THE APPLICATION OF THE LEIDEN ALGORITHM:

CLASSIFYING TARGET COUNTRIES BASED ON HROS' SHAMING PREFERENCE

by

SHANSHAN LIAN

(Under the Direction of Justin Strait)

ABSTRACT

A topic of interest in IR (International Relations) is in understanding how HROs (Human Rights International Non-Governmental Organizations) with different goals in human rights interact with countries. This requires us to think about communities of HROs based on issue preference associated with countries, as opposed to the group of all HROs. By identifying HROs' shaming countries as network events, the project combines a community detection method and IR literature to create a new measure at the country-year level, issue preference. The variable indicates the issue prioritized in HROs' shaming a country. The new measure is valid based on internal and external validity checks. Potential inquiries in HROs regarding issue preference can offer insights for policymakers and HROs in the current political environment.

INDEX WORDS: human rights, international non-governmental organizations, community detection, bipartite networks, naming and shaming

THE APPLICATION OF THE LEIDEN ALGORITHM:

CLASSIFYING TARGET COUNTRIES BASED ON HROS' SHAMING PREFERENCE

by

SHANSHAN LIAN

MA, Shanghai Institutes for International Studies, 2017

A Thesis Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment

of the Requirements for the Degree

Master of Science

ATHENS, GEORGIA

2023

THE APPLICATION OF THE LEIDEN ALGORITHM:

CLASSIFYING TARGET COUNTRIES BASED ON HROS' SHAMING PREFERENCE

by

SHANSHAN LIAN

Major Professor: Justin Strait
Committee: Amanda Murdie
Ray (Shuyang) Bai

Electronic Version Approved:

Ron Walcott
Vice Provost for Graduate Education and Dean of the Graduate School
The University of Georgia
August 2023

DEDICATION

The work is dedicated to my parents, Quanfa Lian and Dongmei Shan, and my husband, Jie Lian. Thank you for constantly encouraging me to pursue the life I want.

ACKNOWLEDGEMENTS

I am deeply indebted to Dr. Justin Strait, the major advisor for my thesis. This endeavor would not have been possible without his guiding and mentoring. Also, I could not have undertaken this journey without the generous support from Dr. Amanda Murdie, my major advisor in my Ph.D. program. In addition, I would like to thank Dr. Ray Bai, whose teaching inspires me a lot. Further thanks are extended to Dr. Baekkwan Park and Dr. Robert Brathwaite, who were crucial contributors in sample collection.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

HROs (Human Rights International Non-Governmental Organizations) are INGOs (International Non-Governmental Organizations) focusing on human rights issues. Some of their efforts in transnational advocacy have impressively pressured governments to improve local human rights (Keck and Sikkink 1998; Risse and Sikkink 1999). By networking with different supporters, such as liberal countries, international organizations, and local activists, for instance, HROs can pressure governments to improve human rights through policymaking in environmental protection and women's rights (Keck and Sikkink 1998).

There exist selection effects, where the actors in IR, such as countries and international organizations, "self-select" into preferred cases (Von Stein 2005; Fuhrmann and Lupu 2016), in HRO-to-government interactions. IR (International Relations) scholarship suggests many factors driving HROs' self-selection. HROs tend to initiate naming and shaming of the countries where human rights violations are severe (Keck and Sikkink 1998). The level of issue emergency could also affect HROs' advocacy agenda towards countries (Carpenter 2007). Also, HROs' concerns about authority could be another reason for country selection (Stroup and Wong 2017). In addition, Chaudhry (2019) points out that the probability of success in issue areas can affect HROs' report bias in India.

The aim of this thesis is to understand how HROs working on different human rights issues self-select into interactions with certain countries. In particular, I aim to address the following questions: How do HROs' issue selection preferences, motivated by these factors, project on countries? In other words, given HROs' motivations to prioritize certain problems, which issues within a country tend to be preferred by HROs?

To address these questions, I conceptualize *issue preference* as the issue prioritized by HROs to a target country. *Issue preference* suggests the top problem emphasized by HROs in a country. For instance, if most HROs spotlight police violence in the U.S., the *issue preference* of the country is police accountability.

By implication, *issue preference* is an indicator describing countries regarding the issues attacked by HROs. Hence, it is challenging to create *issue preference* at the country level through the attribute of HROs, preference on issues or goals since one country tends to be attacked by different HROs with distinct goals. Specifically, since HROs' goals can be identified as one source of heterogeneity, HROs can be further classified through the specific type of human rights they aim to improve. Further, TSMO (Transnational Social Movement Organizations) data (Hughes et al. 2018; Smith et al. 2019) summarize different goals of HROs. Even for an HRO, the organizations can set priorities for different goals, such as goal 1, goal 2, and goal 3, coded in TSMO data. In this project, although I only focus on the primary goal of HROs, goal 1 in the TSMO data, to define *issue preference*, the diffusing focus across HROs makes it difficult to identify the most preferred issue/s towards a country that can interact with HROs with various specific goals. This is the first challenge in measuring issue *preference*.

Another challenge is that, if relying on the counts of interactions with a government from different types of HROs to identify the preferred issue, the dominance of leading HROs, ones with significant activeness in interacting with countries (Stroup and Wong 2017), could bias the outcome of *issue preference*. For many countries, the top issues might come from a small group of leading HROs, which actively and significantly name and shame different countries.

However, for other HROs, even though the frequencies of interaction against countries are lower, their efforts can still be influential if, for instance, their naming and shaming are publicized in news reports. Therefore, it is necessary but challenging to incorporate the contributions to issue preference from both leading and non-leading HROs.

To address these problems, I rely on tools for community detection, a technique in network analysis, to create the measure of *issue preference*. Network analysis has been introduced to IR scholarship (e.g., Kim 2023; Kinne 2013; Kinne 2018; Kinne and Maoz 2023). However, according to my knowledge, community detection, especially in the data with nodes of two different types, has not been used to study HROs.

In this project, the logic of community detection is to classify HROs and countries based on the frequency of interactions. The number of communities can be determined based on (1) tuning the hyperparameter of my selected community detection algorithm and (2) integrating the subject-matter knowledge about overall preferences of HROs.

In community detection, the community of communities for a given sample can be hard to infer. For instance, for a sample, 5 communities and 6 communities can both be meaningful. Since it is hard to pick up an optimal number for communities simply through statistical modeling, I further depend on the expertise of IR to decide the number of communities. In other words, the communities formed based on the selected number should be interpretable in the context of IR. As a result, after detecting communities containing both HROs and countries, I use HROs' attributes in goals to identify which issue the organizations prioritize in attacking the countries. By implication, the countries in the same group can be understood as the ones preferred by HROs in the issue. The logic of the community detection process for each yearly data set can be found in **Figure 1**. The project focuses on HROs' shaming of countries, a typical form of HRO-to-country interaction.

Figure 1: The Logic of Measuring HROs' *Issue Preference* at the Country Level

This project is a new attempt to detect communities of HRO-to-country interactions in the case of shaming. The introduction of network statistics addresses the diffusion of HRO goals and the dominance of leading HROs. Moreover, the measure of HRO *issue preference* at the country level can help IR scholars further the exploration of HROs and countries. Instead of hierarchical models, which can include attributes of HROs and countries but separate two types of actors and mainly count the interactions, I offer an alternative to study the interactions between two different actors in IR. In this project, I transfer the attributes at the HRO level to the country level, which offers a convenient data set with the country-year unit of analysis for IR scholars.

Further research on *issue preference* could help tease out the patterns of HRO-country interactions. Especially in the current civil society repression (Carothers and Brechenmacher 2014; Glasius et al. 2020), figuring out the impact of issue preference on country behaviors could offer insightful suggestions for the survival of HROs.

The paper proceeds as follows. First, I will describe the data in the sample of HROs' shaming. After the explanatory analysis of the data, I will introduce the community detection method, the Leiden algorithm, and discuss the selection of the number of communities through subject-matter knowledge. Later, I use an example to illustrate how I apply community detection to create *issue preference*. Finally, I conduct external validity checks on the new measure, which can justify *issue preference* for future inquiries in IR literature.

CHAPTER 2

DATA DESCRIPTION

A typical approach to HROs' success is shaming or naming and shaming, where HROs rhetorically spotlight and spread governments' human rights violations (Meernik et al. 2012; Franklin 2008; Hafner-Burton 2008; Murdie and Davis 2012a). In this project, I will study *issue preference* in the case of HROs' shaming.

The sample includes five variables: *Year*, *Country*, *HRO*, *Goal*, and *Count*. The description of each variable is presented in **Table 1**. *HRO* and *Goal* are from the TSMO data, which include "organization-year data on (NGO) founding year; membership structure; headquarters location; goals; ties to intergovernmental organizations; ties to nongovernmental organizations; and country-location of members" (Smith et al. 2019). The TSMO data are coded based on the *Yearbook of International Organizations*, a credible but non-open resource for NGOs (UIA 2022).

The TSMO data include NGOs across different primary issue focuses or social movement industries such as various human rights, public health, miscellaneous conservative, world language, and animal rights (Smith et al. 2019: 10-11). Since I explore HROs in this project, I only include NGOs with the value "1" of *hrights*, a binary variable in the TSMO data to identify whether an NGO mainly works on human rights, 1 for HROs and 0 for non-HROs. For my final dataset, I gather 3,929 HROs from the TSMO data.

Table 1: Variable Description in HRO Shaming Sample

Variable	Description	Source
Year	Ranging from 2001 to 2012, indicating which year each HRO names and shames a country	-
Country	Country name, indicating a target country in a given shaming event	Created based on the *Lexis-Nexis* media database
HRO	Coded as *orgname* in the TSMO data, indicating the name of HRO shaming a country in a fixed year	TSMO data (Smith et al. 2019)
Goal	Coded as *goal 1* in the TSMO data, indicating an HRO's primary goal regarding human rights improvement	TSMO data (Smith et al. 2019)
Count	Continuous variable, indicating how many times an HRO name and shame a target country in a given year	Created based on the *Lexis-Nexis* media database

HRO, identified as *orgname* in the TSMO data*,* is a variable specifying the names of organizations, such as ActionAid, ANTI-APARTHEID MOVEMENT, Transparency International, etc., *Goal* indicates a critical attribute of HRO coded based on the aims statement of the organizations in TSMO data. The TSMO data include four goals for each NGO, coded from *goal 1* to *goal 4* (Smith et al. 2019: 5-10). Smith et al. (2019) use 108 values to indicate different goals of NGOs, such as "Local/regional authority initiatives, coordination," "Paradigmatic or system change/ Change in power distribution or culture (i.e., eco-centered culture)," and "Ethical Banking or Investment/Alternative financing arrangements."

I extract the first goal, *goal 1* coded in the TSMO data as one variable, *Goal*, to indicate HROs' attributes for the purpose of community detection. **Table 2** displays all the goals recorded in the variable *Goal* in the sample. Notably, the last value, listed in the last row of **Table 2**, is "general human rights" (shaded in blue). Compared to other values with specific aims, the goal of general human rights can hardly offer a specific issue area of HROs. Some leading HROs, such as AMNESTY INTERNATIONAL and Human Rights Watch, are marked as the ones with "general human rights." Since these leading HROs can be involved in different human rights-related issues with sufficient resources and expertise (Stroup and Wong 2017), it is challenging to use a specific issue to define their primary goals of human rights. Therefore, I exclude the HROs with the goal of "general human rights," since this goal can hardly indicate a specific preference for HROs' shaming. It, to some extent, reduces bias caused by the significant activeness of leading HROs in community detection. The unit of analysis for the cleaned 4041 observations is HRO-country-year.

Unfortunately, the dyadic relationship between HROs and countries has not been updated since the year 2001. To capture the frequency of HRO-to-country shaming, *Count* is a new variable created based on the TSMO data. *Count* is yielded based on the articles containing the names of 3,929 HROs in the TSMO data. And the coding materials are collected from the *Lexis-Nexis* media database, which includes "newspaper and magazine stories, transcripts of TV broadcasts, and summaries of public records fillings"[1].

To count the events of HROs' shaming as well as figure out the target countries from the database, the NLP (Natural Language Processing) technique (Braithwaite and Park 2019; Murdie et al. 2020) is used to extract the information framed as *who* did *what* to *whom*, *when*, and *where* from each article. Specifically, relying on the tools of Named Entity Recognizer (NER) to locate the HRO names and Basic Dependency from Stanford Core NLP to capture grammatical relations between words in a sentence (Manning et al. 2014), *Count* and *Country*, two key variables in the sample, are produced through this data mining process. *Country* demonstrates the target country an HRO is shaming in an article, while *Count* indicates the number of shaming events from an HRO listed in TSMO data to a target country in a given year.

[1] More information can be found at: https://multimedia.journalism.berkeley.edu/tutorials/lexis-nexis/#:~:text=LexisNexis%20is%20a%20huge%20electronic,summaries%20of%20public%20records%20filings.

