful for explaining interstate shaming behavior. For example, existing literature views international norms as contested, evolving, and subject to selective appropriation (Dixon 2017; Dukalskis 2023; Meyerrose and Nooruddin 2025). If states actively shape human rights norms and prioritize different topics (Glendon 1997, 1156, Jensen 2016), then their issue preferences in shaming should also diverge in systematic ways (Terman and Búzás 2021). If shaming behavior is shaped by these sources of dependence, fixed-effects OLS is likely to obscure, rather than reveal, systematic cross-issue differences in shaming decisions.

This study reassesses the role of issue domain in interstate shaming at the Universal Periodic Review (UPR) by moving beyond issue-by-issue regressions toward a unified hierarchical pooled additive and multiplicative effects (AME) model.[1] This framework explicitly incorporates sender heterogeneity, target heterogeneity, and latent dyadic dependence, providing a clearer test of whether political affinity shapes shaming differently across issues and whether states exhibit meaningful variation in their issue priorities. The pooled specification also allows for direct comparison of issue effects—central to evaluating theories that hinge on cross-issue differences.

---

1. Modified from (Hoff 2015).

The analysis proceeds in three steps. First, I revisit the core claim that states systematically shame adversaries more than allies, and assess whether that relationship holds across all issue areas. In doing so, I closely follow the empirical strategy of Terman and Byun (2022) using their issue-by-issue regression design, and modify the model specification into a pooled version, which allows direct comparison of issue effects. The analysis produces results that diverge from the theory's predicted cross-issue pattern: instead of selectively shaming adversaries only on *costly* issues and allies on *safer* ones, the pooled model reveals limited differentiation between issue domains.

Second, I test whether states exhibit heterogeneous preferences over issue domains once political alignment, target characteristics, and network dependencies are accounted for. The hierarchical specification allows sender–issue random effects to vary across all states and issues simultaneously. These random effects show substantial variance, and model comparison using PSIS-LOO and WAIC strongly favors the full hierarchical specification over a model with sender effects that do not vary across issues. This provides direct evidence for Hypothesis 2: that states differ systematically in the issues they emphasize when shaming.

Third, I explore whether these heterogeneous issue emphases reveal underlying normative structure. Using clustering methods applied to the posterior distribution of sender–issue random effects, I detect both global clusters of generalist over- and undershamers and local communities of states that share distinctive conditional issue priorities even after removing political affinity. These communities capture patterned variation not reducible to geopolitical alignment and provide evidence for Hypothesis 3: that issue preferences cluster into coherent normative blocs.

Taken together, the findings challenge the assumption that issue sensitivity and issue preference are uniform across states. Instead, the evidence points to a system where political alignment shapes shaming in a largely issue-consistent manner, but states nevertheless differ in how they distribute attention across issue domains. These differences are coarse at the global level but reveal meaningful normative cleavages when examined in relative terms. This reorients the study of international shaming away from a strict cost-based logic and toward an understanding of how heterogeneous normative agendas shape the politics of accountability.

2 Hypotheses

Scholars consistently show that shaming is a political exercise. Across multiple arenas, including the UN Human Rights Council's Universal Periodic Review (UPR), regional intergovernmental organizations, and voting in the UNHRC, states are far less likely to criticize allies or politically aligned partners (Lebovic and Voeten 2006; Donno 2010; Hug and Lukács 2014; Terman and Voeten 2018; Park and Park 2025). In the analyses, the role of the issue domain had been muted as a control variable.

However, recent work offers a first systematic account about the role of issue domain in explaining interstate shaming behavior. To address the tension between political selectivity and the widely shared desire of states to present themselves as impartial defenders of human rights, Terman and Byun (2022) advances a compelling theory: states tend to direct regime-threatening criticisms at adversaries while limiting their criticism of allies to domains perceived as less sensitive. According to this account, shaming is most consequential when it challenges the target regime's legitimacy, particularly by highlighting highly stigmatized violations of physical integrity rights (4). It is also sensitive when it involves compliance with costly norms, such as civil and political rights related to free expression and political participation. Migrant rights, in the context of rising anti-immigrant sentiment, have similarly become a politically and financially burdensome issue (5). By contrast, they suggest that governance and public services related issues, socioeconomic rights, women's rights, and children's rights are generally safer domains for criticism, as they are less directly threatening to regime stability (5). Other areas they classify as less regime-threatening include rights related to governance and the justice system (Figure 1). Their theory does not classify Race and Ethnic Discrimination into either category, although they present these results (11).[2]

Hypothesis 1 (Replication). *States are more likely to shame adversaries on regime-threatening issues— civil-political rights, physical integrity rights, migration— and friends on soft issues—governance and public services, socioeconomic rights, women's rights, and children's rights.*

---

2. They do not form expectations about LGBTQ rights but include it as a subset of the Protection of Vulnerable Populations cluster.

Importantly, this theory rests on two key notions. First, it suggests that states share a relatively uniform set of issue preferences as senders: holding constant their relationship with a potential target, they are assumed to be similarly willing to enforce norms around protecting LGBT populations and ensuring economic rights. Second, it implies that the target states also share homogeneous issue preferences across issues. That is, what is *sensitive* or *soft issue* to be shamed at an international peer-review session is constant across countries.[3]

While these assumptions help the theory establish a coherent baseline, the framework leaves open important questions about whether such uniformity of issue sensitivity and preference can be taken for granted across diverse international contexts. Human rights may represent a "new standard of civilization" (Donnelly 1988), but this standard is itself shaped through the practices of states as members of the international system. Like any norms, human rights are dynamic social products, actively made and contested (Finnemore and Sikkink 1998; Jensen 2016). States use speech, voting behavior, and coalition-building at the United Nations to influence which rights are prioritized (Dixon 2017; Dukalskis 2023; Meyerrose and Nooruddin 2025).[4] States promote new standards or resist existing ones across diverse domains, from freedom of religion (Baderin 2003) and torture (McKeown 2009) to online freedoms (R. Hulvey 2023) and socioeconomic rights (Alston 1993; Arat 1999; Sceats and Breslin 2012; Berger 2023). Consistent with this view, descriptive evidence from the UPR shows that states on average focus on different issues, though without fully theorizing how these issue preferences interact with geopolitical alignment (Terman and Búzás 2021).

This paper therefore challenges the common assumption that states have uniform preferences over which human rights issues they emphasize when shaming adversaries or refraining from shaming allies. The analyses that follow are intended to be theory-generating rather than definitive tests of a fully specified logic. Drawing on a scattered but suggestive body of issue-specific and country-level qualitative studies, I examine whether states display heterogeneous issue preferences in their shaming behavior after accounting for political affinity and other structural factors. Explaining the

---

3. A country may have different preferences as to which issue they advocate outwardly vs. which issue they are shamed inwardly.

4. At the public opinion level, shaming in a given issue area increases support for the international norm (Bassan-Nygate 2025).

origins of this heterogeneity falls outside the scope of this paper, but existing case studies and issue-focused research indicate that political context, regime type, and evolving international norms can meaningfully shape states' priorities and rhetorical choices in global human rights discourse (Alston 1993; Baderin 2003; McKeown 2009; Donno, Fox, and Kaasik 2022; Berger 2023).

Hypothesis 2 (Corrective)**.** *Controlling for affinity, states exhibit heterogeneous preferences over which issue areas they emphasize in shaming.*

If issue preferences are heterogeneous, how is that variation structured? Prior work shows that international institutions are not merely arenas for material bargaining but also sites where states align into communities of practice, reinforcing particular human rights commitments (Terman and Búzás 2021). Building on this insight, I expect that states will cluster into meaningful communities with distinct normative profiles that cut across geopolitical affinity.

Hypothesis 3 (Exploratory)**.** *Issue preferences cluster into communities not reducible to material alignments.*

3   Data

To evaluate the hypotheses, I use data from the Universal Periodic Review (UPR), compiled by UPR Info and widely employed in research on international human rights politics (Terman and Voeten 2018; Terman and Búzás 2021; Terman and Byun 2022; Kim 2023; Meyerrose and Nooruddin 2025). The UPR is a peer-review process in which every UN member state's human rights record is evaluated by all other members on a rotating four and a half year cycle. Prior to each review, the state under examination and relevant stakeholders including regional organizations and civil society groups submit reports that are compiled by the Office of the High Commissioner for Human Rights (OHCHR). During the review session, member states issue recommendations identifying areas for improvement. A state may issue none, one, or several recommendations, and the government under review must formally respond by either supporting or "noting" (rejecting) each recommendation. Following the

session, a report summarizing the dialogue and recommendations is published, and states may optionally submit a mid-term update before their next review.

The dataset covers the first two full review cycles (2008–2016). It includes 54 issue labels hand-coded by UPR Info experts, which are grouped into seven thematic clusters by Terman and Byun (2022): Civil-Political Rights, Governance and Public Service, Migration, Physical Integrity Rights, Race and Religious Discrimination, Socio-Economic Rights, and Vulnerable Populations.[5] Following Terman and Byun (2022), I exclude sender–target dyads in which no recommendations were made in either cycle, in order to focus only on cases where shaming could occur.

4 Findings

4.1 Hypothesis 1

Fixed effects model   To assess whether states are more likely to shame adversaries than allies across different human rights issue domains, I closely follow the approach developed in the existing study, with one modification in model specification. Using the same data, I include all observations regardless of the issue label of the shaming recommendation and estimate a single ordinary least squares (OLS) regression model. This model adds fixed effects for issue while retaining all other control variables from the replicated study, including geopolitical affinity between reviewer and target (Voeten, Strezhnev, and Bailey 2009), the shamer's and target's physical integrity rights protections (Fariss, Kenwick, and Reuning 2020), shared region, and fixed effects for sender, target, and year. Estimating one OLS model, rather than seven separate models for each thematic cluster, allows the analysis to account for systematic differences across issue domains while producing directly comparable coefficient estimates.

Figure 1 presents the β coefficients capturing the marginal effects of political affinity on shaming across issue areas. Hypothesis 1 predicts negative effects for civil–political rights, physical integrity rights, and migration, and positive effects for governance and public services, socio-economic rights,

---

5. Appendix A, Table A.1 lists the specific issue labels included within each thematic cluster.

6

and the protection of vulnerable populations.[6] The results provide little support for Hypothesis 1. For three of the six issue areas with directional expectations the estimated effects are statistically indistinguishable from zero. For one issue area, the estimated effect runs counter to theoretical expectations. Full regression estimates appear in Appendix Table B.1.

