Supplementary Materials for
Who Securitizes? Climate Change Discourse in the United Nations

Sabrina B. Arias*

November 28, 2021

Contents

*Ph.D. Candidate, Political Science, University of Pennsylvania, sarias@sas.upenn.edu

Speech Data

I utilize the data collected by Baturo, Dasandi, and Mikhalylov (2017), which consists of all speeches given by state representatives in the General Debate from 1970-2014. Not only are speeches good indicators of country preferences and priorities, by securitizing issues, they also perform an agenda-setting function. Each year at the opening of the UNGA in September, the General Debate gives the opportunity for each state to speak in a largely unconstrained setting (Smith, 2006). Because General Debate speeches are not linked to particular resolutions or votes, which are traditionally used by researchers to measure state preferences, they are more informative about a country's underlying priorities and positions.[1] Every country has equal opportunity to speak, affording small states a "rare moment for seizing the spotlight and putting a point of view that might otherwise be ignored," (Nicholas, 1971, 108). The audiences for these speeches include domestic and foreign publics, bureaucrats at the UN, and members of other state delegations. States take the General Debate seriously: each year, nearly all countries who are can do so choose to give speeches in the UNGA plenary session.

States send high-level representatives to the session, with 44.3% represented by heads of state or government, 49.3% by vice-presidents, deputy prime ministers, and foreign ministers, and only 6.4% by country representatives to the UN (Baturo et al. 2017, 3). An institutional norm restricts speech-time to 15 minutes. While some countries ignore the limitation on length, speech-length has indeed declined over time. We may thus consider speech-time as a limited resource – countries are simply unable to address every issue in a given speech because of time considerations. Allocating the scarce resource of speech-time to discuss a given issue is a signal that a country considers it to be of great importance.

One may worry that the speeches delivered in the General Debate are not independent observations, that is, the order in which the speeches are given may have effects on their substantive contents, or speeches may be influenced by previous years' contents. Strategic coordination and political sources of influence are widespread in state voting records in the UN (Voeten, 2013), but procedural constraints of the General Debate make this an unlikely concern in this speech data. Speeches are uniquely crafted each year to reflect current events and themes highlighted by the Secretary-General. The speeches are then submitted in advance of the General Debate to allow for translation into the official languages of the UN, and to circulate the text to the press and the other delegates of the UNGA. As such, speech content is determined in advance of the General Debate rather than crafted in response to statements by earlier speakers. Further, because states consider the General Debate a consequential platform, many people are involved in the speechwriting process from country missions and governments, and therefore the content is determined well in advance of the General Debate.

The process in which the order of speakers is determined also weighs against a strategic selection process in which the order of speeches may affect their contents. Per tradition, Brazil and the United States always give the first speeches of the General Debate. Subsequently, the order of countries is determined by the importance of the delegation's speaker, with heads of state prioritized. Only after these factors are used in ordering are other factors taken into consideration in setting the speech order, including

[1] Interviews with officials from state Permanent Missions to the UN inform and support the claims made in this section.

individual country preferences for speaking order and geographical balance. Based on variation in these factors, the order of speakers and the speech content varies from year to year.

The data used in this analysis is pre-processed by a text tiling procedure, which identifies features to divide speeches into substantively coherent units ("segments") that are analogous to paragraphs (Hearst, 1997). This procedure is necessary because in many of the documents, natural paragraph delimiters are not available and formatting indicators (i.e. line breaks) are inconsistent across units, so semantic similarity is used to determine paragraph-like units.

After separating the speeches into segments, I extract the speeches on climate change from the full corpus by identifying segments that contain any of a set of keywords, shown in Table 1. The set of key terms was composed in several stages. I began with words that appeared frequently in the context of climate change discussions in the course of my research. I then expanded the set of key words by iterating on the initial set, finding the words whose occurrence correlated the most highly with the initial set and were substantively related to climate change. These discovered words were added to the initial set to create the final set of filtering words. States that do not speak on climate change at all, states that do not exist after 1984, and speeches that do not represent a specific state are dropped from the analysis – this includes speeches given on behalf of the European Union, the European Community, Yugoslavia, Czechoslovakia, Democratic Yemen, and the German Democratic Republic. All other states are included.

Table 1: Terms Used to Create Climate Filter

1	climate change	7	climate politics	13	global average temperature
2	global warming	8	framework convention on climate change	14	greenhouse effect
3	cap and trade	9	bali roadmap	15	kyoto protocol
4	unfccc	10	bali action plan	16	ipcc
5	paris accord	11	greenhouse gas	17	greenhouse effect
6	emissions trading scheme	12	ghg	18	intergovernmental panel on climate change

Descriptive statistics of the full corpus and the climate subset are shown in Tables 2 and 3. The overall length of speeches declines over time, dropping off particularly steeply after 2000. In any speech, the greatest number of segments that specifically discuss climate change is 16 (a speech by Samoa in 2015), while the average is 4 segments. SIDS discuss climate change at higher rates than other states with more segments about climate change and greater speech proportions on the topic, but they are not necessarily early adopters, picking up the climate discourse at the same time as other states. There are few outliers in terms of speech length: one exceptionally long speech was given by Libya in 2009 (100 segments). Only four other countries gave more than one speech longer than 50 segments (Russia, USA, Cuba, and Germany), and multiple long speeches were given in 1973, 1976, 1978, 1983, 1984, and 2009.