Table 2: Values in the Variable *Goal*

#	Goal	#	Goal
1	Development/ empowerment/social justice	22	Trade in Hazardous products/ vs. abuse and/or use
2	Environmental protection general	23	Consumer rights/ protection (including product labeling)
3	Global economy/ economic justice/ fair trade/ NIEO/ climate justice	24	Regional government/ integration
4	Freedom of expression/ press	25	Public health/ health rights, equity
5	Women's rights/ empowerment	26	Anti-hunger/ food security
6	Anti-discrimination, minority rights, anti-racism	27	Vs. human trafficking, prostitution, sexual crimes
7	Children's rights/ welfare	28	Religious rights/ freedom
8	Democracy/Pro-Democracy & [civil society support/development]	29	Vs Capital Punishment
9	Electronic communication/access/rights to free expression/FOSS	30	Family planning/reproductive choice
10	Human rights particular group/ i.e. ethnic group, disabled	31	Punish crimes vs. humanity
11	Prisoners' rights/Police Accountability	32	Humanitarian assistance
12	Peace general-understanding, solidarity, communication, 'culture of peace', peacebuilding	33	Israel/ Zionist
13	Refugee rights/asylum/immigration	34	Solidarity
14	International political integration/ international law/ [pro-multilateralism]	35	International law/ standards/ justice/ Intl Criminal Court (promotion and implementation)
15	Housing rights	36	Palestinian liberation
16	Indigenous peoples' rights	37	Nationalism/ self-determination
17	Sustainable development	38	Anti-war/ conscientious objection
18	Environmental protection particular species	39	Environmental protection particular region/s or ecosystems
19	Cultural protection/preservation (including linguistic) [old code 56]	40	Economic cooperative/alternative econ structures/solidarity or social economy
20	Gay, lesbian, transgender rights	41	Anti-poverty/employment for poor; [vs. 'social exclusion']
21	Disarmament- vs military bases, weapons- nuke & conventional	42	Food sovereignty/peasant's rights
		43	Human rights general

CHAPTER 3

EXPLORATORY DATA ANALYSIS

Descriptive Statistics

Table 3 displays the descriptive statistics of the sub-sample by year. First, each sub-sample size, the aggregation of the dyadic interactions between HROs and countries, is larger than the sum of the number of total unique country names and the number of unique HRO names. It suggests that some countries are named and shamed by HROs more than once in a given year. It also echoes that some HROs are more active than others (Murdie and Davis 2012b; Stroup and Wong 2017).

Table 3: Descriptive Statistics of Yearly Sub-Sample

Year	Sample Size	# of Country	# of HRO	# of Goal	Count				
					Sum	Mean	St. Dev.	Min	Max
2001	170	58	73	42	300	1.765	2.455	1	22
2002	189	62	84	42	284	1.503	1.315	1	10
2003	249	70	95	42	394	1.582	1.476	1	14
2004	273	77	100	42	384	1.407	1.306	1	14
2005	285	91	98	42	419	1.470	1.143	1	10
2006	265	77	86	42	392	1.479	0.973	1	7
2007	294	86	101	42	458	1.558	1.140	1	11
2008	343	97	123	42	537	1.566	1.316	1	17
2009	397	88	130	42	629	1.584	1.581	1	18
2010	505	129	129	42	3330	6.594	35.999	1	702
2011	533	148	134	42	4812	9.028	55.583	1	1171
2012	538	132	131	42	2858	5.312	30.277	1	629
Total	4041	1115	1284	--	14797	--	--	--	--

In **Table 3**, since the number of *Goal* is constantly 42 across the years, HROs in all specific issue areas actively interact with countries each year. It infers that every year from 2001 to 2012, we can find shaming events across all the specific goals of HROs.

Figure 2 presents the number of NGOs and countries from 2011 to 2002. More countries, denoted in pink color, have been shamed over the years. Even though the number of target countries slightly decreased in 2006, 2009, and 2012, there is still a growing trend.

Indicated in blue color in **Figure 2**, the number of HROs also grown substantially from 2001 to 2012. Interestingly, from 2001 to 2009, HROs are more than countries each year. Starting from 2010, the number of HROs was not larger than that of countries each year. Given that the HRO number seemed not to grow, it infers that from 2010 to 2012, HROs, as a whole group, involved more countries in spotlighting the issues.

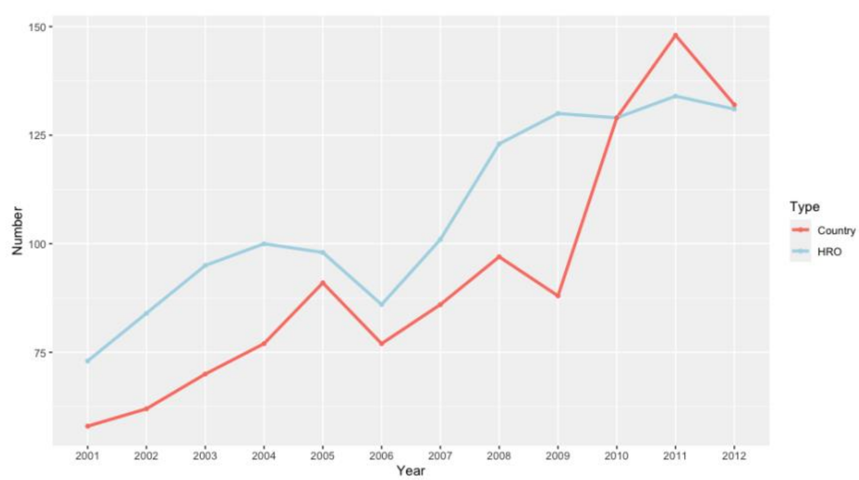

Figure 2: HROs and Countries Involved in Shaming Events by Year

In addition, I visualize the total number of shaming events by year in **Figure 3**. From 2001 to 2009, although the events kept increasing, compared to the HRO number and the country number, the increase was slight. A breakthrough appears in 2009. The number of shaming events rocketed to around 5000 in 2011 from less than 1000 in 2009. This drastic change is inconsistent with the number of HROs in **Figure 2**, where the HROs increase has stagnated since 2008. It suggests that HROs became more active in shaming through the growth of the event number between 2009 and 2011. The significant growth of users on social media, Twitter for instance[2], in 2010 could contribute to this change since online platforms make it easier for HROs to spread shaming. Given that more countries were targeted during the same period, it demonstrates that HROs were active in shaming countries by targeting more countries and initiating more shaming events.

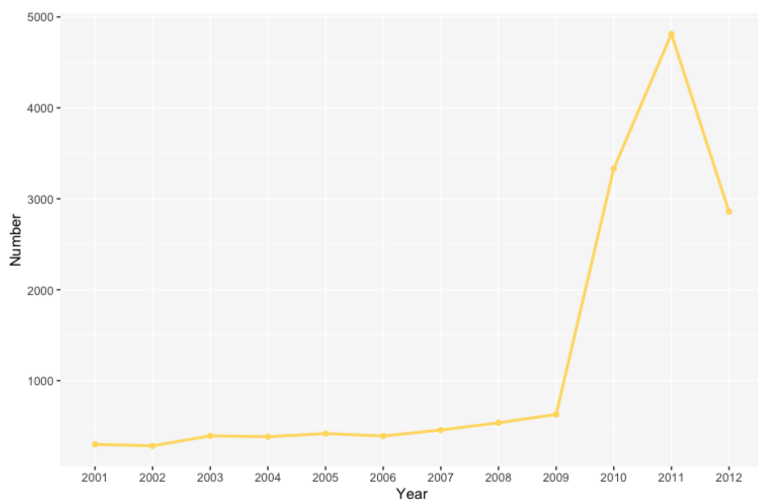

Figure 3: Total Count of Shaming Events by Year

Another interesting implication is that the number of shaming events plummeted in 2012 (see **Figure 3**). Compared to **Figure 2**, where fewer countries were targeted in the same year and the number of HROs was almost unchanged, HROs became less active regarding the frequency of shaming events and the number of target countries in 2012. This finding also makes sense

[2] More information can be found at: https://backlinko.com/twitter-users

since HROs tend to choose silence rather than radical behavior when civil society repression, which tends to target at HROs, increases over a threshold (Lian and Murdie 2013). Since the mid-2000s, there has been a global pushback to NGOs, where governments set legal barriers to restrain NGOs' activities (Carothers and Brechenmacher 2014). It is possible that the repression is cumulative enough to change HROs' attitudes.

Exploratory Data Analysis Regarding Issue Preference

Then, what can the data tell us about HROs' *issue preference* when shaming countries? **Figure 4** displays the aggregated number of reported HROs' shaming of governments by their specific goals. To put it another way, the number at the right end of each bar indicates how often the HROs with the goal denoted on the left side named and shamed countries from 2001 to 2012. For instance, the number 1041 corresponding to "Humanitarian assistance" means all the HROs with the goal of "Humanitarian assistance" shame countries 1401 times in total based on the sample.

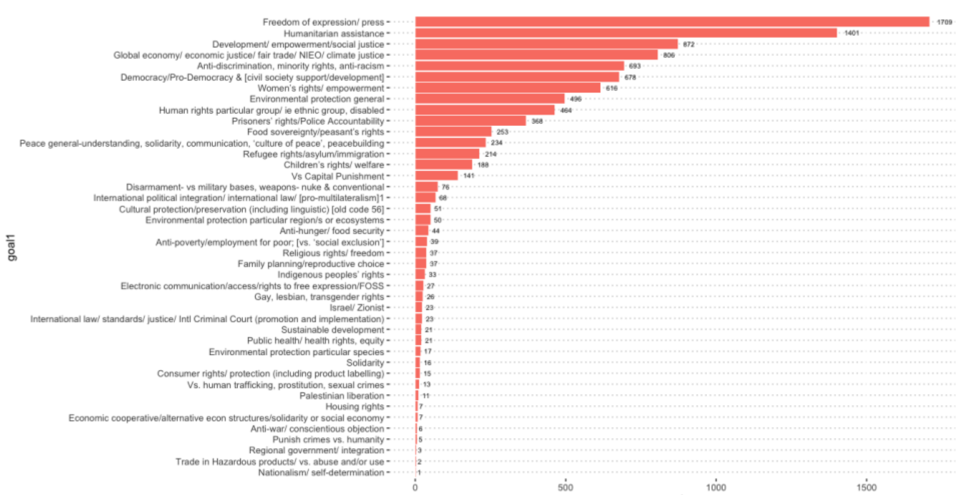

Figure 4: Goal Frequency of HROs, Whose Shaming on Countries is Reported in Media

According to **Figure 4**, on the one hand, HROs present various focuses on human rights, from freedom of expression/press (with the highest frequency of 1709) to nationalism/self-determination (with the lowest frequency of 1). On the other hand, some issues seem to be more prevalently addressed by HROs' shaming, such as freedom of expression/press, humanitarian assistance, development, etc.

The data raises numerous questions of interest. How do these HROs select the countries? What are the issues preferred by *HROs* to every country underlying the aggregated preference? Are HROs with the same focus, women's rights/empowerment, for example, more likely to name and shame certain countries in a given year?

I propose that HROs present different interests in countries. In other words, HROs tend to prioritize different issues to countries. I conceptualize the issues in which countries are more likely to be attacked by HROs as *issue preference*. Even within similar countries regarding political regime, geographical affinity, and the frequency of shaming, two countries, the US (see **Figure 5**) and Canada (see **Figure 6**), for instance, indicate different *issue preferences* of HROs based on the aggregated shaming events from 2001 to 2012.