These patterns sharply diverge from earlier findings. Prior work suggests that allies are more frequently criticized over the treatment of vulnerable populations; here, however, the effect is negative and statistically significant ($\hat{\beta} = -0.018, p = 0.029$). Likewise, the expectation that states shame adversaries more over civil and political rights is not supported ($\hat{\beta} = -0.001, p = 0.895$). Similarly, the results diverge from the theory that allies are more frequently targeted than adversaries on governance- and justice-related issues or socio-economic rights ($\hat{\beta} = 0.008, p = 0.675$ and $\hat{\beta} = -0.018, p = 0.189$).
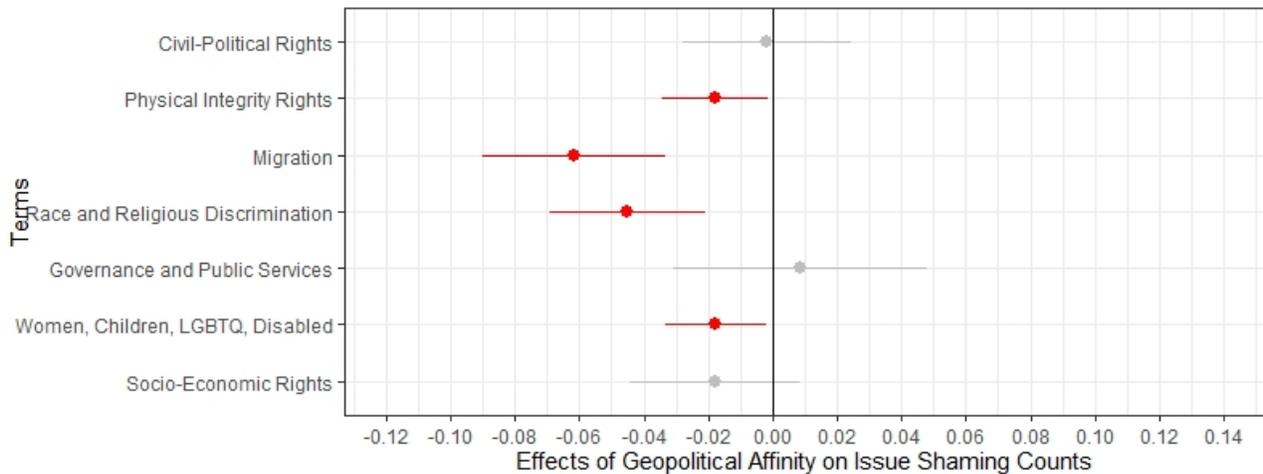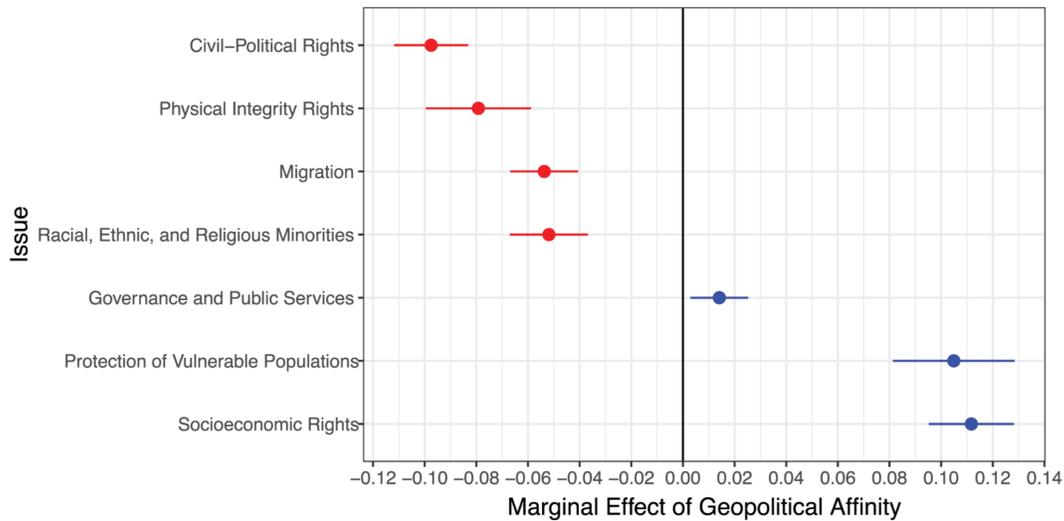

Figure 1: Effects of Geopolitical Affinity on Issue Shaming Counts

Note: The estimates of the β coefficients show that the marginal effect of geopolitical affinity on the number of recommendations offered by a given reviewer for each issue category, controlling for the same factors as the original study, is mostly negative and statistically significant. Some issue areas, such as Governance and Public Service have non-negative estimates, but the standard error is very large.

The difference between the original and the replicated studies, despite their reliance on the same data and OLS regression as an estimation strategy, arises from differences in the model specification. Figure 2, taken from the original study, reports estimates and their margins of error (represented by red, grey, and blue circles and bars) to show whether states differentiate between adversaries and

---

6. The authors do not advance a prediction for Race and Religious Discrimination.

Figure 2: Main Figure from the Previous Study

Note: Reported figure from the original paper, based on seven OLS models estimated separately for each of the seven thematic issue clusters. I present this figure rather than the one based on 54 OLS models (one per issue label), as the latter is too large to display and the seven-model specification enables an easier comparison with my corrected estimates in Figure 1 and Figure 3.

allies in their shaming behavior *when the data are subset by issue category*. However, because these estimates are generated from 7 separate OLS models, the estimates for each issue are not directly comparable.[7]

Random effects model   The above replication with modification provides estimates that are comparable across issues. However, the inherent limitations of having sender and target fixed effects in the model limit our ability to account for an important dynamic accounted in the literature: that senders prioritize different human rights norms, and these priorities may shape their shaming behavior. Numerous studies indicate that states interpret human rights differently and actively work to make and shape international human rights norms in line with their values (Finnemore and Sikkink 1998; Jensen 2016; Cordell et al. 2020; R. A. Hulvey 2022; Dukalskis 2023). In the context of the Universal Peri-

---

7. In other words, we cannot evaluate whether the same actors face different incentive structures across issue domains, because the cross-issue differences are absorbed into issue-specific baselines—comparing 7 separate OLS models is equivalent to one giant fully saturated model interacting the issue variable with all covariates, including fixed effects.

odic Review, where states either endorse or undermine norms by choosing to shame or remain silent, descriptive evidence shows that countries are *fragmented* across issues, holding divergent normative positions (Terman and Búzás 2021). For example, while women's and children's rights are broadly endorsed, norms related to sexuality and migration tend to be championed by only a subset of states. These findings suggest that shaming frequency in a given issue area may be correlated across dyads when they share the same sender. In the same vein, states as targets may exhibit different sensitivities to shaming in different issues. That is, shaming frequency in the dyads that share the same target may be correlated.[8]

Recent research also reports network propensities: due to the UPR's institutional design and the role of NGOs in information sharing and diplomatic networking, states are more likely to shame the potential target that other third-party states shame (Kim 2023). When analyzing such shaming data using an OLS regression model with fixed effects for sender, target, and year, these valuable information— sender heterogeneity, target heterogeneity, and third-order network effects—is lost in the analysis.

To address these while providing robust evidence for Hypothesis 1, I estimate a pooled random effects model that allows for sender-issue random effects, target-issue random effects, and a dyad-level latent variable, adopted and modified from Hoff (2015). The seven clusters of issue areas, included covariates, and all other measurements follow that of Terman and Byun's (2022). While Hoff (2015) is a model for a single network, Equation 1 adaptation models multiple networks (one for each issue domain) simultaneously with issue-specific parameters. As an outcome layer, I use Negative Binomial to account for the asymmetric, ordinal, and overdispersed nature of our shaming count data. The AME structure proposed by Hoff (2015) is highly flexible and has been adapted across a wide range of applications (e.g., Gade, Hafez, and Gabbay 2019; Gade et al. 2019; Rozenas, Minhas, and Ahlquist 2019; Adams, Weschle, and Wlezien 2021; Kinne 2024). Their ordinal extension example in the paper employs a semi-parametric outcome layer treated as a statistical nuisance, "for modeling of more general types of ordinal variables, such as those that are continuous, or those for which the number of levels is not pre-specied," which the authors describe as "somewhat limiting" (Hoff

---

8. Overall dyad-level shaming frequency (regardless of the issue domain) is already accounted for as a control variable.

2015, 27). Equation 1 modifies this approach for our purposes by fully parameterizing the outcome distribution to handle overdispersed count data directly. The tradeoff is computational. The fully Bayesian specification is expensive: using the No-U-Turn Sampler (NUTS), a type of Markov Chain Monte Carlo (MCMC) algorithm, the estimation took approximately 13 hours on an NVIDIA A100 GPU.

$$y_{ijk} \sim \text{NegativeBinomial}(\mu_{ijk}, \phi),$$

$$\log \mu_{ijk} = \eta_{ijk},$$

$$\eta_{ijk} = \alpha_k + \beta_{\text{aff},k} \, \text{Affinity}_{ij} + \boldsymbol{\beta}_k^\top \boldsymbol{X}_{ij} + \boldsymbol{u}_i^\top \boldsymbol{v}_j + a_{ik} + b_{jk}, \tag{1}$$

$$\text{where } a_{ik} \sim \mathcal{N}(0, \sigma_a^2), \quad b_{jk} \sim \mathcal{N}(0, \sigma_b^2), \boldsymbol{u}_i, \boldsymbol{v}_j \overset{i.i.d.}{\sim} \mathcal{N}_2(\boldsymbol{0}, I_2),$$

$$\alpha_k \sim \mathcal{N}(0, 1), \quad \boldsymbol{\beta}_k \sim \mathcal{N}(\boldsymbol{0}, I), \quad \phi \sim \text{Gamma}(2, 0.1).$$

The results, shown in Figure 3, also provide limited support for Hypothesis 1. Of the six issue domains with directional expectations, three yield effects that run counter to the predicted signs. The general pattern of negative relationships between geopolitical affinity and shaming, however, is consistent with prior evidence that states tend to be more lenient toward those with whom they share closer political ties (Lebovic and Voeten 2006; Donno 2010; Hug and Lukács 2014; Terman and Voeten 2018).

4.2   Hypothesis 2

Up to this point, I have shown that Hypothesis 1 receives only weak support under both a corrected OLS specification with sender and target fixed effects and an alternative random-effects formulation. To investigate more deeply how issue characteristics shape interstate shaming behavior at the UN Human Rights Council, I now turn to the question of whether sender–issue heterogeneity exists once we account for the set of targets a state is simultaneously choosing to shame.