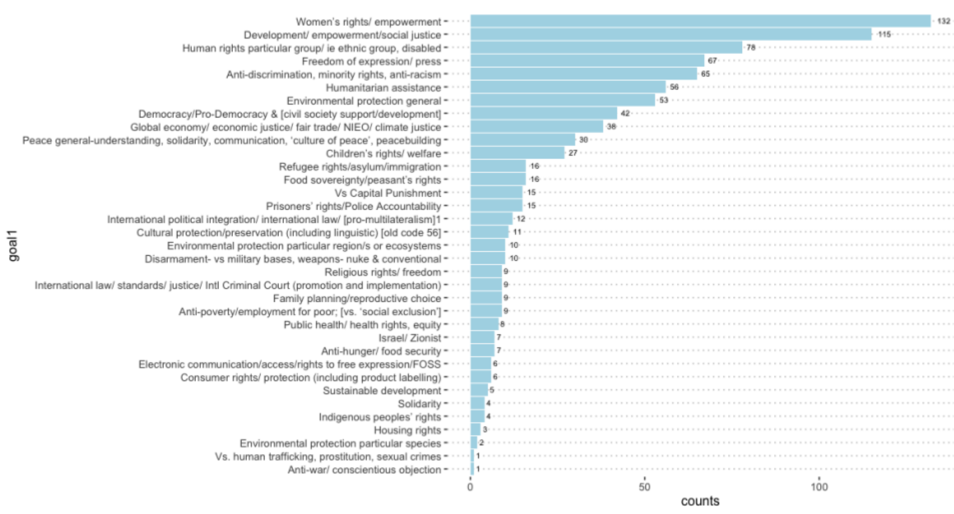

Figure 5: Shamed Frequency of US by HROs' Goals

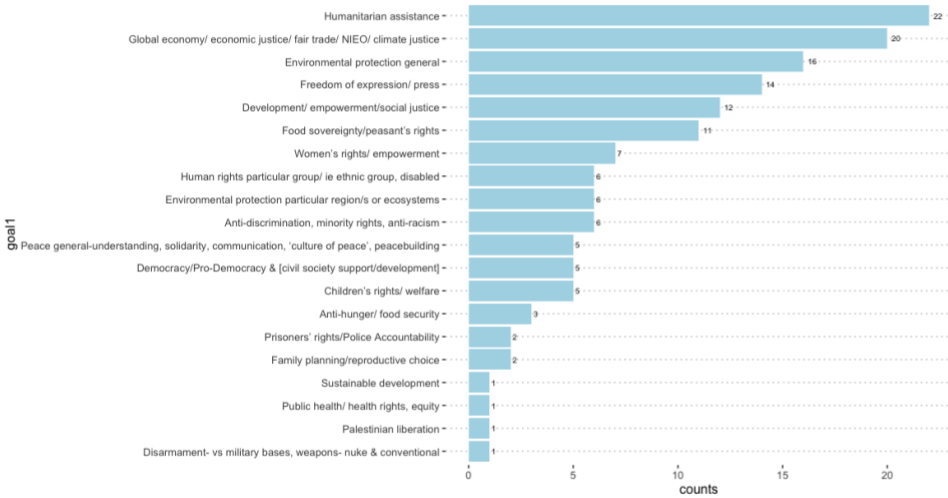

Figure 6: Shamed Frequency of Canada by HROs' Goals

Moreover, the scholarship suggests that HROs tend to self-select into scenarios where governments are more likely to violate human rights (Keck and Sikkink 1998). That explains why more repressive countries tend to become the targets named and shamed by HROs. However, according to **Figure 7**, it is puzzling that the top countries shamed by HROs reported in the 2001-2012 media include many consolidated democracies such as the US, the UK, Switzerland, Canada, France, and Australia. Meanwhile, some countries, almost "ignored" by the HROs in the 12 years, tended to demand international attention to local human rights violations. For instance, Jamaica, which was only mentioned once in the data set, was notorious for discrimination against LGBT groups in 2010.[3] What are *issue preferences* across countries? For a specific issue, which countries are more likely to be attacked by HROs in a given year?

[3] More information can be found at: https://www.amnestyusa.org/reports/annual-report-jamaica-2010/

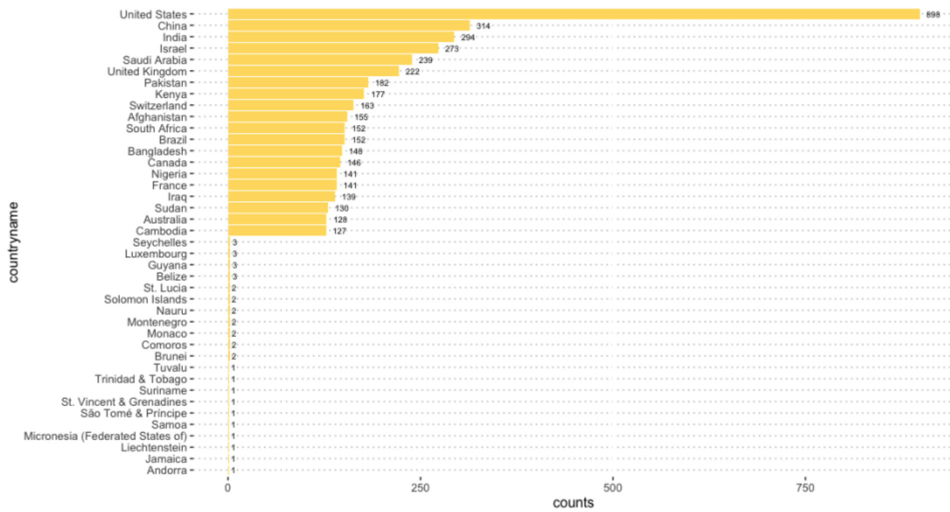

Figure 7: Shamed Frequency of Some Countries, Including the Top 20 and Last 21

To address the questions in the exploratory analysis, I will introduce the Leiden algorithm to cluster countries based on the issues preferred by HROs.

CHAPTER 4

COMMUNITY DETECTION

Network analysis has been used in IR research to address different questions (e.g., Kim 2023; Kinne 2013; Kinne 2018; Kinne and Maoz 2023), including the puzzles of NGOs (Hadden 2015; Murdie 2014; Murdie and Davis 2012a; Murdie and Polizzi 2016; Wilson et al. 2016). Community detection, a critical technique in network analysis, has been applied to IR for clustering (Cheng et al. 2021). However, there has not been much application for community detection in a bipartite network, where two different types of nodes form a network based on their interactions. One successful attempt is the application of lobbying in legislative politics by Kim and Kunisky (2021).

Analyzing data framed in a bipartite network is important in IR study. Critical actors in IR are defined at different levels: the international level, the national level, and the individual level. Conventionally, relying on hierarchical modeling can integrate the attributes of actors rather than the interactions between the actors at different levels. Unfortunately, the current technique used for community detection for a bipartite network is not convenient for IR scholars to apply to new research. For example, the Python package developed by Kim and Kunisky (2021) for bipartite network clustering can mainly be used for lobbying data rather than a general form of IR data. Also, Kim and Kunisky (2021) fail to specify how to “infer” the community numbers based on the literature. This is a fallout in applying network analysis to a bipartite network, which requires expertise to justify the community formation.

Therefore, in this project, I introduce the Leiden algorithm (Traag et al. 2019) to group the interactions between HROs and countries. The Leiden algorithm offers an option to detect communities in bipartite networked data. The algorithm has packages in R and Python softwares. The operational convenience through R is more accessible for IR scholars. In this project, I rely on the "leiden" package in R for community detection (Kelly 2022). In the following parts, I will specify community detection steps that are visualized in the flow chart in **Figure 8**.

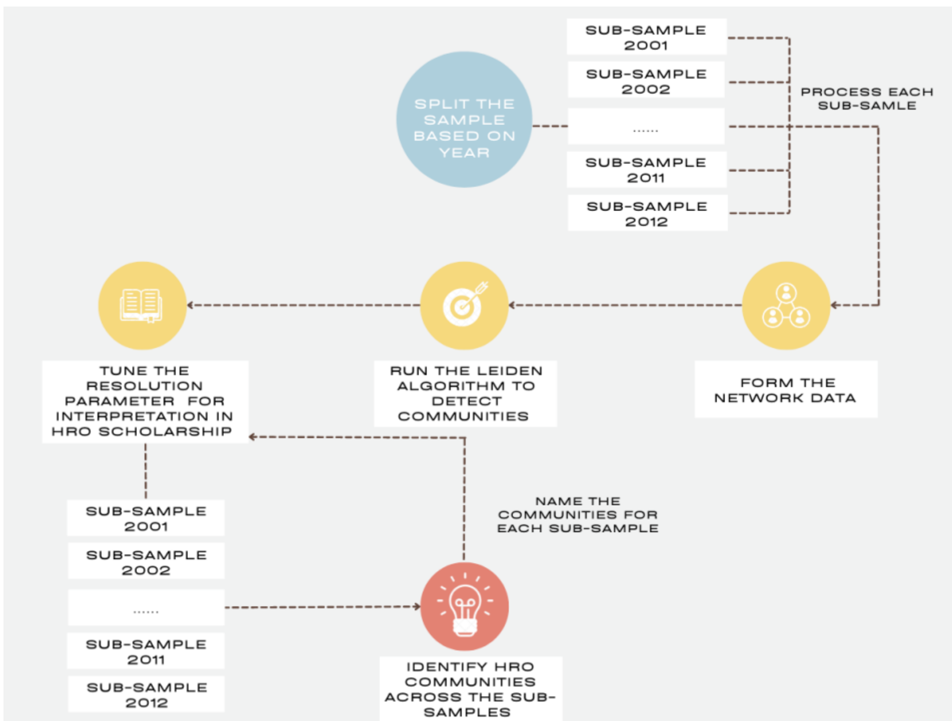

Figure 8: Flow Chart of Community Detection Steps

Form a Network Based on Yearly Subsample

I first to create a data set where the unit of analysis is country-year since this unit of analysis is extensively used in IR. On the one hand, I focus on countries since IR is a discipline built upon the assumption that countries are the central actors. Even nowadays, with the prosperity of studying different non-state actors, such as NGOs, MNCs (Multinational Corporates), and terrorist groups, countries are still important actors in IR scholarship. Also, the extensive study of countries means there are numerous available datasets at the country level. Therefore, country-level data can offer convenient data to empirically test the research involving state actors. On the other hand, the yearly dataset offers variation at the country level, which is always used by IR scholars to explore patterns of country performance across years.

To create *issue preference*, accordingly, I disaggregate the data based on year. Then, to detect the communities in each yearly sub-sample, I transfer HROs' shaming countries into a network. Since I identify HROs and countries as two different nodes, the data can form a bipartite network. For the edges, if one HRO names and shames a country, an edge connects the nodes. The weight of the edge is defined as the frequency of HROs' shaming. Notably, there is no edge within each type of node.

To illustrate how the network is identified in this case, I visualize the network based on a sub-sample, which includes some observations in the freedom of expression-focused issue community in 2005, in **Figure 9**. Each pink circle on the left indicates an HRO, which name is specified on the circle. I use blue squares on the right for countries whose names are marked on the squares. If a circle is connected with a square by a grey line, the edge, it means that an HRO shames the country. The thickness of the edges or the weight of the edges demonstrates how many times the countries are attacked by the HROs. The thicker an edge is, the more frequently the HRO name and shame the country. For instance, OXFAM's interaction with the US is more frequent than the one with Sri Lanka in **Figure 9**.

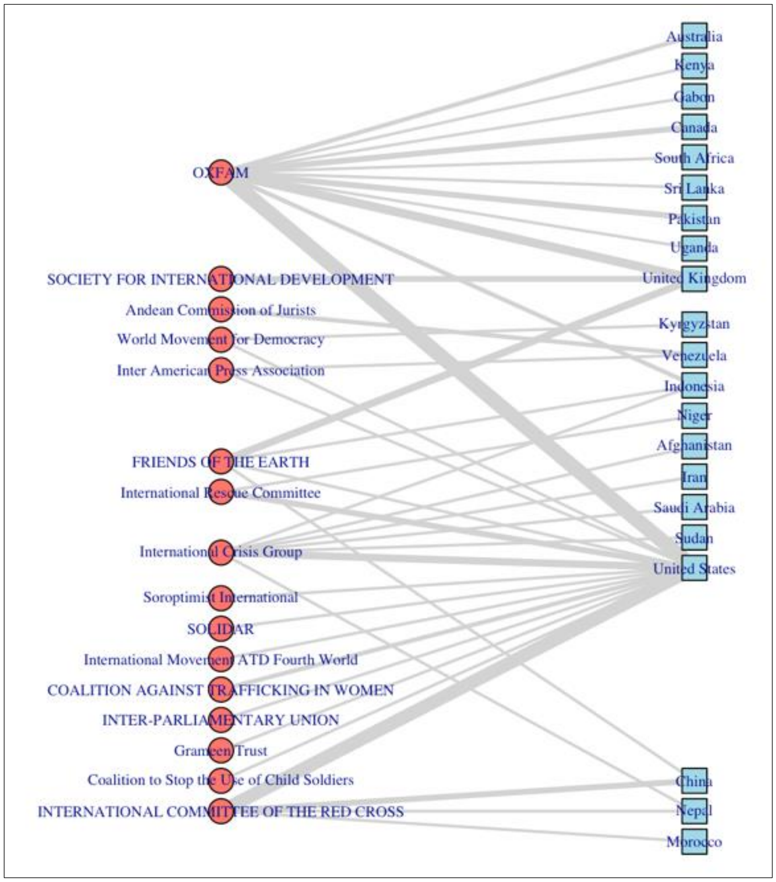

Figure 9: An Example of the Bipartite Network Based on Some Observations in the Freedom of Expression-Focused Issue Community in 2005 Detected by the Leiden Algorithm

Community Detection Algorithm

After modeling the network, I will run the community detection algorithm on each sub-sample disaggregated by year, respectively. The justification of yearly data, in this case, comes from the following aspects. First, the IR scholarship suggests that HROs tend to update their target countries. As INGOs, HROs' transnational activities could be affected by funding (Cooley and Ron 2002) and the pursuit of global fame (Stroup and Wong 2017). Both financial assistance to HROs and international influence can be affected by hot buttons, which could diffuse over time. For instance, the MeToo movement can gather the attention of the HROs focusing on women's rights to the problem in the US. After a while, the global diffusion could make women's rights become a new hot button in a geopolitical country since human rights movements tend to present the spillover effect in neighboring countries (Lee and Murdie 2021).

In the Leiden algorithm, I will optimize modularity, an indicator of node density within and between communities. First, the function to produce modularity is:

$$\frac{1}{2m}\sum_{ij}(A_{ij} - \gamma n_i n_j)\delta(\sigma_i, \sigma_j),$$

where m is the total edge weight, A_{ij} is the weight of edge $(i, j), \gamma$ is the resolution parameter, n_i is the node weight of node i, σ_i is the cluster of node i and $\delta(x, y) = 1$ if and only if $x = y$ and 0 otherwise. By setting $n_i = k_i$, the degree of node i, and dividing γ by $2m$, I can gain an expression for modularity (Traag et al. 2019). Correspondingly, in R software, I need to input the data in the network form, specify the weights of nodes, and tune the resolution parameter to yield the modularity in the Leiden function.

The next step is to optimize the modularity by selecting a method in the Leiden algorithm. Conventionally, the Louvain algorithm, which is another approach for community detection, forms some partitions by locally moving the nodes which are assumed as communities of one at the beginning. Then, the Louvain algorithm will directly aggregate the locally formed partitions (Blondel et al. 2008). It makes that some communities are internally poorly connected. It is because some bridging nodes for two sub-communities can be moved out of the community. The improvement of the Leiden algorithm is to add one step called refinement of partitions. It can guarantee that the bridging nodes stay within a community to well connect the sub-communities. In sum, the Leiden algorithm (1) locally creates partitions, (2) refine the partitions by ensuring well-connection within communities, and (3) aggregates partitions to detect communities.

In this project, I use the modularity vertex partition as the function of optimization to explore the community structure, which can be achieved by setting "ModularityVertexPartition.Bipartite" as the partition type in R for the bipartite network in this project. This method can achieve two expected goals in community detection – positive values of the edge weights and the optimization of the modularity. First, I need to define the edge weights with positive rather than negative values since the frequency of HRO-to-country interactions in the research is larger than zero. Second, I expect denser node connections with communities but sparse node connections between modules, which is indicated by high modularity in networks (Newman 2006). Therefore, I'd like to maximize the modularity, which can boost the strength of community detection.[4]

However, the Leiden algorithm can produce different partitions with significance. Once the resolution parameter changes, the number of communities could vary, or the nodes in some communities would flow out to other communities. In other words, although different resolution parameters could lead to the same classification of communities, there can be many options of node classification. By implication, instead of selecting an "optimal" partition, I attempt to find an acceptable number of communities by relying further on subject-matter input.

Two Strategies for Tuning the Resolution Parameter

The key step to an acceptable Leiden partition is to select an appropriate (gamma) resolution parameter. The selection, on the one hand, needs to demonstrate the statistical

[4] More information can be found at: https://leidenalg.readthedocs.io/en/stable/reference.html?highlight=modularity%20vertex#modularityvertexpartition

significance of partitions. On the other hand, given the multiple significant partitions, integrating contextual knowledge can help better pick up a partition interpretable in the literature.

Since I aim to create a categorical measure, I would like to create 5 to 7 communities, which would be interpretable (Kim and Kunisky 2021), across all the sub-samples. Therefore, I make two strategies for tuning the resolution parameter in each sub-sample when considering statistical modeling.

Strategy 1. I will detect communities of more than 7 in each sub-sample to pool community types for selection. I usually start tuning the resolution parameters for communities from 7 to 10. For the same number of communities, I would also try partitions for interpretation since the consisting of nodes in some communities could change in a given partition. I start examining 7 communities since small number of communities could contain more NGOs within a community, which could make it less interpretable. On the other hand, for large number of communities, I can merge some communities with same issues.

Notably, if it is hard to figure out the parameter within 10 communities regarding the subject matter expertise, I will then try the range of 11 to 20 communities. In other words, my approach is to start from smaller clusterings, which will make it easier to check the interpretation of clustering in IR. For instance, in **Figure 10**, the smallest community number is 10. Therefore, after knowing that 10 communities are interpretable in IR in this case, I will tune the parameter to form 14 communities, falling in the range of 11 to 20 communities.

Strategy 2. The increase in community numbers will lead to some small communities. In thissituationn, I need to guarantee that the smallest community need to have at least two nodes since the Leiden algorithm can ensure that the nodes in a community of two will come from different types. In this case, if a community contains two nodes, one must be an HRO and the other a country, which will capture the HRO-country interaction within each community. However, the Leiden algorithm can also produce a community of one, either an HRO or a country. Since a community of one is uninterpretable for *issue preference*, I set the smallest community to include at least two nodes.

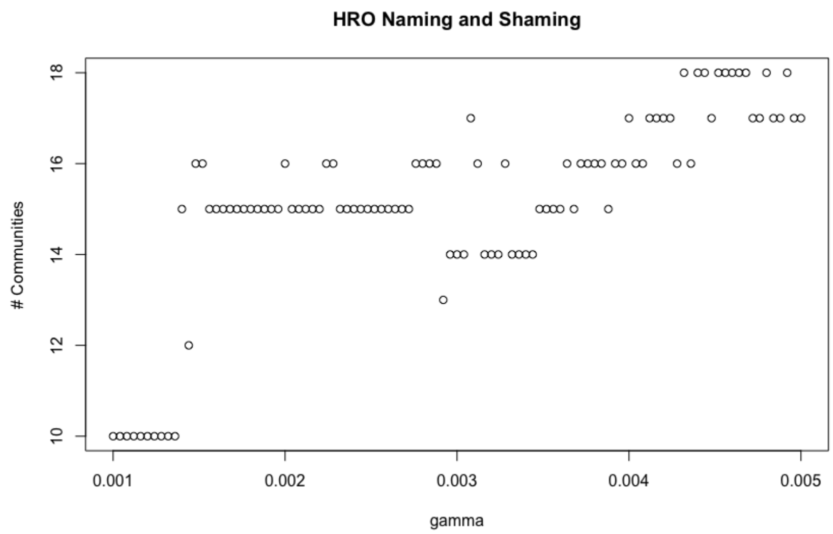

Figure 10: Selected Range of Resolution Parameter in the Sub-Sample 2006

Integrating Literature: The "Dominance-Relevance" Standard

Given that many partitions are statistically significant, I need to further use the IR literature to select the acceptable one. I conceptualize my approach to integrating expertise into community detection as the "dominance-relevant" standard.

In each community, there are different HROs regarding specific goals in human rights improvement. I count the number of HROs with an identical goal within each community. To put it another way, I identify the goals as preferred issues and the number of HROs with an identical goal as issue frequency to the countries within the same community. In **Figure 11**, for example, which displays one detected community based on the sample of 2003, two different HROs aiming for “freedom of expression/press” are present in this community. It suggests that the countries in this community are preferred by this type of HRO.

Since one community can contain many HROs with varying frequencies of issues in shaming countries, it is challenging to identify one *issue preference* for each community. My solution is the “dominance-relevant” standard. Based on this standard, I first figure out the dominant issues preferred by HROs. Still, in **Figure 11**, it is easy to find that “women’s rights/empower” HROs name and shame most in the community. Then, I will check the relevance of other issues to ensure that most of the issues in one community are highly relevant based on the literature. In other words, most of the other HROs’ goals must be connected to women’s rights. It makes sense to discuss women’s rights in the framework of most other issues presented in **Figure 11**, such as “refugee rights/asylum/immigration” and “human rights particular group/ ie ethnic group, disabled.”

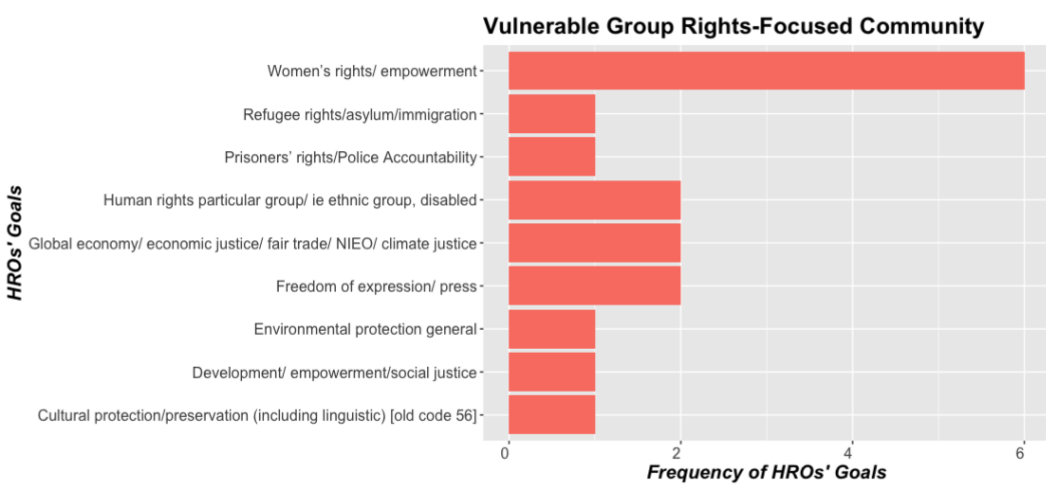

Figure 11: Distribution of Issue Frequency Focused by the HROs’ Shaming Within the Vulnerable Group-Focused Community in 2003

Since the issue of women's rights is classified into the first community in **Table 4**, I will name this community "Vulnerable Group Rights-Focused Community." Correspondingly, the countries clustering in this community will be coded as "vulnerable group rights" under the variable of *issue preference*. More discussion about community names will be found later.

However, it can be challenging to figure out which type of HRO is dominantly active in one community. Also, it could happen that most HROs' goals in one community are not highly relevant based on the literature, especially when many different types of HROs are classified into one cluster. My solution is to tune the resolution parameter of the Leiden algorithm to reduce the nodes in the communities or increase the community numbers. In most cases, it will make it easy to point out the issue appearing most in the new community. If the dominant issue can still hardly be identified, I will further guarantee that most HROs in one community are highly related based on the literature. In **Figure 12**, I present a typical case where all the issues can be discussed in the community of vulnerable group rights defined in **Table 4**.

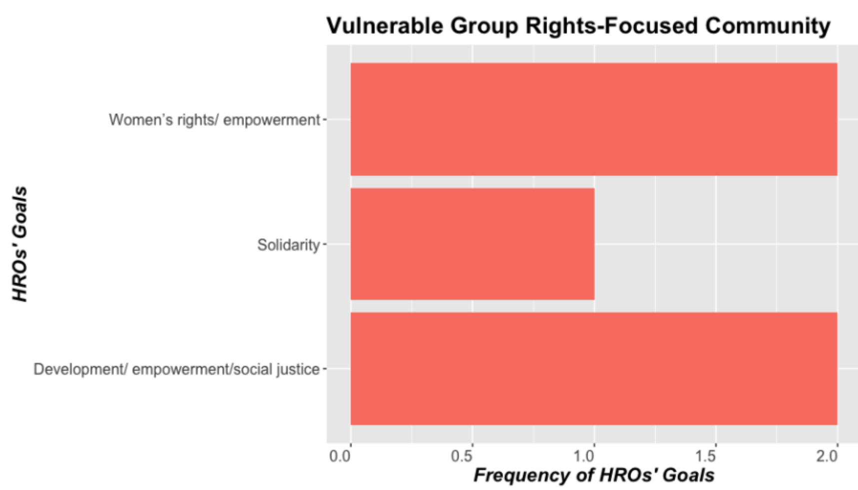

Figure 12: Distribution of Issue Frequency Focused by the HROs' Shaming Within the Peace-Focused Community in 2009

Creating Issue Preference

By tuning the resolution parameter for each yearly sample[5] and integrating the literature on HROs, I first gain 7 issues focused on by the communities across all the sub-samples. They are vulnerable group rights, peace, sustainable development, minority rights, democracy, freedom of expression, and economy in **Table 4**.

Table 4: The Pool of Community Type Based on HROs' Goals

Type	Community	Goals
1	Vulnerable Group Rights	Women's Rights, Family Planning, Children's Rights
2	Peace	Disarmament, Solidarity, Humanitarian Assistance, Refugee Rights, Peacebuilding
3	Sustainable Development	Environmental Protection, Sustainable Development, Food Security
4	Minority Rights	Religious Rights, Ethnic Groups, Indigenous People, the Disable, LGBT Rights, Anti-Discrimination, Prisoner's Rights, Peasant's Rights
5	Democracy	International Law, Democracy, Nationalism
6	Freedom of Expression	Freedom of Expression
7	Economy	Economic Justice, Social Economy, Global Economy

However, I further merge the issue community of freedom of expression to that of democracy. I also include the community of economy to that of sustainable development. The updated community types are vulnerable group rights, peace, sustainable development, minority rights, and democracy (See **Table 5**).