Prior research has described the relationship between political costs and normative commitments in shaming as "difficult to disentangle" (McCarty 2011, 43; Terman and Búzás 2021, 4). The previous
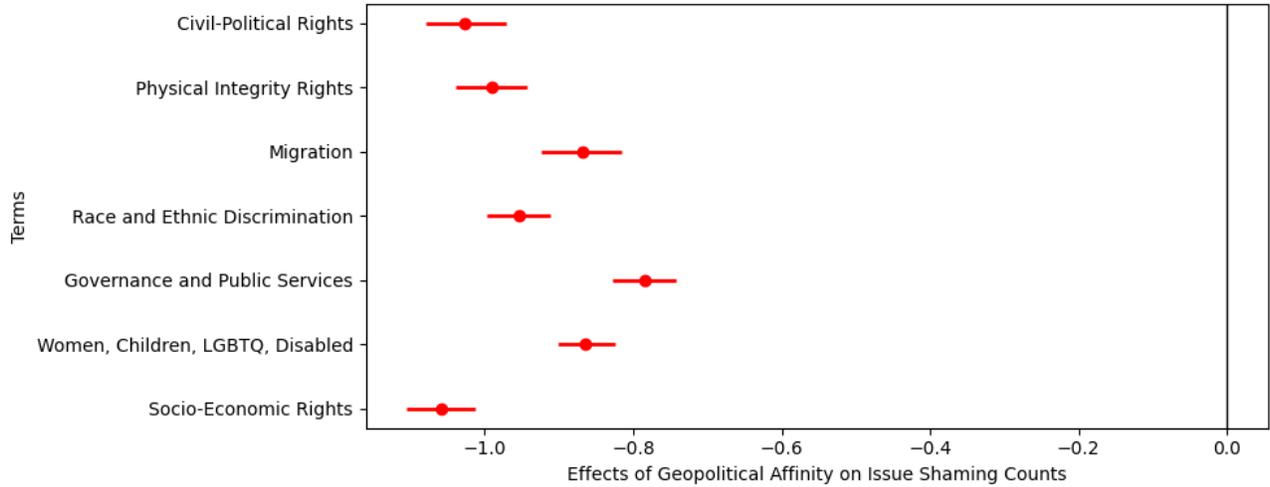
Figure 3: Effects of Geopolitical Affinity on Issue Shaming Counts

Note: The points are the posterior means of $\hat{\beta}_{\text{aff},k}$, representing the marginal effect of geopolitical affinity on the log expected shaming counts between the given dyad for each issue category, controlling for the same factors as the original study and additional random effects and the latent dyadic variable. Credible intervals are constructed at the 95% using the posterior data.

section's pooled-AME model (Equation 1) address this directly, by isolating issue emphasis specific to the sender ($a_{ik} \sim \mathcal{N}(0, \sigma_a^2)$) from other predictors. The variance of this sender-issue random effects term serve as the descriptive evidence for the heterogeneity. The posterior draws show that σ_a has a posterior mean of 0.55 with the 95% credibility interval of [0.52, 0.59], consistent with substantive and statistically meaningful heterogeneity in how states allocate shaming across issue domains.

If sender-issue heterogeneity is sufficiently strong, the pooled AME model in Equation 1 should fit the data better than a nested model that replaces ($a_{ik} \sim \mathcal{N}(0, \sigma_a^2)$) with a simpler sender-level effect that does not vary across issues. Comparing these two models therefore provides a direct assessment of how much issue-specific variation contributes to explaining interstate shaming patterns.

Methods As outlined in Equation 1, the full model includes sender–issue random effects, target–issue random effects, issue-specific intercepts, and a dyadic latent factor structure. The restricted model retains all of these terms except for the sender–issue random effects, which is replaced with a simpler sender-level effect that does not vary across issues. Because both models are estimated in a fully Bayesian framework using MCMC, classical likelihood-ratio testing is not applicable: their

11

marginal likelihoods are not tractable, and they differ in effective rather than fixed parameter counts. Instead, I compare the two models' performance using various measurements of their predictive accuracy. Specifically, I compare whether the full model's expected log predictive density (elpd) is greater than the restricted model. Elpd measures how well a model predicts new, unseen data by averaging the log predictive density over the data points and the posterior distribution of the model parameters (Vehtari and Ojanen 2012; Vehtari, Gelman, and Gabry 2015). To estimate the metric, I include common estimators that explicitly penalize the number of effective parameters and that indirectly account for model complexity through predictive performance, to measure the same property. Notably, these estimators and the measurement are mostly used in the context of model evaluation and selection rather than strict hypothesis testing.

The first method, Pareto-smoothed importance sampling leave-one-out (LOO) cross-validation estimates a model's out-of-sample predictive accuracy by approximating the leave-one-out log predictive density using importance sampling, with a smoothing step to stabilize high-variance weights (Vehtari, Gelman, and Gabry 2015). Instead of an explicit penalty term based on the number of parameters, the method indirectly accounts for model complexity through predictive performance: models that overfit (i.e., are too complex) tend to perform worse on held-out data, resulting in lower PSIS-LOO scores (Vehtari, Gelman, and Gabry 2015). The second method, the Widely Applicable Information Criterion (WAIC) is an information criteria analogous to AIC and BIC in the frequentist space. WAIC estimates out-of-sample predictive accuracy by averaging the log-likelihood over the posterior, penalized by the variance of the log-likelihood, which serves as an effective number of parameters (Watanabe 2009).

Again, a substantial increase in elpd for the full model over the restricted model therefore signals that allowing sender effects to vary across issue domains captures real structure in the data rather than idiosyncratic noise. This provides a principled way to determine whether the richer hierarchical specification is warranted and whether sender–issue heterogeneity meaningfully improves model performance beyond what can be achieved with only target–issue heterogeneity and dyadic latent factors.

	LOO			WAIC		
Model	**elpd**	**Δelpd**	**Weight**	**elpd**	**Δelpd**	**Weight**
Full	−80857.45	0	0.975	−80846.42	0	0.976
Restricted	−83321.37	2463.92	0.025	−83313.65	2467.23	0.024

Table 1: Model comparison using PSIS-LOO and WAIC

Note: Higher elpd values indicate better predictive performance. The full model is overwhelmingly preferred under both criteria. The standard deviation of the pointwise log predictive densities across all data points, divided by the square root of the number of data points

Results   Both the LOO and WAIC strongly favor the full model (see Table 1). Under LOO, the full specification achieves an elpd of –80,857 compared to –83,321 for the restricted model, a difference of roughly 2,464 in favor of the full model.[9] WAIC yields effectively identical results, with a Δ elpd of approximately 2,467. Model weights indicate that the full model receives 97–98% of the predictive support, while the restricted model receives only 2–3%. These differences far exceed conventional thresholds for meaningful improvement, demonstrating that sender–issue heterogeneity substantially enhances predictive accuracy even after accounting for target–issue variability and dyadic latent affinity. Substantively, this provides another evidence that a country's propensity to shame varies across issue areas in systematic ways that cannot be captured by static sender effects, and that modeling this structure directly yields a significantly better representation of shaming behavior.

4.3   Hypothesis 3

If states' shaming preferences diverge, do they nevertheless form meaningful communities based on those preferences? The UPR is designed as an interactive process, where working group sessions, the involvement of regional groups, and the participation of civil society actors provide opportunities for socialization (Kim 2023). To capture this dynamic, I again use the sender-issue level random effects (a_{ik} from Equation 1). The matrix consists of estimates for each sender-issue combination, each estimate with its own confidence intervals. As noted, the raw a_{ik} represents how much the given sender overshames or undershames in absolute terms for this issue, relative to what the model predicts.

---

9. elpd is typically negative because log-likelihoods for probability densities less than 1 are negative, especially when summing over many data points.

The distribution of the posterior means of these values are presented in Figure 5.

In addition to the seven standard issue clusters, I treat LGBT rights as a distinct subcategory, originally nested under the vulnerable populations cluster. Given the rising salience and contestation of LGBT rights in global discourse,[10] this sub-issue merits separate analytical attention (Alter, Hafner-Burton, and Helfer, 2019, 457; Park, Greene, and Colaresi, 2020; Pauselli and Urzúa, 2024; Toosi, 2025;). Within the UPR process, states diverge sharply on whether to raise LGBT rights in shaming. Of the 195 countries that engage in shaming during at least one UPR cycle, 150 never mention the lack of LGBT protections in other countries. Yet, eight states have issued over 200 LGBT-related statements, and seven more have issued over 150, resulting in an average of 88.15 LGBT-related shaming statements (SD = 99.40). The weak correlation between scaled mean residuals for the vulnerable populations cluster and LGBT rights (Pearson's $r = 0.26$) further justifies treating LGBT rights as a separate dimension in the community detection analysis.[11]

Figure 4: Distribution of Shaming Count on LGBT Issue across Countries

Note: Top defenders of LGBT rights at the UPR are Canada, Netherlands, France, Spain, Argentina, Norway, Chile, Czech Republic (> 200), Slovenia, Uruguay, USA, Australia, United Kingdom, Brazil, and Sweden (> 150).

Figure 5 presents a weighted bipartite graph where one set of nodes represents countries and the other set represents issue areas. Edges connect countries to the issues they disproportionately emphasize or de-emphasize, with edge weights corresponding to the magnitude of the posterior mean

---

10. While recognizing disagreement, e.g. Mertus 2007.

11. The UPR data source, UPR Info, applies multi-label classification, assigning multiple issue labels to each statement based on its content.

of the sender-issue random effects. That is, each edge in the network represents the extent to which a country disproportionately emphasizes (blue line) or de-emphasizes (red line) a given issue area in their shaming, after accounting for factors that were included as covariates in the model. Thicker line means a high absolute value of the sender-issue effects. The visualization only depicts a part of the matrix created, as it is set to include edge absolute weights higher than 1.2 for simplicity in visualization that focuses on meaningful variation.[12] In this way, each country's shaming profile is a size 8 vector, while each estimate carries a confidence interval.

---

12. Issue emphasis is interpreted as prioritizing provision of the particular rights, and de-emphasis is interpreted as de-prioritizing and disregarding the provision of the particular rights. An exception to that interpretation is a shaming made in May 2008 in UPR Session 2 during Cycle 1, to Tonga by Bangladesh. Their recommendation is made to undermine LGBT rights: "Continue to criminalize consensual same sex, which is outside the purview of universally accepted human rights norms, according to Tonga's national legislation."
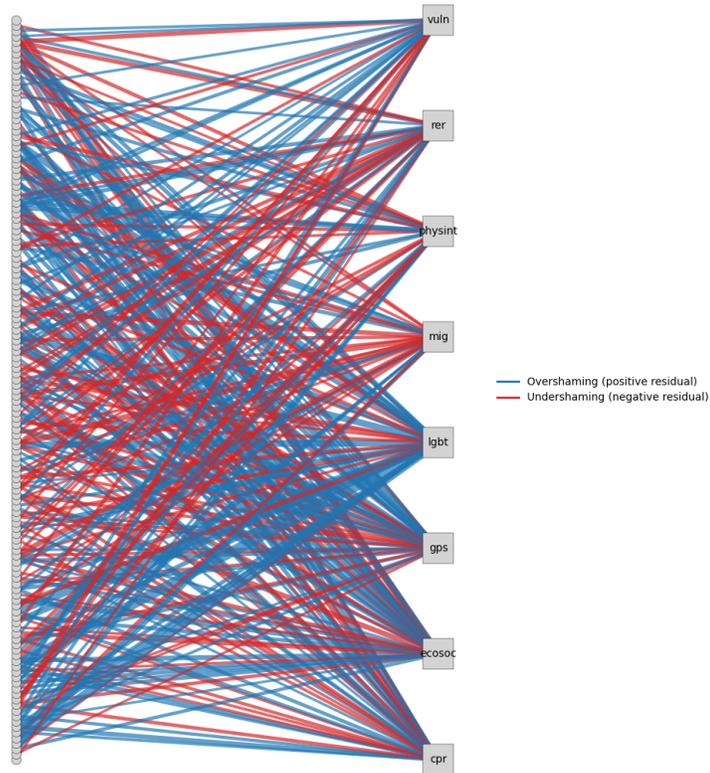
Figure 5: Raw Average Residual Shaming per Country–Issue

Note: Edges represent the posterior means of the demeaned sender-issue random effects ($\tilde{a}_{ik} = a_{ik} - \frac{\sum_k a_{ik}}{K}$) from Equation 1. Blue lines denote over-shaming relative to model expectations, red lines denote under-shaming. Edge thickness reflects the absolute magnitude of the term. To simplify visualization, only edges with a magnitude greater than 1.2 are presented. Credibility intervals are omitted.