[5] More information can be found at: https://people.duke.edu/~jmoody77/snh/2021/CommunitiesSNH2021.nb.html

Table 5: Final Community Type Based on HROs' Goals

Type	Community	Goals
1	Vulnerable Group Rights	Women's Rights, Family Planning, Children's Rights
2	Peace	Disarmament, Solidarity, Humanitarian Assistance, Refugee Rights, Peacebuilding
3	Sustainable Development	Environmental Protection, Sustainable Development, Food Security, Economic Justice, Social Economy, Global Economy
4	Minority Rights	Religious Rights, Ethnic Groups, Indigenous People, the Disable, LGBT Rights, Anti-Discrimination, Prisoner's Rights, Peasant's Rights
5	Democracy	International Law, Democracy, Nationalism, Freedom of Expression

The justification of the combination is based on the IR literature, where the differences between freedom of expression and democracy and between economy to sustainable development are not significant since freedom of expression is one indicator of democracy and economy issues are included in sustainable development. Moreover, compared to a partition with 7 communities, five communities can make each community miss fewer issues (See **Table 6**). In addition, the justification of the 5 final issue communities is also demonstrated in the external validity checks. I will discuss the legitimacy in the external check section.

Table 6: Communities Within Each Sub-Sample

Year	# of Commu-nity	Resolution Parameter	Community Type				
			Vulnerable Group Rights	Peace	Sustainable Development	Minority Rights	Democracy
2001	9	0.0043	✓	✓	✓	✓	✓
2002	19	0.0035	✓	✓	✓	✓	✓
2003	10	0.0021	✓		✓	✓	
2004	9	0.002	✓		✓	✓	✓
2005	16	0.0035	✓	✓	✓	✓	✓
2006	17	0.005	✓		✓	✓	✓
2007	13	0.003	✓		✓	✓	✓
2008	11	0.0023	✓	✓	✓	✓	✓
2009	13	0.0022	✓	✓	✓	✓	✓
2010	16	0.0025	✓		✓	✓	✓
2011	8	0.00211	✓	✓		✓	✓
2012	19	0.00254	✓	✓	✓	✓	✓

I will use IR expertise to further demonstrate the internal validity of the 5 detected communities. First, the distribution of HROs' goals in each community is interpretable, which is consistent with the "dominant-relevant" standard. I will use IR literature to check whether most issues within the same community can be framed into the issue defining the community type. Second, the ability of the community detection algorithm to correctly identify the community for leading HROs implies its legitimacy. According to the IR scholarship, to demonstrate moderation towards states, the leading HROs tend to fall in the communities less likely to irritate governments (Stroup and Wong 2017). In the project, for instance, Transparency International, an HRO advocating against corruption, is classified into the community of sustainable development rather than the rights-related communities, the more radical issues against countries. Third, I will also examine whether countries in a community face the issue used to the community type.

After finalizing the five communities based on vertex modularity across all the years, I then go back to the communities in each sub-sample to tag the communities, including both HROs and countries, with the 5 community names in **Table 5**. Notably, the community names can be repeated in a sub-sample. For instance, in a sample divided into 10 communities, it is possible that there are five communities named as vulnerable groups' rights. I note that there are instances where certain communities do not appear in a sub-sample (e.g., the peace community in 2003). The community types within each sub-sample can be seen in **Table 6**.

CHAPTER 5

AN EXAMPLE: COMMUNITY DETECTION BASED ON THE SUB-SAMPLE IN 2001

To better explain how I create *issue preference* by combining the Leiden algorithm and IR literature, I will use the sub-sample in 2001 to illustrate the community detection. I take this sub-sample as an example is since the typical problems can be found in this case. Also, according to **Table 5**, the five communities can all be detected in this sub-sample. It can present how to summarize the five communities. The descriptive statistics of the 2001 sub-sample are in **Table 3**.

Community Detection

Figure 13 displays the number of communities based on different values of gamma, the resolution parameter. When the gamma is smaller than 0.006, the community number ranges from 4 to 12. The plot fails to indicate which community partition I should prioritize. Therefore, I have to compare different Leiden partitions of 7 to 10 communities first. Also, these partitions tend to appear more in the plot, indicated by the number of dots in a horizontal line. Then, I will use the expertise of HROs to check each community under the partitions of 7, 8, 9, and 10 communities.

Worth noting, even with an identical partition, for instance, 9 communities, when slightly changing the parameters, each community could contain different nodes. For instance, in the sub-sample of 2001, when the gamma is 0.0043, there are 44 nodes in the first detected community and 26 nodes in the third community. However, when tuning the parameter to 0.0041, given the same number of communities (9 communities) denoted in **Figure 13**, the consisting of the two communities changes – 6 nodes from the 3rd community move out to the 1st detected community while other communities stay the same.

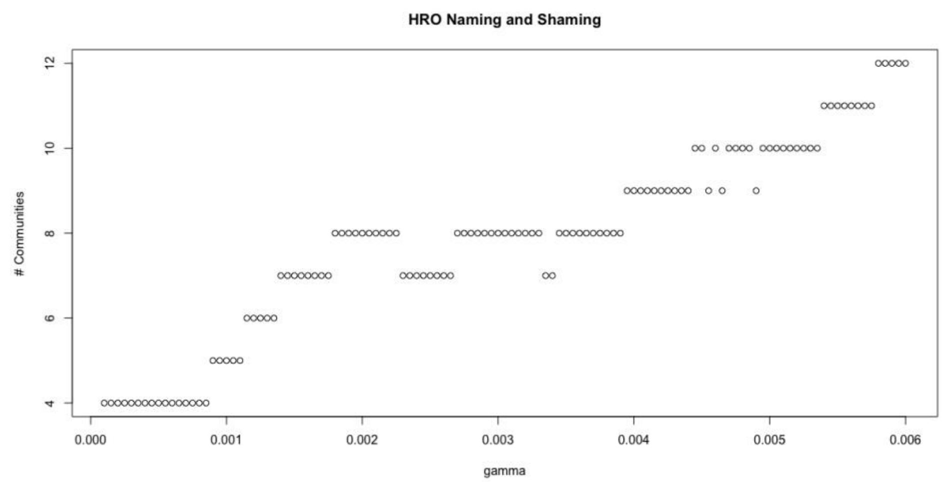

Figure 13: Number of Detected Communities Across Different Values of Gamma Resolution Parameter in 2001

Why Exclude Other Leiden Clusterings

After figuring out the range of Leiden partitions, the selection of the resolution parameter hinges on the "dominance-relevance" standard. I will use an example to illustrate how I exclude some Leiden partitions.

Figure 14 displays the 3rd community detected in the sub-sample of 2001 when setting the resolution parameter to 0.0045. It fails to satisfy the "dominance-relevance" standard within the community since even the three dominant issues can hardly be framed into an issue defining the community type. The sensitivity of Leiden partitions highlights the importance of subject-matter input.

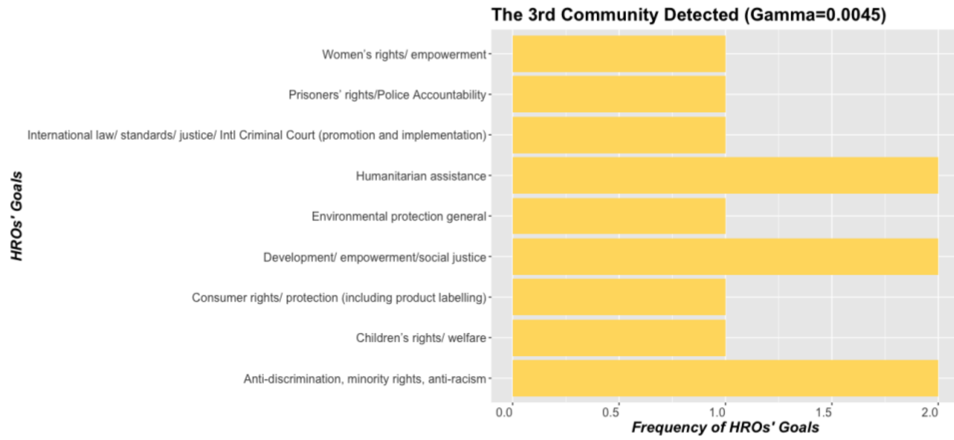

Figure 14: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Sustainable Development-Focused Community (the 3rd Community Detected with Resolution Parameter 0.0045 in 2001)

Final Outcome

I go through HROs' goals in each community within each Leiden partition. Based on the "dominance-relevance" standard, I prefer 9 communities for the sub-sample in 2001 with a resolution parameter of 0.0043. **Figures 15** to **23** display the goal frequencies of all the HROs and country names for the 9 communities. The visuals help me identify community types in the sub-sample of 2001.

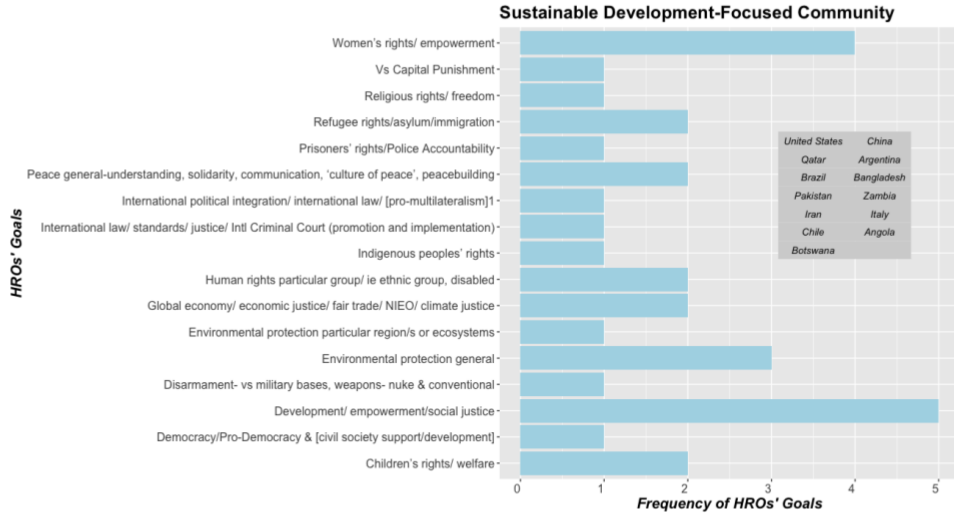

Figure 15: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Sustainable Development-Focused Community (the 1st Community Detected with Resolution Parameter 0.0043 in 2001)

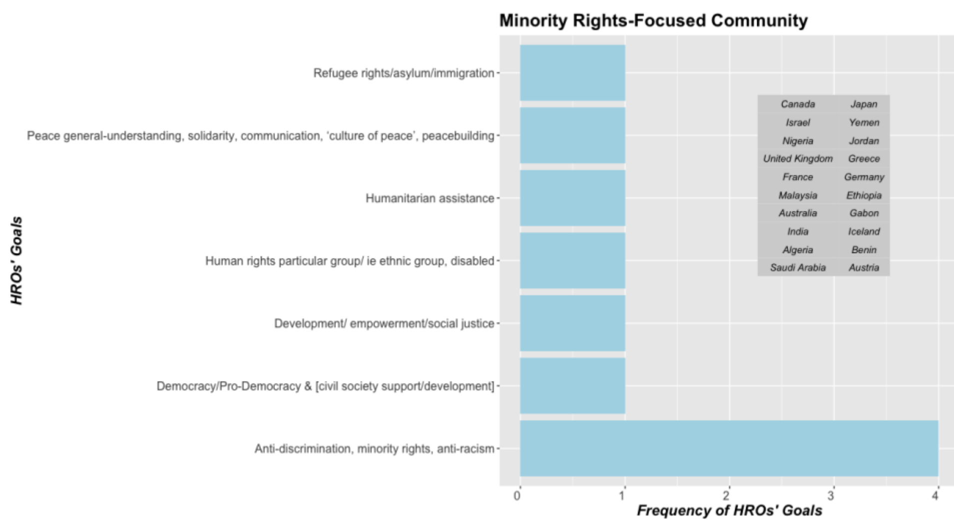

Figure 16: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Minority Rights-Focused Community (the 2nd Community Detected with Resolution Parameter 0.0043 in 2001)