Global patterns To assess whether states exhibit systematic cross-issue variation in their shaming emphases net of political alignment and network structure, I analyze the sender–issue random effects (a_{ik}) from Equation 1. Because the goal is to identify global patterns in how states prioritize different issue domains, rather than merely local similarities among small subsets of states, I cluster the full posterior distribution of the a_{ik} matrix rather than relying on a single point estimate. I employ a Gaussian Mixture Model (GMM), a widely used unsupervised clustering method that provides a parametric representation of latent group structure. GMM assumes that observations arise from a mixture of Gaussian components—a defensible approximation once the draws are standardized and aggregated across many posterior samples. Moreover, GMM flexibly accommodates covariance structures and yields probabilistic cluster assignments, allowing uncertainty in each state's cluster membership to be captured directly.

The number of mixture components is chosen by minimizing the Bayesian Information Criterion (BIC) evaluated on the posterior-mean a_{ik} matrix, ensuring an appropriate balance between parsimony and fit. Applying this model to many posterior draws produces a collection of posterior partitions, which I summarize through a posterior co-clustering matrix that records how often each pair of states is assigned to the same cluster. Following Dahl (2006), I then identify the single posterior draw whose partition is closest to this co-clustering matrix and treat it as the consensus global clustering. This procedure recovers a broad, system-wide structure in states' conditional emphasis on human rights issues.[13]

Figure 6 summarizes the average issue profile for the discovered clusters. Three clusters' patterns in states' conditional issue emphases reveal that the variance in shaming sender-issue random effects can largely be understood by systematic undershamers and overshamers. For example, Cluster 0 represents a group of states that adopt a generally low-intensity engagement with the UPR, either due to limited diplomatic capacity, weaker commitments to vertical accountability, or strategic restraint. They issue substantially fewer shaming statements than expected on all issue domains. Cluster 1 consists of states that consistently overshame across all issue domains, with especially strong emphasis

---

13. I compute per-state stability as an average co-clustering with others in the same consensus cluster (Dahl 2006), and the final cluster membership probabilities, which are reported in Appendix NA.
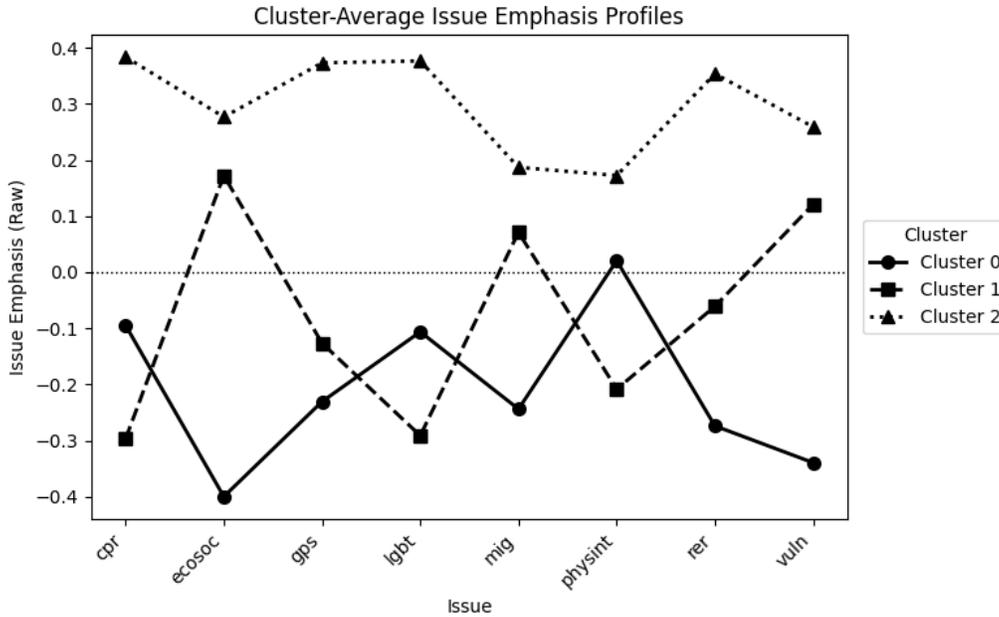
Figure 6: GMM Results

on socioeconomic rights and governance and public services issues. These states behave as general rights entrepreneurs, assigning above-expected attention across a wide spectrum of rights issues. Cluster 2 captures a middle category with a more selective de-emphasis pattern than Cluster 0, where shaming emphasis remains near the zero-line, but undershaming especially in socioeconomic rights and governance and public services issues.

Taken together, these global patterns indicate that, once political affinity and network dependencies are accounted for, the heterogeneity in issue emphasis is relatively coarse. The dominant cleavage is not between states that specialize in particular rights domains but rather among three broad groups: generalist overshamers, generalist undershamers, and a set of states that selectively disengage from socioeconomic and governance-related issues. This suggests that the principal dimension of variation reflects overall shaming activism, while genuine issue specialization plays a secondary role. Substantively, the results imply that most states behave less as active shapers of the international human rights agenda, articulating differentiated normative priorities, and more as takers, varying mainly in their general willingness to engage in shaming rather than in issue-specific strategic or principled commitments.

Local patterns   While these global clusters capture the broad, system-wide structure of states' shaming emphases, they mask more fine-grained relationships among individual states. Many states differ not in how much they shame overall but in the specific issues they prioritize. To uncover this local structure, meaning the patterns of similarity that do not appear in the global partitions, I apply a cosine similarity network approach. This method identifies smaller communities of states whose issue profiles have similar directional patterns, even when their overall levels of shaming place them in different global clusters.

To capture each sender's relative emphasis across issues, I demean the a_{ik} term at the sender level, i.e. $\tilde{a}_{ik} = a_{ik} - \frac{\sum_k a_{ik}}{K}$. Systematic deviations from zero indicate that a shamer over- or under-emphasizes that issue relative to its own baseline. By removing each sender's average level of shaming, the transformed matrix highlights how states distribute their attention across issues relative to their own baselines. Then, I compute cosine similarity between all pairs of countries. Cosine similarity captures the extent to which two issue profiles point in the same angle, and not the centering, in the eight-dimensional space. In other words, this metric reflects the relative pattern of emphasis and de-emphasis across issues, independent of overall levels of shaming of that sender. Table 2 illustrates this approach by displaying the $\tilde{a}_{ik}$ for the United States alongside those of the countries most similar and most dissimilar to its shaming profile in terms of cosine similarity.

Issue Area	United States	Sweden	Luxembourg
Civil-Political Rights	0.592	0.219	-0.232
Socioeconomic Rights	-1.239	0.121	-0.477
Governance and Public Services	0.293	0.121	-0.015
LGBT Rights	1.832	0.680	-0.851
Migration	-0.069	-0.074	0.095
Physical Integrity Rights	-0.319	-0.101	-0.094
Racial, Ethnic, and Religious Minorities	-0.505	-0.145	-0.034
Vulnerable Groups	-0.586	-0.222	0.221

Table 2: Standardized issue emphasis profiles for selected countries

Note: Sweden is the closest to the United States in terms of cosine similarity (0.997) in the 8-dimensional issue preference space. Luxembourg ranks as the most distant by cosine similarity (-0.918).

The resulting cosine similarities between countries are projected onto a one-mode network, as the weight of the edges connecting pairs of countries (see Figure 7). Two countries are linked more

strongly the more similar their issue profiles are. To identify groups of countries that emphasize similar types of issues in their shaming behavior, I apply the Louvain community detection algorithm (Newman, 2004; Blondel et al., 2008). Modularity-based, unsupervised method such as Louvain is an appropriate choice, because I don't have to prespecify the number of communities or the distance metrics arbitrarily. Also, my interest is not in the absolute magnitude but in the relational structure that optimizes for within-group cohesion and between-group separation. Validation of the community detection is provided in Appendix NA.
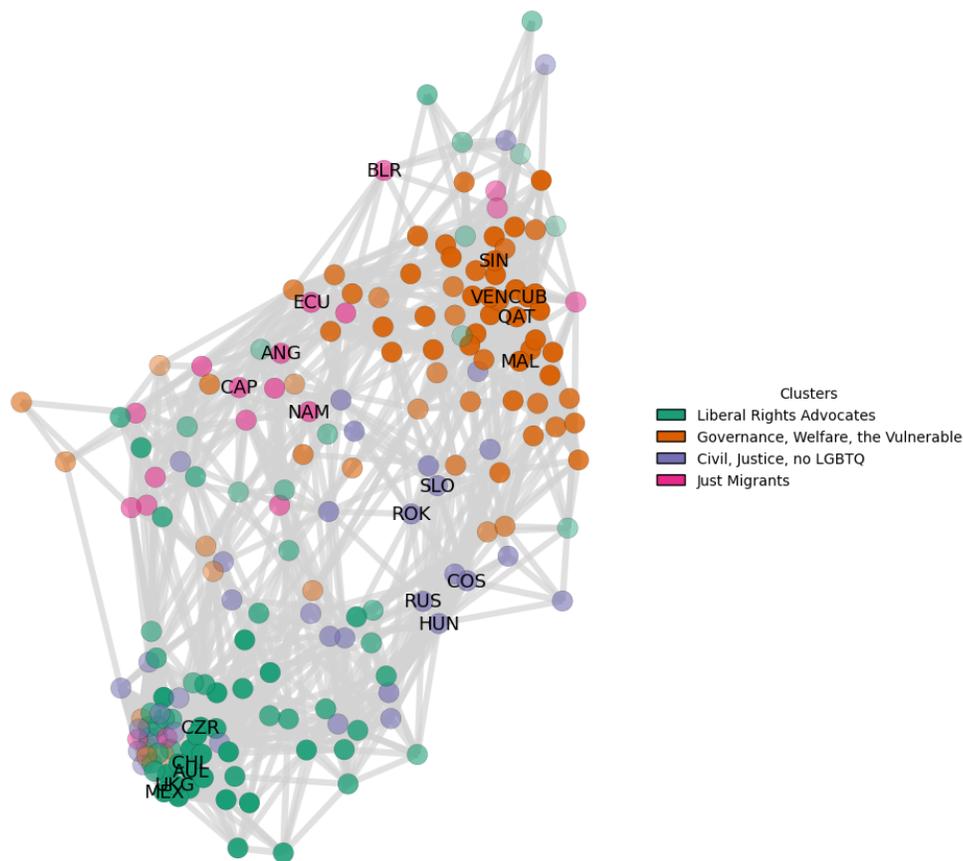

Figure 7: Cosine-similarity based clustering

Note: Gradation reflects stability scores, reflecting each country's average co-clustering with others in the same cluster across draws. Countries from each cluster with top 5 stability scores are labeled with their Correlates of War country code. A high stability score indicates that the state is consistently grouped with the same cluster-mates across posterior samples, whereas a low score suggests that its cluster membership is uncertain or varies substantially across draws.
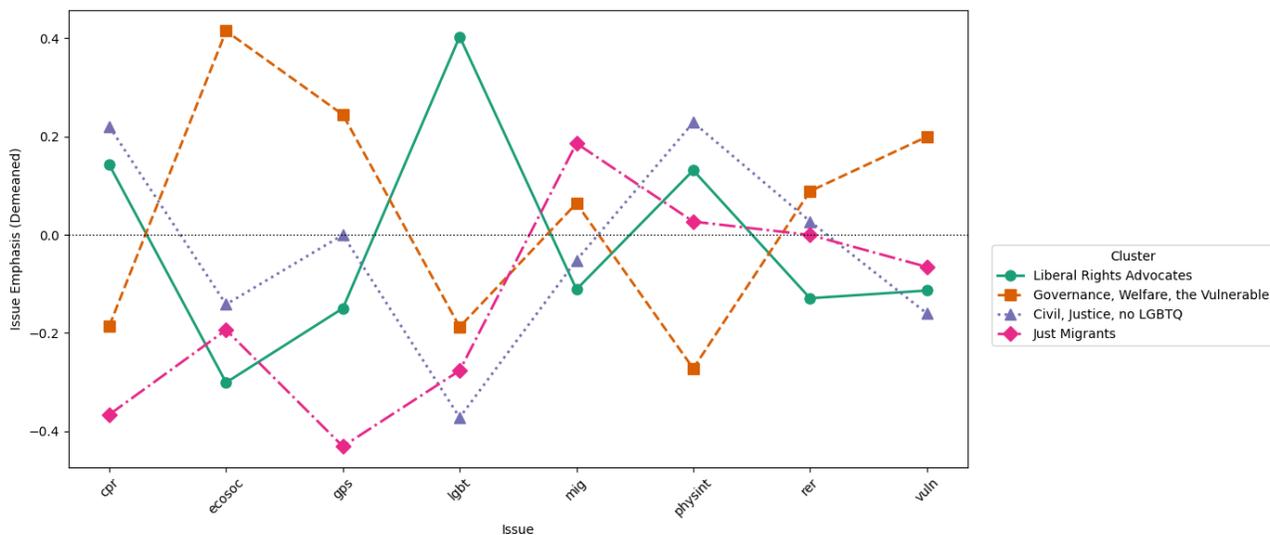
Figure 8: Characterization of cosine similarity clusters

Note: Centered values reflect deviations from each sender's average emphasis across issues. The number is obtained by averaging posterior mean values of a_{ik} for countries that belong to each cluster. The exact values are presented in Appendix Table D.1.

Figure 8 presents a snapshot of the relative issue emphasis of different local clusters identified. Table 3 complements this by showing the most prototypical countries per each cluster, based on their cosine similarity with the centroid of each cluster in the eight dimensional space. The *Liberal Rights Advocates* cluster is characterized by their relative issue emphasis in LGBTQ rights, civil and political rights, and physical integrity rights, and de-emphasis in socioeconomic rights and governance and public service-related issues. The United States, United Kingdom, and 70.4% of the 27 EU countries are in this bloc, with the remaining eight countries split between the *Governance, Welfare, the Vulnerable* cluster (Finland, Luxembourg, Portugal, Slovenia) and the *Civil, Justice, no LGBTQ* cluster (Germany, Hungary, Poland, Slovakia).

Governance, Welfare, the Vulnerable cluster, characterized by their issue emphasis in socioeconomic rights and governance and public service-related rights and de-emphasis in civil and political rights and physical integrity rights, aligns with increasing scholarly descriptions of some states' normative agendas within the international human rights network, headed by China under Xi Jinping. China under Xi Jinping has famously sought to redefine the global human rights discourse by prioritizing state-led development and stability over individual civil and political rights. As the Chinese

21

Communist Party's official ideological journal goes: "without the right to peace and the right to development, other human rights cannot even be brought up" Mitchell (2021, 618-619). Pauselli, Urdínez, and Merke (2023) argues that China's approach to human rights is also relativism, "which holds that the interpretation of human rights should be based on the specific circumstances of each country," which implies a natural prioritization of governance and public services related shaming and government-centered approach to human rights. Top countries in this bloc are often authoritarian and maintains a close tie with China (see Table 3).

Civil, Justice, no LGBTQ cluster's shaming profile is very similar to that of the *Liberal Rights Advocates*, except for one issue: LGBTQ rights. Countries included in this cluster includes Russia, who has consistently enacted laws that discriminate against LGBTQ individuals in the name of traditional values and nation building against foreign interference (Wilkinson 2014; Chandler 2020; Chaney 2018). While the majority of this cluster is authoritarian countries (Pauselli and Urzúa 2024), it includes East Asian democracies that do not legalize same-sex marriage, such as Japan or South Korea.

Lastly, the *Just Migrants* cluster prioritizes shaming in migrant's rights and relatively deprioritizes most common issue domains such as the civil and political rights, socioeconomic rights, governance and public services related rights, while not showing much emphasis or de-emphasis in other issue domains such as physical integrity rights, discrimination based on race and religion, or protection of vulnerable populations. The most prototypical countries in this cluster are Ecuador and Cape Verde, which are all significant migrant exporting or transit countries. The full list of countries in each community is shown in Appendix Table D.2.

Cluster	Countries (Cosine Similarity)
Liberal Rights Advocates	Brazil (0.969), Mexico (0.963), Australia (0.957), Belgium (0.939), France (0.938)
Governance, Welfare, the Vulnerable	Qatar (0.976), Bahrain (0.924), Oman (0.920), Cuba (0.896), Singapore (0.883)
Civil, Justice, no LGBTQ	Russia (0.928), Costa Rica (0.831), Botswana (0.810), Hungary (0.808), Tunisia (0.796)
Just Migrants	Ecuador (0.894), Cape Verde (0.886), Kazakhstan (0.879), Nepal (0.851), Angola (0.734)

Table 3: Prototypical Countries

Note: The countries were selected based on their cosine similarity with the centroid of each cluster. The number in the bracket shows the cosine similarity between the country and the centroid of each cluster.

Taken together, the community detection results show that the human rights regime contains structured and persistent normative blocs rather than a homogeneous field of universal standards. These communities arise from systematic similarities in states' issue-specific shaming profiles, after removing the effects of political alignment and network position. As such, they offer a principled map of why some issues receive sustained international attention while others remain marginal.[14] The resulting structure also provides analytic traction for theorizing how distinct interpretations of human rights diffuse, compete, and become institutionalized within the UPR process. Notably, the alignment between certain communities' shaming practices and their domestic human rights priorities—particularly where political rights or minority protections are weakened—echoes concerns in existing scholarship. Prior work argues that governments sometimes support international norms strategically to mitigate future challenges to their own authority (Moravcsik 2000). The communities identified here are consistent with that logic: states may bolster select global norms while diverting attention away from rights that pose greater domestic constraints.

4.4 Discussion

4.4.1 Issues and their relationships

To identify which issues most strongly differentiate the detected communities on average, I compute the variance of standardized shaming emphasis across communities for each issue area. This approach follows ANOVA logic: If a particular issue shows high between-group variance in standardized residuals, it is more likely to be divisive, i.e., emphasized by some communities and downplayed by others. This method provides a simple, interpretable ranking of issue areas by their contribution to community separation and highlights the dimensions along which normative preferences are most sharply contested. Table 4's first column reports this.

To complement the variance-based approach, I train a random forest classifier to predict each country's community membership based on its standardized issue profile (Breiman, 2001; Breiman

---

14. Because the clustering is derived from fully Bayesian posterior draws of sender–issue effects, and consensus membership is determined through a principled Dahl method, the resulting communities are closer to latent groupings rather than simple exploratory clusters.

et al., 2018). To assess which issue areas are most influential in distinguishing communities, I use the model's built-in measure of feature importance: the mean decrease in classification accuracy. This is computed by permuting each issue variable in the out-of-bag (OOB) samples and observing the increase in prediction error. If shuffling an issue leads to a significant drop in accuracy, it indicates that the variable plays a key role in the model's ability to differentiate normative communities. This model-based analysis captures interactions and nonlinearities in the data, offering a robust complement to the variance-based ranking of divisive issues. This is displayed in Table 4's second column. The results in Table 4 consistently shows that LGBTQ rights, physical integrity rights, and civil and political rights are among the top issues that differentiate between the four local clusters identified above.

Issue	Between-Group Variance	Mean Decrease Accuracy
LGBTQ Rights	0.086	0.150
Physical Integrity Rights	0.050	0.144
Civil-Political Rights	0.048	0.054
Socioeconomic Rights	0.058	0.027
Protection of Vulnerable Populations	0.014	0.025
Governance and Public Services	0.033	0.025
Migration	0.026	0.023
Racial, Ethnic, and Religious Minorities	0.008	0.021

Table 4: Between-group variance and variable importance by human rights issue

Note: Between-group variance reflects how much standardized shaming emphasis varies across detected communities, following ANOVA logic. Mean decrease in accuracy comes from a random forest classifier predicting community membership; it captures the loss in predictive performance when an issue is permuted, with higher values indicating greater importance for distinguishing community assignment.

Given the distinct issue profile clusters, are there any patterns that summarize the tensions and contestations in different human rights issue domains? For example, one of the long-standing divisions in global human rights discourse is that between the civil and political rights and socioeconomic rights (Jensen 2016). In the contemporary debates, China has revived a concept of rights hierarchy that foregrounds socioeconomic development and stability over mechanisms of individual accountability and procedural protections (Mitchell 2021).

To further examine how issue areas relate to one another, Figure 9 presents principal component analysis (PCA) on the posterior means of the 193 by 8 matrix of sender-issue random effects. The

figure provides a compact way to examine how issue areas relate to one another. Because PCA identifies the directions of greatest covariance in the sender–issue effects matrix, issues that tend to rise and fall together across states load in similar directions in the reduced space. When plotted as arrows in the biplot, issues that are strongly positively correlated appear close together, issues that are uncorrelated appear roughly orthogonal, and issues that are negatively related point in opposite directions. We can see that civil and political rights, physical integrity rights, and LGBT rights tend to vary together across the sender, with an orthogonal relationship with the other cluster of issues consisted of socioeconomic rights, migrant rights, governance and public services related rights, race and religious minority rights, and rights of the vulnerable populations. The presented dimensions, PC1 and PC2 explain about 55% of the data.

The correlation matrix generated with posterior draws of the sender–issue random effects also shows similar issue clustering patterns. (see Appendix Figure E.1).

4.4.2 Influential Countries

To further assess the robustness of the network structure while also finding interesting patterns like the influence of individual countries, I performed a jackknife resampling procedure in which each country was iteratively removed and the resulting community assignment compared to the original using the Adjusted Rand Index (ARI). Lower ARI values indicate that a country's removal disrupts the topology of the network, causing reconfigurations in cluster boundaries, implying greater structural influence. The distribution is well balanced around the median (see Figure E.2).