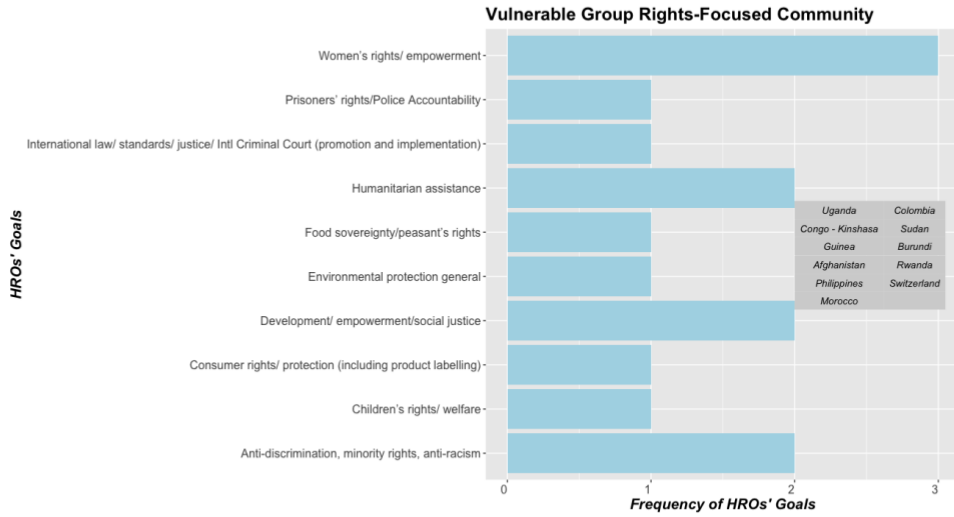

Figure 17: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Vulnerable Group Rights-Focused Community (the 3rd Community Detected with Resolution Parameter 0.0043 in 2001)

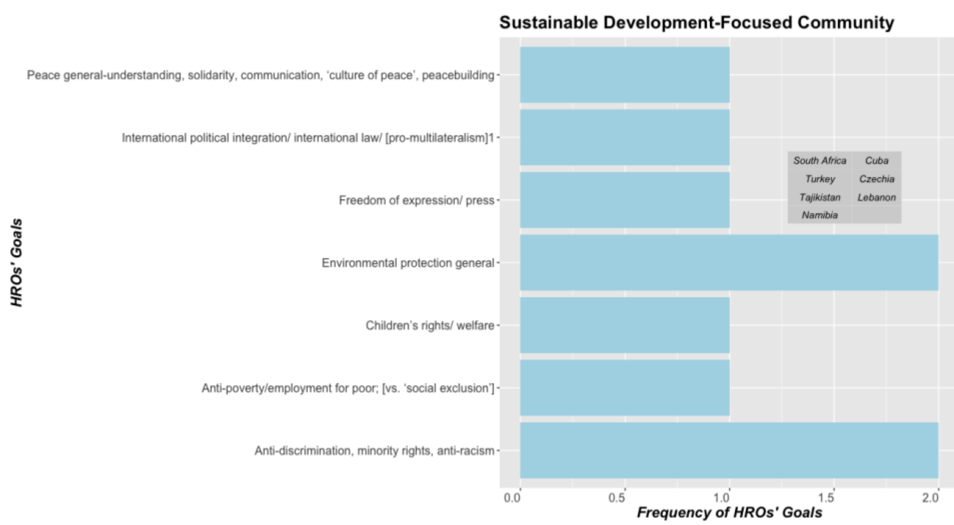

Figure 18: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Sustainable Development-Focused Community (the 4th Community Detected with Resolution Parameter 0.0043 in 2001)

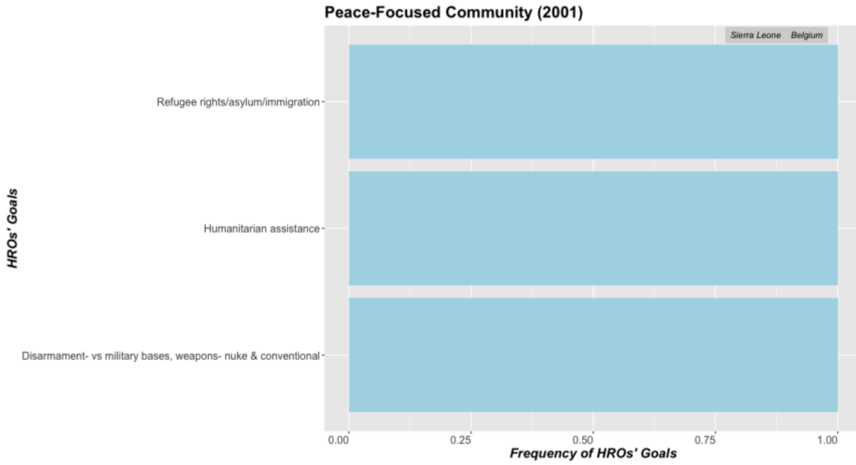

Figure 19: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Peace-Focused Community (the 5th Community Detected with Resolution Parameter 0.0043 in 2001)

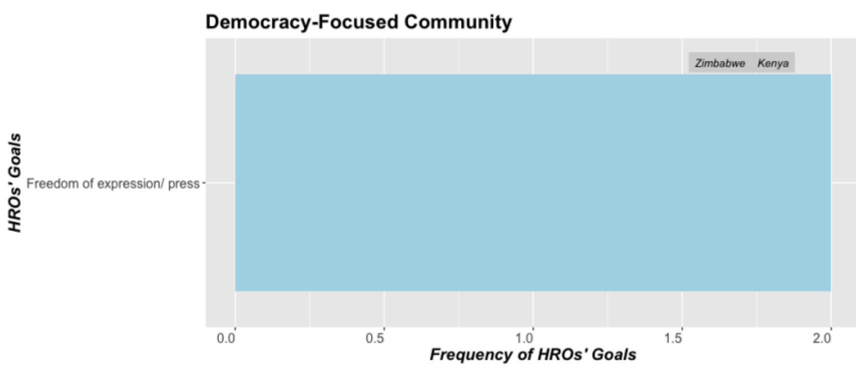

Figure 20: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Democracy-Focused Community (the 6th Community Detected with Resolution Parameter 0.0043 in 2001)

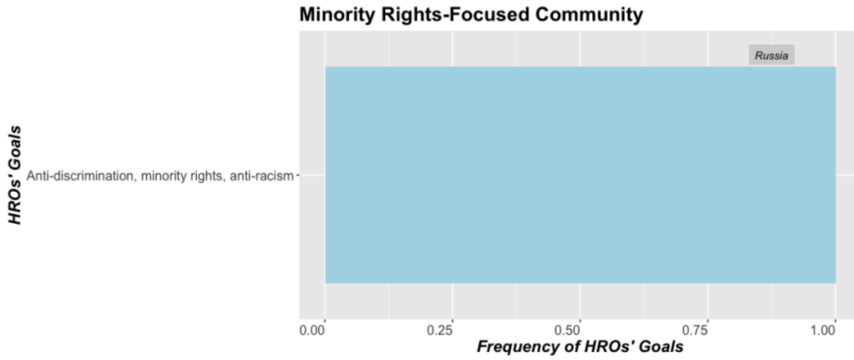

Figure 21: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Minority Rights-Focused Community (the 7th Community with Resolution Parameter 0.0043 Detected in 2001)

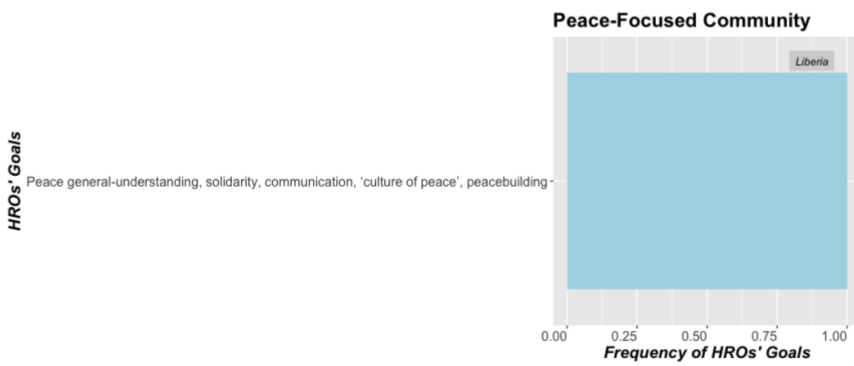

Figure 22: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Peace-Focused Community (the 8th Community with Resolution Parameter 0.0043 Detected in 2001)

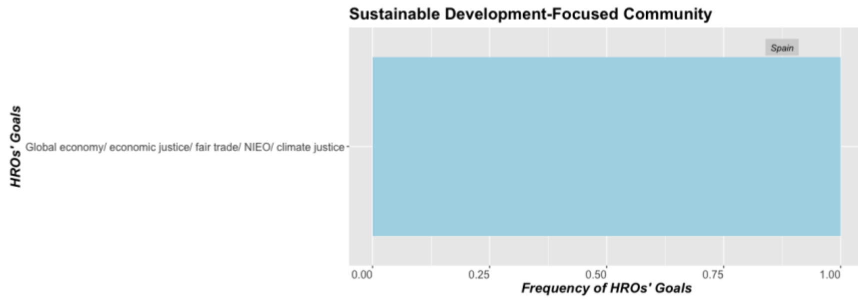

Figure 23: Distribution of Issue Frequency Focused by the HROs and the Target Countries Within the Sustainable Development-Focused Community (the 9th Community with Resolution Parameter 0.0043 Detected in 2001)

The summary statistics of the accepted partition are presented in **Table 7**. It suggests the internal validity of the partition in 9 communities. First, each community type can be interpreted with the issue highlighted by HROs in countries with the same community. For instance, it seems to be contentious that Sierra Leone and Belgium are located in the same community focusing on peace (see the rows shaded in pink in **Table 7**). Since HROs can raise different concerns regarding peace, where Sierra Leone could contribute to peace within its region, while Belgium could better assist in peace issues outside the borders, for instance, it makes sense to include these two countries in the same community.

Second, the pattern of the leading HROs echoes the IR literature. On the one hand, leading HROs tend to moderate their interactions with countries (Stroup and Wong 2017). In the first detected community (highlighted in blue in **Table 7**), leading HROs, such as OXFAM and FRIENDS OF THE EARTH, tend to be clustered in the issue of less political salience, sustainable development. On the other hand, leading HROs are more likely to attract other HROs' support, which is caused by leading HROs' sufficient resources for public exposure and organizing expertise (Stroup and Wong 2017). This explains why leading HROs are in the community with most HROs.

Third, community types could explain why some countries perform counterintuitively in their interaction with HROs. For instance, in the current "closing space," a global phenomenon where countries use administrative tools to repress NGOs (Carothers and Brechenmacher 2014; Clasius et al. 2020), some democracies and the countries with significant democratic attributes narrow down the working space of HROs to varying extent. The community focused on minority rights in yellow in **Table 7** could offer an explanation. As a politically salient issue, minority rights are more likely to challenge a government's legitimacy and further could lead to more repression of HROs.

Table 7: Summary Statistics of HROs and Target Countries in Each Issue-Focused Community Based on the Sub-Sample of 2001

	Community Type	HRO		Country	
		#	Example	#	Example
1	Sustainable Development	31	OXFAM	13	United States
			FRIENDS OF THE EARTH		China
			Network Women in Development Europe		Brazil
2	Minority Rights	10	Migrants Rights International	20	India
			INDIGENOUS AFFAIRS		Canada
			International League Against Racism and Anti-Semitism		Israel
3	Vulnerable Group Rights	15	INTERNATIONAL COMMITTEE OF THE RED CROSS	11	Afghanistan
			International Save the Children Alliance		Rwanda
			Catholic Relief Services		Uganda
4	Sustainable Development	9	Women's Environment and Development organization	7	South Africa
			Franciscans International		Cuba
			Practical Action		Turkey
5	Peace	3	Coalition to Stop the Use of Child Soldiers	2	Sierra Leone
			Migrante International		Belgium
			International Crisis Group		
6	Democracy	2	World Press Freedom Committee	2	Zimbabwe
			INTERNATIONAL FEDERATION OF JOURNALISTS		Kenya
7	Minority Rights	1	INTERNATIONAL LESBIAN AND GAY ASSOCIATION	1	Russia
8	Peace	1	Mano River Women Peace Network	1	Liberia
9	Sustainable Development	1	Clean Clothes Campaign	1	Spain

CHAPTER 6

VALIDITY CHECK

Since the number of communities is not fixed and interpretations might lead to different acceptable classifications, more checks on the indicator's validity are needed. After checking the internal validity of partitions within each sub-sample as I did in the example of 2001, I will rely on the IR literature to demonstrate that the classification in this project is reasonable.

The IR (International Relations) scholarship assumes that countries are more likely to increase civil society repression, a type of repression against civil society organizations (CSOs), if the legitimacy is challenged (Dupuy et al. 2016; Kiyani and Murdie 2020). Therefore, I contend that HROs' focus on politically salient issues, such as rights and democracy, is more likely to increase the repression by irritating the governments.