For example, Mauritius produced the largest drop in clustering consistency (ARI = 0.279), suggesting that it plays a central role in anchoring the *Governance, Welfare, and the Vulnerable* cluster of states with similar shaming profiles. In contrast, China ranks 84th out of 193 in influence in sustaining the cluster. China's relatively low influence score in the jackknife analysis does not contradict its well-documented role in promoting a governance-centered reinterpretation of human rights norms. Rather, it reflects the distinction between ideological leadership in discourse and structural centrality in the observed shaming behavior network. China's sender–issue profile, although emblematic of the broader normative project associated with state-led development and political stability, is not
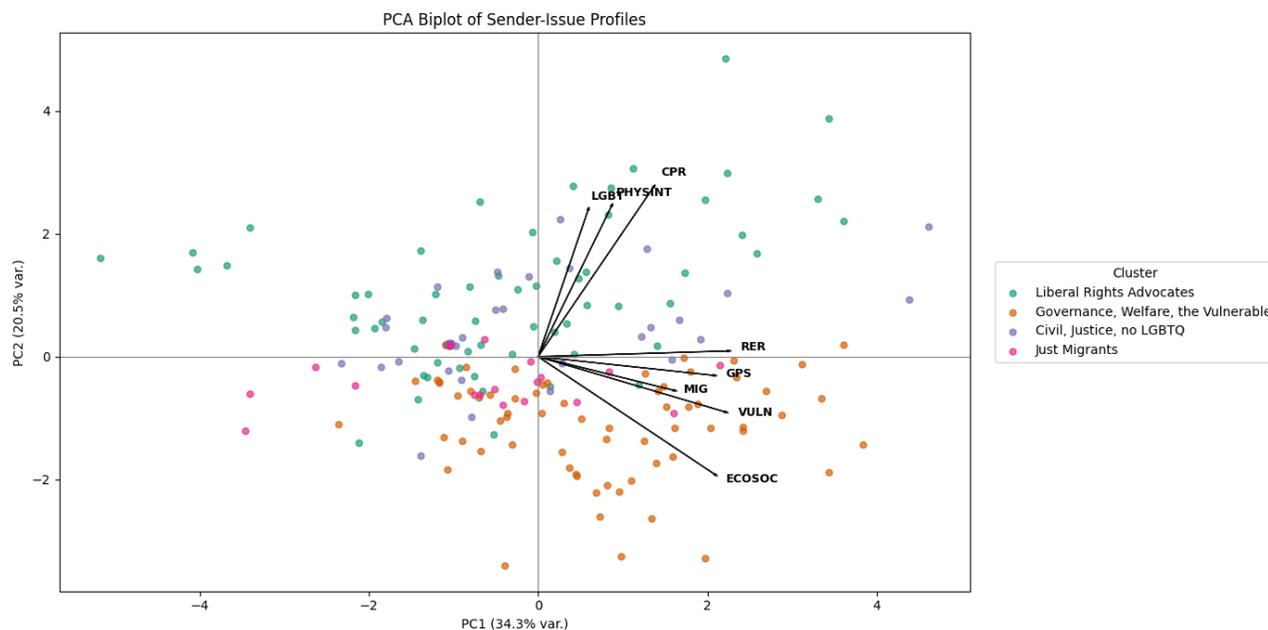
25

Figure 9: Visualization of country-country, country-issue, issue-issue relationships

Note: In the PCA biplot, the location of each state represents its overall pattern of conditional issue emphasis projected into the two-dimensional PCA space. The issue vectors indicate the direction in which that issue increases most strongly across states. If a state lies in the same direction as an issue vector, it means the state places above-average emphasis on that issue relative to other issues. Moving along the arrow means moving toward higher relative emphasis on that issue. The closer the state lies to the tip of the vector, the stronger that emphasis. Longer vectors reflect issues that explain more of the total variance, and shorter vectors indicate issues that vary little across states. Small angles between vectors indicate strong positive correlation.

uniquely definitive within the cluster. Many states in this community—including several smaller authoritarian or hybrid regimes—exhibit highly similar or even more internally consistent patterns of emphasizing socioeconomic rights and public services while downplaying civil-political rights. As a result, the removal of China does not significantly alter the community structure, because the cluster is redundantly anchored by numerous states whose shaming profiles closely resemble China's. Table 5 presents top structurally important countries for each cluster.

Most countries emerge as unexpectedly central to each cluster. Their influence stems not from diplomatic reach, but from the consistency and distinctiveness of their shaming practices, which lend normative stability to their respective communities. This pattern aligns with findings in the foreign aid literature: minor powers, with limited influence in global politics, often emphasize human rights in their foreign policy as a principled stance (Dietrich and Murdie 2017). In contrast, major donor

Country	Jackknife ARI Score	Cluster
Australia	0.281	*Liberal Rights Advocates*
Mauritius	0.279	*Governance, Welfare, and the Vulnerable*
San Marino	0.286	*Civil, Justice, no LGBTQ*
Cape Verde	0.283	*Just Migrants*

Table 5: Countries with Highest Structure Influence

Note: Lower ARI scores indicate countries whose removal has the greatest effect on the community structure, that its removal shifts other countries' community assignments.

governments, which are more entangled in strategic interests, tend to be guided by "realpolitik" considerations that can override human rights concerns (Dietrich and Murdie 2017). These findings highlight the need to look beyond the general overshaming and undershaming patterns and traditional power politics to understand how international norms are shaped, challenged, and embedded in global institutions.

5   Conclusion

This paper set out to reassess how issue domain shapes interstate shaming at the Universal Periodic Review and to evaluate a prominent claim that states reserve regime-threatening criticism for adversaries while using safer issues to discipline allies. The first set of analyses revisited this claim (Hypothesis 1) using both corrected OLS specifications and a hierarchical pooled AME model that permits direct comparison of issue effects. Across these approaches, the evidence for selective, issue-contingent shaming is weak. Affinity exerts a broadly uniform effect on shaming across issue areas, and the predicted pattern of adversaries being targeted primarily on costly issues while allies are criticized on soft issues does not emerge. Instead, the results support a simpler affinity-based logic: states tend to be more lenient toward allies across most domains, rather than finely calibrating criticism according to issue sensitivity.

The second contribution concerns heterogeneity in issue preferences (Hypothesis 2). By estimating a pooled AME model with sender–issue random effects, the paper explicitly tests whether states differ in their relative emphasis on particular issues once political alignment, target characteristics,

and network structure are controlled for. The variance of the sender–issue effects, combined with model comparison using PSIS-LOO and WAIC, shows that allowing sender effects to vary across issue domains substantially improves predictive performance. This indicates that states do exhibit systematic, issue-specific shaming profiles that cannot be captured by static sender effects alone. In other words, shaming behavior is not only a function of who is shaming whom, it also reflects how states prioritize different categories of rights within a shared institutional setting.

The third contribution is to characterize the structure of this heterogeneity (Hypothesis 3). Clustering the posterior distribution of sender–issue random effects reveals that, once affinity and network dependencies are accounted for, global variation in issue emphasis is relatively coarse. At the system level, the dominant cleavage lies between generalist overshamers, generalist undershamers, and a set of states that selectively disengage from socioeconomic rights and governance and public services. A further layer of analysis based on cosine similarity uncovers more fine-grained communities, such as a liberal rights cluster that emphasizes LGBT rights, civil and political rights, and physical integrity rights, and a governance and welfare cluster that prioritizes socioeconomic and governance-related issues while downplaying political rights. These local communities are not simple reflections of material alignments, but represent distinct normative packages that shape how states use shaming to define and contest human rights standards.

Taken together, the findings suggest a more nuanced view of the international human rights regime. Political alignment clearly matters for whether states shame at all, yet much of the remaining structure in shaming behavior reflects heterogeneity in normative emphasis rather than the sharp cost-based differentiation posited by issue sensitivity theories. Most states appear to vary primarily in their overall shaming activism, while a subset align into communities that advance particular constellations of rights and selectively neglect others. These communities are often anchored not only by major powers, but also by mid-sized and smaller states whose consistent and distinctive shaming profiles help stabilize the internal cohesion of normative blocs. Their influence derives less from material capabilities than from patterned engagement with human rights discourse.

Conceptually, the paper argues for an account of shaming that treats norms as heterogeneous and structured rather than uniform and exogenous. Methodologically, it demonstrates how hierarchical

network models and posterior-based clustering can recover latent patterns in issue emphasis that conventional regression designs obscure. Future research can build on this framework by tracing how these normative communities evolve over time, how domestic politics shape states' positions within them, and how institutional features of the UPR and other review mechanisms mediate the interaction between geopolitical affinity and heterogeneous human rights agendas.

References

Adams, James, Simon Weschle, and Christopher Wlezien. 2021. "Elite interactions and voters' perceptions of parties' policy positions." *American Journal of Political Science* 65 (1): 101–114.

Alston, Philip. 1993. "The importance of the inter-play between economic, social and cultural rights, and civil and political rights." *Human Rights at the Dawn of the 21st Century* 59:61.

Alter, Karen J, Emilie M Hafner-Burton, and Laurence R Helfer. 2019. "Theorizing the judicialization of international relations." *International Studies Quarterly* 63 (3): 449–463.

Arat, Zehra F. 1999. "Human rights and democracy: expanding or contracting?" *Polity* 32 (1): 119–144.

Baderin, Mashood A. 2003. *International human rights and Islamic law*. OUP Oxford.

Bassan-Nygate, Lotem. 2025. "Champions and Pariahs: Image Effects of Naming and Shaming." *Journal of Politics.*

Berger, Tobias. 2023. "Human Rights beyond the Liberal Script: A Morphological Approach." *International Studies Quarterly* 67 (3): sqad042.

Blondel, Vincent D, Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. "Fast unfolding of communities in large networks." *Journal of statistical mechanics: theory and experiment* 2008 (10): P10008.

Breiman, Leo. 2001. "Random forests." *Machine learning* 45:5–32.

Breiman, Leo, Adele Cutler, Andy Liaw, Matthew Wiener, and Maintainer Andy Liaw. 2018. "Package 'randomforest'." *University of California, Berkeley: Berkeley, CA, USA* 81:1–29.

Chandler, Andrea. 2020. "Russia's laws on 'non-traditional' relationships as response to global norm diffusion." *The International Journal of Human Rights* 25:616–638. https://doi.org/10.1080/13642987.2020.1789106.

Chaney, Paul. 2018. "Civil Society, 'Traditional Values' and LGBT Resistance to Heteronormative Rights Hegemony: Analysis of the UN Universal Periodic Review in the Russian Federation." *Europe-Asia Studies* 70:638–665. https://doi.org/10.1080/09668136.2018.1453924.

Cordell, Rebecca, K Chad Clay, Christopher J Fariss, Reed M Wood, and Thorin M Wright. 2020. "Changing standards or political whim? Evaluating changes in the content of US State Department Human Rights Reports following presidential transitions." *Journal of Human Rights* 19 (1): 3–18.

Dahl, David B. 2006. "Model-based clustering for expression data via a Dirichlet process mixture model." *Bayesian inference for gene expression and proteomics* 4:201–218.

Dietrich, Simone, and Amanda Murdie. 2017. "Human rights shaming through INGOs and foreign aid delivery." *The Review of International Organizations* 12 (1): 95–120.

Dixon, Jennifer M. 2017. "Rhetorical adaptation and resistance to international norms." *Perspectives on Politics* 15 (1): 83–99.

Donnelly, Jack. 1988. "Human rights at the United Nations 1955–85: the question of bias." *International Studies Quarterly* 32 (3): 275–303.

Donno, Daniela. 2010. "Who is punished? Regional intergovernmental organizations and the enforcement of democratic norms." *International Organization* 64 (4): 593–625.