I use two different response variables, both approximations of civil society repression from the database of Varieties of Democracy (V-Dem) (Coppedge et al. 2020), to model the effects of HROs' *issue preference* on countries from 2000 to 2012. The response variables are *CSO Repression* (v2csreprss) and *CSO Entry and Exit* (v2cseeorgs) in Models 1 and 2, respectively. *CSO Repression* addresses whether "the government attempts to repress civil society organizations (CSOs)." And *CSO Entry and Exit* is coded based on the answers to "to what extent does the government achieve control over entry and exit by civil society organizations (CSO) into public life?" After I inverted the order of both variables in the original dataset, a higher value in a response variable suggests more government repression in the models. The response variables are also led for one year to eliminate the confounding effects caused by the predictors before the practices in the same country-year.

The key predictor in each model is the measure of HROs' *issue preference*. I set sustainable development, the least politically salient issue in this case, as the baseline of the variable. In Models 1 and 2, I use *issue preference* with 5 communities, the measure I create in this project. I also use issue preference with 7 communities, the original one produced in community detection in Models 3 and 4. I expect to compare the two groups of models (Models 1&2 vs. Models 3&4) regarding different measures of *issue preference* to demonstrate that the measure with 5 communities is more externally valid.

The confounders included in the models are *NGO-to-Government Conflicts*, *Neighborhood Effects*, and *HR Protection Score*, based on the IR literature. The specific information on the controls and the variables discussed above can be found in **Table 8**.

Therefore, the models can be expressed as:

$$Civil\ Society\ Repression_{i,(t+1)} = \alpha + \beta_1 Issue\ Preference_{i,t} + \beta_2 NGO - to - Government\ Conflicts_{i,t} + \beta_3 Neighborhood\ Effects_{i,t} + \beta_4 HR\ Protection\ Score_{i,t} + \varepsilon_{i,t},$$

where the response variable will be updated with different approximations, *CSO Repression* and *CSO Entry and Exit* to form different models. Also, *Issue Preference* in Models 1 and 2 are the ones with 5 values. In Models 3 and 4, Issue Preference include 7 values.

I list the four models in **Figure 24**, I highlight the response variables in pink and key independent variables in blue.

Model 1:

$$CSO\ Repression_{i,(t+1)} = \alpha + \beta_1 Issue\ Preference_{i,t}(5\ communities) + \beta_2 NGO - to - Government\ Conflicts_{i,t} + \beta_3 Neighborhood\ Effects_{i,t} + \beta_4 HR\ Protection\ Score_{i,t} + \varepsilon_{i,t},$$

Model 2:

$$CSO\ Entry\ and\ Exit_{i,(t+1)} = \alpha + \beta_1 Issue\ Preference_{i,t}(5\ communities) + \beta_2 NGO - to - Government\ Conflicts_{i,t} + \beta_3 Neighborhood\ Effects_{i,t} + \beta_4 HR\ Protection\ Score_{i,t} + \varepsilon_{i,t},$$

Model 3:

$$CSO\ Repression_{i,(t+1)} = \alpha + \beta_1 Issue\ Preference_{i,t}(7\ communities) + \beta_2 NGO - to - Government\ Conflicts_{i,t} + \beta_3 Neighborhood\ Effects_{i,t} + \beta_4 HR\ Protection\ Score_{i,t} + \varepsilon_{i,t},$$

Model 4:

$$CSO\ Entry\ and\ Exit_{i,(t+1)} = \alpha + \beta_1 Issue\ Preference_{i,t}(7\ communities) + \beta_2 NGO - to - Government\ Conflicts_{i,t} + \beta_3 Neighborhood\ Effects_{i,t} + \beta_4 HR\ Protection\ Score_{i,t} + \varepsilon_{i,t},$$

Figure 24: Four Models for External Validity Checks

The outputs of the four linear regression models to the time-series cross-sectional panel data from 2001 to 2012 are presented in **Table 9**. In each of the first two models, keeping other predictors fixed, a country preferred by HROs in the issue of vulnerable group rights, compared to the country spotlighted in sustainable development issues, tends to increase CSO repression in the next year by around 0.3 units on average. And ceteris paribus, a country attacked by HROs in democracy issues, comparing the country highlighted with sustainable development problems, is more likely to cause a 0.4-unit increase on average in CSO repression in the next year.

Table 8: Description of Variables in External Validity Check Models

Variable	Type	Description	Expectation for Correlation	Source
CSO Repression	response variable	Continuous variable from 0 to 5, indicating the level of civil society repression in a given country-year	-	V-Dem Database
CSO Entry and Exit	response variable	Continuous variable from 0 to 5, indicating the level of civil society repression in a given country-year	-	V-Dem Database
Issue Preference	key predictor	Categorical variable indicating an issue HROs prioritize in shaming towards a country in a given year	Less politically salient issues (e.g., sustainable development) are less likely to increase civil society repression.	Created in this project
NGO-to-Government Conflicts	predictor	Count of events of NGO-to-government interactions in the news in a given country-year	More reported conflictual attitudes from NGOs against governments tend to increase civil society repression.	the ICREWS (Integrated Crisis Early Warning System) database
Neighborhood Effects	predictor	the percentage of neighboring countries that have adopted administrative crackdown measures on NGOs in a given country-year	More *Neighborhood Effects* tend to increase civil society repression.	Chaudhry (2022)
HR Protection Score	predictor	Continuous variable from 0 to 8, measuring governments' human rights protection in a given country-year. The smaller number indicates less respect for human rights.	Higher HR Protection Score tends to decrease civil society repression.	Fariss et al. (2020)

Table 9: External Validity Check Models of HROs' Issue Preference on Civil Society Repression

	Dependent variable:			
	CSO Repression (t+1) **Model 1**	*CSO Entry and Exit (t+1)* **Model 2**	*CSO Repression (t+1)* **Model 3**	*CSO Entry and Exit (t+1)* **Model 4**
Issue Preference: economy (baseline = "sustainable development")			-0.352	-0.619
			(0.407)	(0.415)
Issue Preference: freedom of expression (baseline = "sustainable development")			**0.403****	**0.419****
			(0.168)	(0.171)
Issue Preference: minority rights (baseline = "sustainable development")	0.189	0.254*	0.162	0.206
	(0.141)	(0.144)	(0.144)	(0.147)
Issue Preference: peace (baseline = "sustainable development")	-0.123	-0.004	-0.150	-0.051
	(0.406)	(0.414)	(0.407)	(0.415)
Issue Preference: vulnerable group rights (baseline = "sustainable development")	**0.300****	**0.315****	**0.273****	**0.268****
	(0.128)	(0.130)	(0.131)	(0.134)
Issue Preference: democracy (baseline = "sustainable development")	**0.404*****	**0.405*****	0.291	0.156
	(0.153)	(0.156)	(0.250)	(0.255)
NGO-to-Government Conflicts	0.035	0.021	0.035	0.021
	(0.024)	(0.024)	(0.024)	(0.024)
Neighborhood Effects	3.291***	2.091***	3.288***	2.087***
	(0.596)	(0.609)	(0.597)	(0.609)
HR Protection Score	-0.483***	-0.351***	-0.481***	-0.348***
	(0.033)	(0.034)	(0.033)	(0.034)
Constant	-1.450***	-1.373***	-1.423***	-1.326***
	(0.120)	(0.122)	(0.124)	(0.127)
Observations	876	876	876	876

R^2	0.263	0.151	0.264	0.154
Adjusted R^2	0.257	0.144	0.256	0.145
Residual Std. Error	1.295 (df = 868)	1.322 (df = 868)	1.296 (df = 866)	1.321 (df = 866)
F Statistic	44.270*** (df = 7; 868)	22.050*** (df = 7; 868)	34.495*** (df = 9; 866)	17.535*** (df = 9; 866)
Note:				*p<0.1; **p<0.05; ***p<0.01

To better understand countries' responses towards HROs' issue preference, I visualize the marginal effects of *issue preference*, which can present how the predicted values of the dependent variable change on average when *issue preference* varies at 90% confidence intervals in **Figure 25**. The visual on the left corresponds to model 1, while the right one is based on model 2. In **Figure 25**, although the difference between freedom of expression and sustainable development and the difference between vulnerable group rights and sustainable development, the values of *issue preference*, in **Table 9** are slight, significance does exist at 90% confidence intervals. In sum, the outcomes of the models, including the significance of issue preference and controls, echo the findings in the IR literature. As a result, the new measure is externally valid.

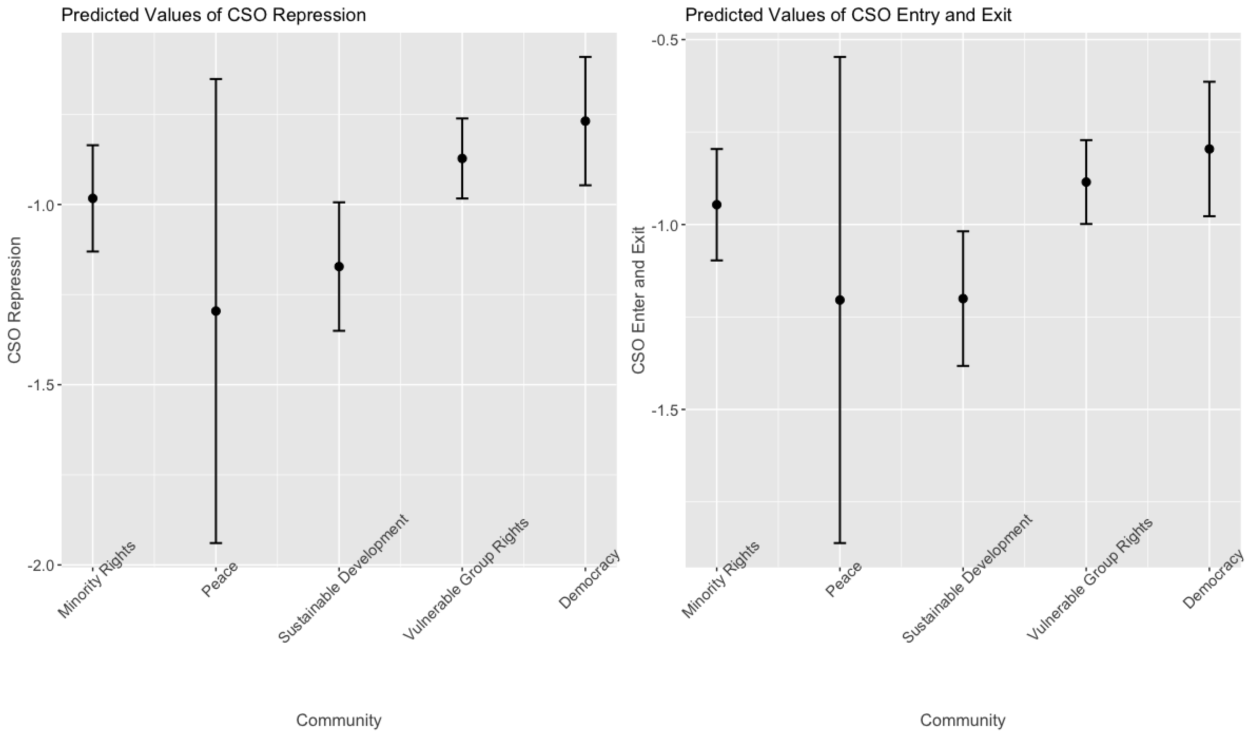

Figure 25: Predicted Civil Society Repression Based on Issue Preference with 5 Communities at the 90% Prediction Intervals (The Left Plot is for Model 1 and the Right for Model 2)

However, for Models 3 and 4, although *issue preference* presents statistical significance, when I plot the marginal effects of *issue preference* in the two models with 90% confidence intervals, the significant difference between vulnerable group rights and sustainable rights in **Table 9** disappears in **Figure 26**. It fails to echo the literature suggested by *issue preference* of 5 communities. It is another reason to choose the partition of 5 communities to measure *issue preference*.

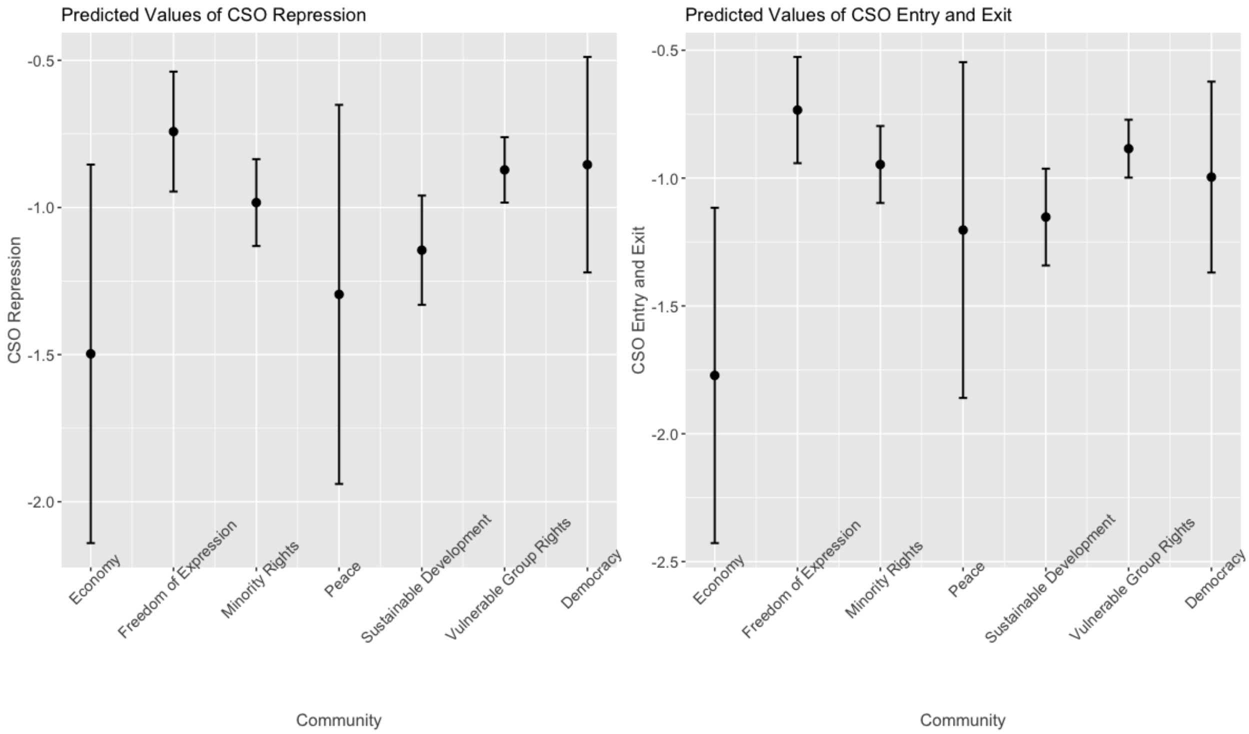

Figure 26: Predicted Civil Society Repression Based on Issue Preference with 7 Communities at the 90% Prediction Intervals (The Left Plot is for Model 3 and the Right for Model 4)

CHAPTER 7

CONCLUSION

In this project, I attempt to further the understanding of how the behaviors of NGOs, especially HROs, affect countries. I create a new measure, *issue preference*, to indicate how HROs prioritize countries in the interactions. Admittedly, different communities can be formed based on (1) the instability of community numbers conditional on resolution parameter selection, and (2) other summaries of community issues. However, this project is an innovative attempt to transfer the attributes at the organizational level to the indicators at the country level by combining statistical modeling and IR expertise. It explains what human rights HROs attempt to improve in a given country.

The country-year measure, a typical and popular unit of analysis in IR, can provide more insights into the study of NGO-government interactions. Besides the question of the effect of issue preference on civil society repression in the validity checks, it would be interesting, for instance, to study the motivations of HROs' *issue preference*. What leads to emphasizing some issues rather than others in a country? What conditions the change of *issue preference*?

Another contribution of the project is that it offers an option to reduce the bias caused by leading HROs' dominant interactions with countries. The Leiden algorithm achieves the highlights of other HROs, which avoids an HRO affecting more than one community each time, and by considering the goal relevance of all the HROs within the same community.

Finally, it provides the study of interactions between two different types of actors in social science with a pattern of classification. The introduction of the Leiden algorithm can not only classify the actors but also incorporate the literature in social science.

REFERENCES

Blondel, Vicent D., Jean-Loup Guillaume, Renaud Lambiotte, and Etienne Lefebvre. 2008. Fast Unfolding of Communities in Large Networks. *Journal of Statistical Mechanics: Theory and Experiment* 10: 10008. https://doi.org/10.1088/1742-5468/2008/10/P10008

Brathwaite, Robert, and Baekkwan Park. 2019. Measurement and Conceptual Approaches to Religious Violence: The Use of Natural Language Processing to Generate Religious Violence Event-Data." *Politics and Religion* 12(1): 81-122. https://doi.org/10.1017/S1755048318000317

Carothers, Thomas, and Saskia Brechenmacher. 2014 *Closing Space: Democracy and Human Rights Support Under Fire*. Carnegie Endowment for International Peace. https://carnegieendowment.org/2014/02/20/closing-space-democracy-and-human-rights-support-under-fire-pub-54503. Accessed 8 November 2022

Carpenter, R. Charli. 2007. Setting the Advocacy Agenda: Theorizing Issue Emergence and Nonemergence in Transnational Advocacy Networks. *International Studies Quarterly* 51(1): 99-120. https://doi.org/10.1111/j.1468-2478.2007.00441.x

Chaudhry, Suparna. 2019. Bridging the Gap: The Relationship Between INGO Activism and Human Rights Indicators. *Journal of Human Rights* 18(1): 11-133. https://doi.org/10.1080/14754835.2019.1579638

Chaudhry, Suparna. 2022. The Assault on Civil Society: Explaining State Crackdown on NGOs. *International Organization*: 1-42. https://doi.org/10.1017/S0020818321000473

Cooley, Alexander, and James Ron. 2002. The NGO Scramble: Organizational Insecurity and the Political Economy of Transnational Action." *International security* 27(1): 5-39. http://www.jstor.org/stable/3092151. Accessed 1 April 2023

Hadden, Jennifer, and Sarah Sunn Bush. 2021. What's Different About the Environment? Environmental INGOs in Comparative Perspective. *Environmental Politics* 30(1-2): 202-223. https://doi.org/10.1080/09644016.2020.1799643

Huimin Cheng, Ye Wang, Ping Ma, Amanda Murdie. 2021. Communities and Brokers: How the Transnational Advocacy Network Simultaneously Provides Social Power and Exacerbates Global Inequalities, *International Studies Quarterly* 65(3):724-738. https://doi.org/10.1093/isq/sqab037

Fariss, Christopher J., Michael R. Kenwick, and Kevin Reuning. 2020. Estimating One-Sided-Killings from a Robust Measurement Model of Human Rights. *Journal of Peace Research* 57(6): 801-814. https://doi.org/10.1177/0022343320965670

Franklin, James C. 2008. Shame on you: The impact of human rights criticism on political repression in Latin America. *International Studies Quarterly* 52 (1): 187-211. https://doi.org/10.1111/j.1468-2478.2007.00496.x

Fuhrmann, Matthew, and Yonatan Lupu. 2016. Do Arms Control Treaties Work? Assessing the Effectiveness of the Nuclear Nonproliferation Treaty. *International Studies Quarterly* 60(3): 530-539. https://doi.org/10.1093/isq/sqw013

Glasius, Marlies, Jelmer Schalk, and Meta De Lange. 2020. Illiberal Norm Diffusion: How Do Governments Learn to Restrict Nongovernmental Organizations?. *International Studies Quarterly* 64(2): 453-468. https://doi.org/10.1093/isq/sqaa019

Hadden, Jennifer. 2015. *Networks in Contention: The Divisive Politics of Climate Change.* Cambridge Studies in Contentious Politics. Cambridge: Cambridge University Press.

Hafner-Burton, E. 2008. Sticks and Stones: Naming and Shaming the Human Rights Enforcement Problem. *International Organization* 62(4): 689-716. https://doi.org/10.1017/S0020818308080247

Hughes, Melanie M., Pamela Paxton, Sharon Quinsaat, and Nicholas Reith. 2018. Does the Global North still Dominate Women's International Organizing? A Network Analysis from 1978 to 2008. *Mobilization: An International Quarterly* 23(1): 1-21. https://doi.org/10.17813/1086-671X-23-1-1

Keck, Margaret E., and Kathryn Sikkink. 1998. *Activists beyond borders: Advocacy networks in international politics*. Cornell University Press.

Kelly ST. 2022. *Leiden: R Implementation of the Leiden Algorithm*. R package version 0.4.3, https://github.com/TomKellyGenetics/leiden.

Kim, In Song, and Dmitriy Kunisky. 2021. Mapping Political Communities: A Statistical Analysis of Lobbying Networks in Legislative Politics. *Political Analysis* 29(3): 317-336. https://doi.org/10.1017/pan.2020.29

Kim, Yooneui. 2023. A Network Analysis of Naming and Shaming in the Universal Periodic Review. *International Interactions*: 1-28. https://doi.org/10.1080/03050629.2023.2172003

Kinne, Brandon J. 2013. Network Dynamics and the Revolution of International Cooperation." *American Political Science Review* 107(4): 766-785. https://doi.org/10.1017/S0003055413000440

Kinne, Brandon J. 2018. Defense Cooperation Agreements and the Emergence of a Global Security Network. *International Organization* 72(4): 799-837. https://doi.org/10.1017/S0020818318000218

Kinne, Brandon J., and Zeev Maoz. 2023. Local Politics, Global Consequences: How Structural Imbalance in Domestic Political Networks Affects International Relations. *The Journal of Politics* 85(1): 107-124. https://doi.org/10.1086/720311

Lian, Shanshan, and Amanda Murdie. 2023. How Closing Civil Society Space Affects NGO-Government Interactions. Journal of Human Rights: 1-20. https://doi.org/10.1080/14754835.2022.2158723

Manning, Christopher, Mihai Surdeanu, John Bauer, Jenny Finkel, Steven Bethard and David McClosky. 2014. The Stanford CoreNLP Natural Language Processing Toolkit. *Proceedings of 52nd Annual Meeting of the Association for Computational Linguistics: System Demonstrations*: 55–60.

Meernik, James, Rosa Aloisi, Marsha Sowell, and Angela Nichols. 2012. The Impact of Human Rights Organizations on Naming and Shaming Campaigns. *Journal of Conflict Resolution* 56 (2): 233-256. https://doi.org/10.1177/0022002711431417

Murdie, Amanda. 2013. The Ties That Bind: A Network Analysis of Human Rights International Nongovernmental Organizations. *British Journal of Political Science* 44(1): 1-27. https://doi.org/10.1017/S0007123412000683

Murdie, Amanda, and David R. Davis. 2012a. Shaming and Blaming: Using Events Data to Assess the Impact of Human Rights INGOs. *International Studies Quarterly* 56(1): 1-16. https://doi.org/10.1111/j.1468-2478.2011.00694.x

Murdie, Amanda, and David R. Davis. 2012b. Murdie, Amanda, and David R. Davis. Looking in the Mirror: Comparing INGO Networks Across Issue Areas. *The Review of International Organizations* 7: 177-202. https://doi.org/10.1007/s11558-011-9134-5

Murdie, Amanda, David R. Davis, and Baekkwan Park. 2020. Advocacy Output: Automated Coding Documents from Human Rights Organizations. *Journal of Human Rights*, 19(1): 83-98. https://doi.org/10.1080/14754835.2019.1671173

Lee, Myunghee, and Amanda Murdie. 2021. The Global Diffusion of the #MeToo Movement. *Politics & Gender* 17(4): 827-855. https://doi.org/10.1017/S1743923X20000148

Murdie, Amanda, and Marc Polizzi, 2017. Human Rights and Transnational Advocacy Networks, In *The Oxford Handbook of Political Networks, Oxford Handbooks*, edited by Jennifer Nicoll Victor, Alexander H. Montgomery, and Mark Lubell: 715-732. https://doi.org/10.1093/oxfordhb/9780190228217.013.31, Accessed 2 April 2023

Risse, Thomas, and Kathryn Sikkink. 1999. The Socialization of International Human Rights Norms into Domestic Practices: Introduction. In *The Power of Human Rights: International Norms and Domestic Change*, edited by Thomas Risse, Stephen C. Ropp, and Kathryn Sikkink, 1-38. Cambridge: Cambridge University Press.

Smith, Jackie, Dawn Wiest, Melanie Hughes, Samantha Plummer, and Brittany Duncan. 2019. Transnational Social Movement Organizations Dataset (TSMOD), 1953–2013. *Computer file]. Harvard Dataverse [distributor]. doi* 10.

Stroup, Sarah S., and Wendy H. Wong. 2017. *The authority trap: Strategic choices of international NGOs*. Cornell University Press.

Traag, Vincent A., Ludo Waltman, and Nees Jan Van Eck. 2019. From Louvain to Leiden: Guaranteeing Well-Connected Communities. *Scientific Reports* 9(1): 5233. https://doi.org/10.1038/s41598-019-41695-z

UIA (Union of International Associations). 2022. *Yearbook of international organizations.*

Von Stein, Jana. 2005. Do Treaties Constrain or Screen? Selection Bias and Treaty Compliance. *American Political Science Review* 99(4): 611-622. https://doi.org/10.1017/S0003055405051919

Wilson, Maya, David R. Davis, and Amanda Murdie. 2016. The View from the Bottom: Networks of Conflict Resolution Organizations and International Peace. *Journal of Peace Research* 53 (3): 442-458. https://www.jstor.org/stable/43920600