Donno, Daniela, Sara Fox, and Joshua Kaasik. 2022. "International incentives for women's rights in dictatorships." *Comparative Political Studies* 55 (3): 451–492.

Dukalskis, Alexander. 2023. "A fox in the henhouse: China, normative change, and the UN Human Rights Council." *Journal of Human Rights* 22 (3): 334–350.

Fariss, Christopher, Michael Kenwick, and Kevin Reuning. 2020. *Latent Human Rights Scores Version 4.* Harvard Dataverse. https://doi.org/10.7910/DVN/RQ85GK. https://doi.org/10.7910/DVN/RQ85GK.

Finnemore, Martha, and Kathryn Sikkink. 1998. "International norm dynamics and political change." *International organization* 52 (4): 887–917.

Gade, Emily Kalah, Michael Gabbay, Mohammed M Hafez, and Zane Kelly. 2019. "Networks of cooperation: Rebel alliances in fragmented civil wars." *Journal of Conflict Resolution* 63 (9): 2071–2097.

Gade, Emily Kalah, Mohammed M Hafez, and Michael Gabbay. 2019. "Fratricide in rebel movements: A network analysis of Syrian militant infighting." *Journal of Peace Research* 56 (3): 321–335.

Glendon, Mary Ann. 1997. "Knowing the universal declaration of human rights." *Notre Dame L. Rev.* 73:1153.

Hoff, Peter. 2015. "Dyadic data analysis with amen." *arXiv preprint arXiv:1506.08237.*

Hug, Simon, and Richard Lukács. 2014. "Preferences or blocs? Voting in the United Nations human rights council." *The review of international organizations* 9 (1): 83–106.

Hulvey, Rachel. 2023. "Polarizing the United Nations: China's Rise and the Dimensions of Ideological Change."

Hulvey, Rachel A. 2022. "Cyber Sovereignty: How China is Changing the Rules of Internet Freedom." *UC Institute of Global Conflict and Cooperation.*

Jensen, Steven LB. 2016. *The making of international human rights: the 1960s, decolonization, and the reconstruction of global values.* Cambridge University Press.

Kim, Yooneui. 2023. "A network analysis of naming and shaming in the universal periodic review." *International Interactions* 49 (3): 287–314.

Kinne, Brandon J. 2024. "Network context and the effectiveness of international agreements." *International Studies Quarterly* 68 (3): sqae099.

Lebovic, James H, and Erik Voeten. 2006. "The politics of shame: the condemnation of country human rights practices in the UNCHR." *International Studies Quarterly* 50 (4): 861–888.

McCarty, Nolan. 2011. "Measuring legislative preferences." *The Oxford Handbook of the American Congress.*

McKeown, Ryder. 2009. "Norm regress: U.S. revisionism and the slow death of the torture norm." *International relations* 23 (1): 5–25.

Mertus, Julie. 2007. "The rejection of human rights framings: The case of LGBT advocacy in the US." *Human Rights Quarterly* 29 (4): 1036–1064.

Meyerrose, Anna M, and Irfan Nooruddin. 2025. "Trojan horses in liberal international organizations? How democratic backsliders undermine the UNHRC: Contribution to the Special Issue "Autocratic Regimes, Democratic Backsliding, and International Organizations"." *The Review of International Organizations* 20 (1): 125–156.

Mitchell, Ryan. 2021. "The Emerging Chinese Model of Statist Human Rights." *Am. U. Int'l L. Rev.* 37:617.

Moravcsik, Andrew. 2000. "The origins of human rights regimes: Democratic delegation in postwar Europe." *International organization* 54 (2): 217–252.

Newman, Mark EJ. 2004. "Analysis of weighted networks." *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics* 70 (5): 056131.

Park, Baekkwan, Kevin Greene, and Michael Colaresi. 2020. "Human rights are (increasingly) plural: Learning the changing taxonomy of human rights from large-scale text reveals information effects." *American Political Science Review* 114 (3): 888–910.

Park, Chun-Young, and Sanghoon Park. 2025. "You scratch my back, and I scratch yours: Autocratic reciprocity in the politics of naming and shaming." *International Interactions* 51 (1): 1–28.

Pauselli, Gino, Francisco Urdínez, and Federico Merke. 2023. "Shaping the liberal international order from the inside: A natural experiment on China's influence in the UN human rights council." *Research & Politics* 10 (3): 20531680231193513.

Pauselli, Gino, and María-José Urzúa. 2024. "Why autocracies fear LGBTQ+ rights." *Journal of Democracy* 35 (2): 80–92.

Rozenas, Arturas, Shahryar Minhas, and John Ahlquist. 2019. "Modeling asymmetric relationships from symmetric networks." *Political Analysis* 27 (2): 231–236.

Sceats, Sonya, and Shaun Breslin. 2012. *China and the international human rights system.* Chatham house.

Terman, Rochelle, and Zoltán I Búzás. 2021. "A house divided: Norm fragmentation in the international human rights regime." *International Studies Quarterly* 65 (2): 488–499.

Terman, Rochelle, and Joshua Byun. 2022. "Punishment and politicization in the international human rights regime." *American Political Science Review* 116 (2): 385–402.

Terman, Rochelle, and Erik Voeten. 2018. "The relational politics of shame: Evidence from the universal periodic review." *The Review of International Organizations* 13 (1): 1–23.

Toosi, Nahal. 2025. *Trump drastically cutting back annual human rights report,* March. https://www.politico.com/news/2025/03/19/trump-human-rights-report-00238581.

Vehtari, Aki, A. Gelman, and Jonah Gabry. 2015. "Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC." *Statistics and Computing* 27:1413–1432. https://doi.org/10.1007/s11222-016-9696-4.

Vehtari, Aki, and Janne Ojanen. 2012. "A survey of Bayesian predictive methods for model assessment, selection and comparison." *Statistics and Computing,* https://doi.org/{https://doi.org/10.1007/s11222-0169696-4.}.

Voeten, Erik, Anton Strezhnev, and Michael Bailey. 2009. *United Nations General Assembly Voting Data.* Harvard Dataverse. https://doi.org/10.7910/DVN/LEJUQZ. https://doi.org/10.7910/DVN/LEJUQZ.

Watanabe, Sumio. 2009. *Algebraic geometry and statistical learning theory.* Vol. 25. Cambridge university press.

Wilkinson, C. 2014. "Putting "Traditional Values" Into Practice: The Rise and Contestation of Anti-Homopropaganda Laws in Russia." *Journal of Human Rights* 13:363–379. https://doi.org/10.1080/14754835.2014.919218.

Part

Supplementary Material

Table of Contents

A Thematic Issue Clusters

Thematic Category	Issues
Civil-Political Rights	Civil Political Rights; Civil Society; Elections; Freedom of Association and Peaceful Assembly; Freedom of Opinion and Expression; Freedom of the Press; Human Rights Defenders
Governance and Public Services	Corruption; Public Security; Human Rights Education and Training; Counter-terrorism
Migration and Migrant Labor	Asylum Seekers and Refugees; Freedom of Movement; Labor; Migrants; Internally Displaced Persons
Physical Integrity Rights	Justice; Death Penalty; Detention; Enforced Disappearances; Extrajudicial Executions; Human Rights Violations by State Agents; Impunity; Torture and Other CID Treatment
Racial, Ethnic and Religious Minorities	Freedom of Religion and Belief; Minorities; Racial Discrimination; Indigenous Peoples
Socio-Economic Rights	Environment; Right to Land; Right to Water; Economic, Social and Cultural Rights; Development; Poverty; Right to Education; Right to Food; Right to Health; Right to Housing
Protection of Vulnerable Populations	Disabilities; HIV / AIDS; Sexual Orientation and Gender Identity; Rights of the Child; Trafficking; Women's Rights

Table A.1: Thematic categories and corresponding human rights issues.

This table is taken from Terman and Byun (2022), Appendix page 4.

B Hypothesis 1

B.1 OLS

Table B.1: Effects of Geopolitical Affinity on Issue Shaming Counts (Regression for Figure 1)

Term	Estimate	Std. Error	Statistic	p-value
(Intercept)	0.679***	0.175	3.881	0.000
Political Affinity	0.012	0.017	0.716	0.474
Civil-Political Rights	0.640***	0.024	26.401	0.000
Governance and Public Service	0.421***	0.027	15.574	0.000
Migration	0.420***	0.023	18.099	0.000
Physical Integrity Rights	0.668***	0.014	47.457	0.000
Race and Religious Discrimination	0.492***	0.025	19.947	0.000
Socio-Economic Rights	0.409***	0.022	18.755	0.000
Women, Children, LGBTQ, Disabled	0.554***	0.023	24.025	0.000
Political Affinity:Civil-Political Rights	-0.000	0.004	-0.126	0.900
Political Affinity:Governance and Public Service	0.006	0.014	0.412	0.681
Political Affinity:Migration	-0.056***	0.010	-5.503	0.000
Political Affinity:Physical Integrity Rights	-0.017**	0.008	-2.070	0.038
Political Affinity:Race and Religious Discrimination	-0.047***	0.012	-4.021	0.000
Political Affinity:Socio-Economic Rights	-0.021	0.016	-1.309	0.190
Political Affinity:Women, Children, LGBTQ, Disabled	-0.021**	0.010	-2.170	0.030
Mean Severity	0.076***	0.013	5.649	0.000
Physint Score (Sender)	-0.037	0.042	-0.878	0.380
Physint Score (Target)	-0.070	0.055	-1.280	0.200
Same Region	0.054**	0.017	3.187	0.001

Continued on next page

Table B.1: Effects of Geopolitical Affinity on Issue Shaming Counts (Regression for Figure 1) (continued)

Term	Estimate	Std. Error	Statistic	p-value
UPR Reviewed Session	0.016	0.026	0.605	0.544

Observations: 14093

R^2: 0.778

Adjusted R^2: 0.771

F-statistic : 126.2 on 379 and 13713 DF, p-value: $< 2.2e - 16$

Residual Std. Error 0.7775 (df=13713)

Note: *p<0.1; **p<0.05; ***p<0.01.

C   Hypothesis 2

C.1   Pooled AME (Restricted)

Model   Difference with the pooled model in the main text: everything else is the same, except for the shaming sender-issue specific random effects a_{ik} is replaced with a_k.

Main results   The effects of geopolitical affinity is very similar to that of the main model in terms of magnitude and credibility intervals.

Figure C.1: Marginal Log-Effects of Geopolitical Affinity on Shaming Counts, by Theme

Note: Credible intervals are constructed at the 95% level using posterior draws.

C.2   Pooled AME (LGBTQ added)

Main results   The effects of geopolitical affinity is very similar to that of the main model in terms of magnitude and credibility intervals.
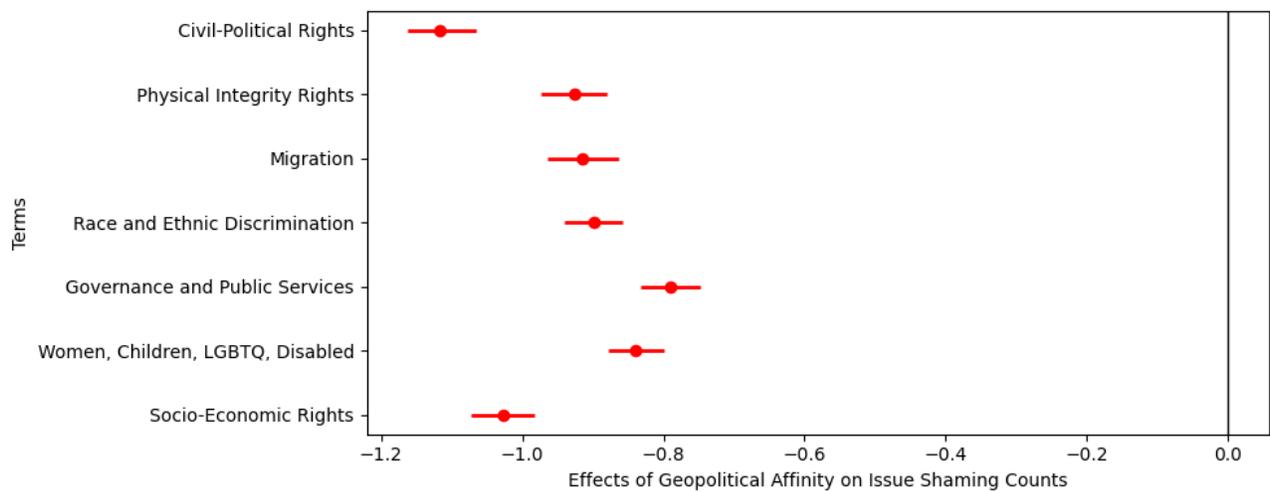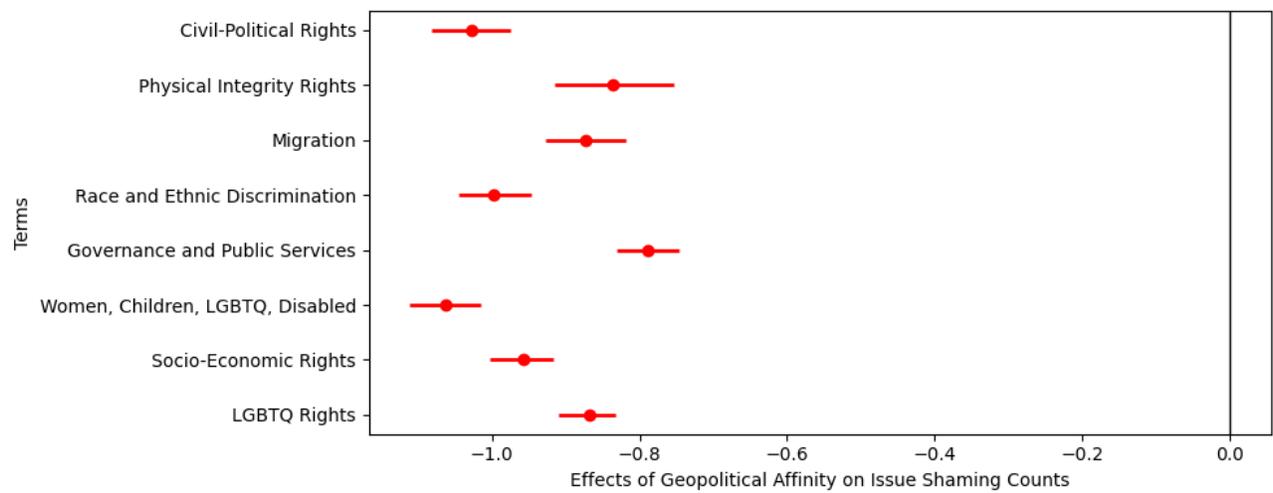
Figure C.2: Marginal Log-Effects of Geopolitical Affinity on Shaming Counts, by Theme

Note: Credible intervals are constructed at the 95% level using posterior draws.

D   Hypothesis 3

D.1 Issue Preferences by Cosine Cluster

Community	Interpretive Label	cpr	ecosoc	gps	lgbt	mig	physint	rer	vuln
0	Liberal Rights Advocates	0.14	-0.30	-0.15	0.40	-0.11	0.13	-0.13	-0.11
1	Governance, Welfare, the Vulnerable	-0.19	0.41	0.24	-0.19	0.06	-0.27	0.08	0.19
2	Civil, Justice, no LGBTQ	0.22	-0.14	-0.00	-0.37	-0.05	0.23	0.03	-0.16
3	Just Migrants	-0.37	-0.19	-0.43	-0.28	0.19	0.03	-0.00	-0.07

Table D.1: Average Issue Emphasis by Cluster

Note: The number is obtained by averaging the posterior mean values of a_{ik} for countries that belong to each cluster. Higher values indicate greater relative emphasis on a given issue, net other factors.

D.2   Full List of Community Membership

Liberal Rights Advocates

Afghanistan (0.111)	Albania (-0.161)	Andorra (-0.204)	Argentina
Australia	Austria	Barbados	Belgium
Belize	Brazil	Burundi (0.177)	Bulgaria
Canada	Central African Republic	Chile	Colombia
Congo (Brazzaville) (-0.014)	Croatia	Cyprus	Czech Republic
Denmark	Dominican Republic (-0.254)	Estonia	Ethiopia
East Timor (-0.323)	Fiji (-0.270)	France	Federated States of Micronesia
Gambia	Guinea-Bissau	Greece	Honduras
India (-0.044)	Ireland	Israel	Italy
Kenya	Latvia	Lithuania (0.188)	Madagascar
Malawi	Mexico	Mali	Malta (0.159)
Montenegro	Mongolia	Marshall Islands	Mozambique
Nauru	Nigeria (0.068)	Netherlands	Paraguay
Peru (-0.145)	North Korea (-0.055)	Romania	South Africa
St. Kitts and Nevis	Spain	Sao Tome and Principe	Swaziland
Tanzania (-0.448)	Thailand (0.170)	Tonga	United Kingdom
Ukraine	USA	Samoa	Zambia

Governance, Welfare, the Vulnerable

Algeria	Azerbaijan	Bahrain	Bahamas (0.118)
Bhutan	Bangladesh	Bolivia	Brunei
Cambodia	China	Cuba	Djibouti
Dominica (-0.730)	Egypt	Equatorial Guinea	Finland
Grenada (-0.753)	Haiti	Iceland (0.121)	Iran
Jamaica	Jordan	Kuwait	Kyrgyzstan
Laos	Liberia (-0.714)	Lesotho	Libya
Luxembourg	Malaysia	Mauritius	Moldova
Monaco	Morocco	Myanmar	New Zealand
Nicaragua	Norway	Oman	Pakistan
Panama	Philippines	Portugal	Qatar
El Salvador	Saudi Arabia	Senegal	Seychelles
Singapore	St. Lucia (-0.300)	Slovenia	Solomon Islands
Somalia	Sri Lanka	South Sudan (0.016)	Sudan
Suriname (-0.735)	St. Vincent and the Grenadines (0.173)	Syria	Tajikistan
Turkmenistan	Trinidad and Tobago	Turkey	United Arab Emirates
Uganda (0.193)	Uruguay	Venezuela	Serbia
Zimbabwe			

Civil, Justice, no LGBTQ

Antigua & Barbuda (0.154)	Armenia	Benin	Botswana
Cameroon	Côte d'Ivoire	Chad	Costa Rica
Democratic Republic of the Congo	Gabon (0.140)	Ghana	Germany
Georgia	Guyana (0.104)	Hungary	Indonesia
Iraq	Japan	Kiribati (0.111)	Lebanon
Liechtenstein (0.166)	Mauritania (-0.146)	Maldives	Papua New Guinea (0.092)
Poland	South Korea	Russia	Sierra Leone
Slovakia	San Marino (0.110)	Tunisia	Tuvalu (0.115)
Uzbekistan	Vanuatu (-0.500)		

Just Migrants

Angola	Burkina Faso	Belarus	Bosnia and Herzegovina
Cape Verde	Comoros (-0.094)	Ecuador	Eritrea (-0.262)
Guatemala	Guinea (-0.220)	Kazakhstan	Macedonia
Namibia	Nepal	Niger	Palau (-0.243)
Rwanda	Togo	Yemen (-0.070)	

Table D.2: Countries grouped by detected community

Note: Countries that have less than 0.2 cosine similarity with the centroid of each cluster have an additional note about the exact similarity inside the parentheses. Countries with cosine similarity noted implies these countries are at the border of each community, not sharing much similarity in issue profiles with the rest of the community members, although classified into this group.

E Discussion

E.1 Issue to Issue Relationship

Figure E.1 visualizes the posterior mean correlation matrix of the sender–issue random effects. Each correlation coefficient is computed across the full set of posterior draws, which ensures that the heatmap reflects not only point estimates but also the underlying uncertainty in how issue emphases co-move. The resulting structure highlights several broad regimes of normative contestation. Issue areas traditionally emphasized in liberal and procedural democratic frameworks – civil and political rights and physical integrity rights – exhibit consistently weak or negative correlations with socioeconomic rights.

Figure E.1: Association between shaming in each issue, net political costs
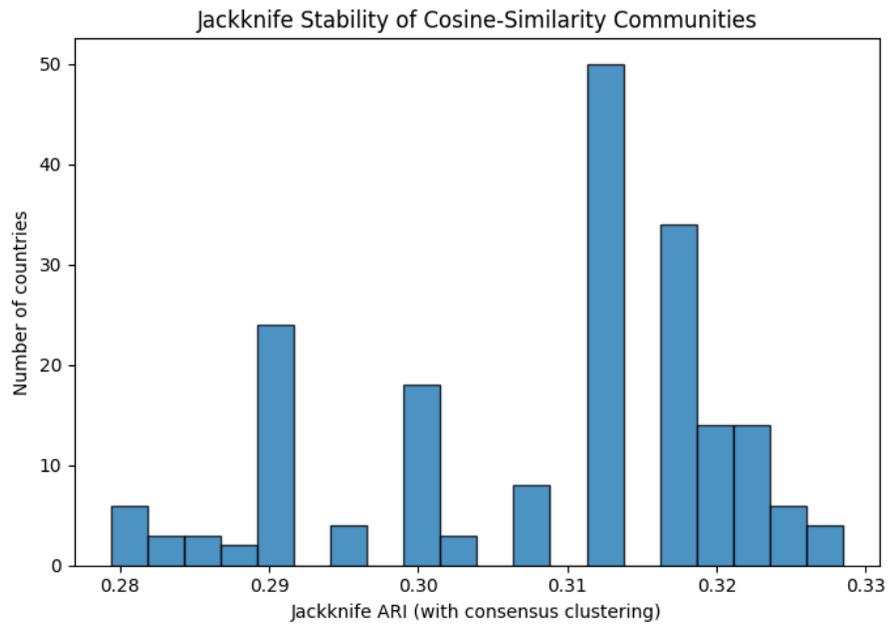
E.2 Jackknife stability

Figure E.2: Jackknife Stability (ARI)

Note: The ARI values are symmetric around the median of 0.313, meaning the community assignment is relatively stable.

Appendix References

Terman, Rochelle, and Joshua Byun. 2022. "Punishment and politicization in the international human rights regime." *American Political Science Review* 116 (2): 385–402.