To prepare this speech data for unsupervised text analysis, additional pre-processing steps were needed. I remove word stems that occurred in fewer than 1% of documents or in more than 95% of documents, common stop words, as well as documents that contained only unique word stems (that is, shared no features with other documents). I also remove the terms 'united', 'nations', 'general', and 'assembly', as these will occur too frequently to be informative.

Table 2: Descriptive Statistics: All Segments

	# Segments per Speech	# Speeches per Year	Avg. # Segments per Year	Total Segments by Year	Avg. Segments Per Country
Min.	4	150	17.1	2963	11.2
1st Qu.	15	191	19.1	3271	22.6
Median	19	193	20.8	3552	28.0
Mean	20	191	20.9	3565	27.6
3rd Qu.	23	194	21.2	3662	31.9
Max.	93	196	40.4	5550	51.9

Table 3: Descriptive Statistics: Climate Segments Only

	# Segments per Speech	Prop. of Speech	# Speeches per Year	Avg. # Segments per Year	Total Segments per Year	Avg. Segments per Country
Min.	1	0.01	1	1.0	1	1.0
1st Qu.	2	0.09	125	3.3	270	2.3
Median	3	0.16	147	3.8	360	3.1
Mean	4	0.19	126	3.5	315	3.6
3rd Qu.	5	0.26	157	4.1	404	4.3
Max.	16	0.73	169	4.4	492	8.3

Topic Models

Methodology and Estimation

Structural Topic Model (STM) is a variant of Latent Dirichlet Allocation (LDA), which is a Bayesian generative model of language. LDA, and the STM variant of LDA employed in this analysis, assumes a hierarchical system of distributions, with an underlying Poisson distribution of words (N) and of Dirichlet topic probabilities (θ) across topics (k). Conditional on these priors of words and of topics, each of N words in a document w_n is drawn, and each topic z_n and word w_n from a Multinomial. The β parameter is a KxV matrix of word probabilities (Grimmer and Stewart 2013). STM adjusts this the LDA procedure to allow the topic proportions of θ (referred to as topical prevalence) and the observed words n (referred to as topical content) to be drawn from a document-specific prior rather than a universal prior. Covariates associated with each document can inform the distribution of topics and words across documents (Roberts et al. 2016). To allow for variation in the content and prevalence of topics over time, I include year as a parameter in the STM. To allow the prevalence of topic to vary non-linearly over time, I fit a spline on years in topical prevalence. To allow topical content to vary over time but to impose some constraint, I fit a factor on decade in topical content.[2]

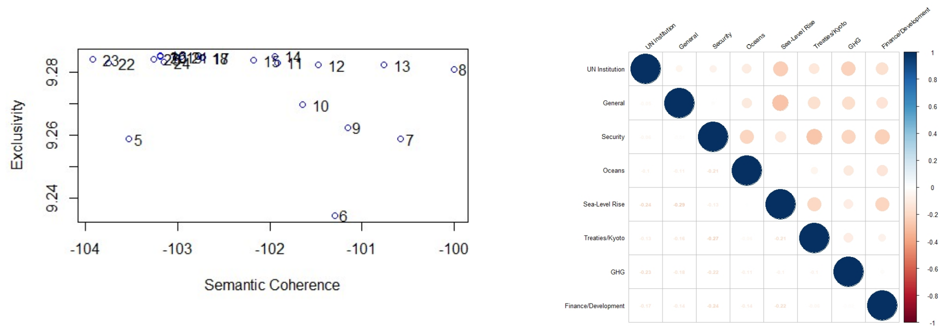
I estimate the STM models at the standard value value of the prevalence hyperparameter (α) as $1/k$, though changing the value of α had little effect on the results. I employ a Spectral initialization for stability. Because the corpus is already limited to the particular topic of climate change, I employ a smaller k than would be typical for analysis across an entire corpus. I test values of k ranging from 2 to 15 and assess topical coherence manually to find the optimal number of topics. I estimate the STM with $k = 8$ topics based on manual evaluation of topical coherence, as well as maximization of semantic coherence and exclusivity, seen in the left panel of Figure 1. The distinctiveness of the

[2]Results were largely consistent across different parameterization choices of year and decade; the final modeling choices were made to minimize the residuals in the model.

different topics obtained in the final model is validated by an analysis of the correlation between the different topics, which was found to be low across all topical pairs, shown in the right panel of (Figure 1). The ultimate topics obtained were quite consistent across all the different model specifications described above.

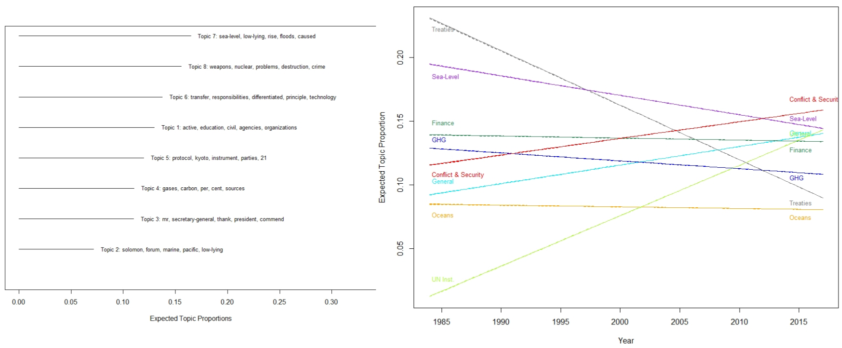
The topic proportions are shown in the left panel of Figure 2 and Table 1 of the main text. The conflict and security topic is the second most frequently employed. It is also used more over time, while other frames, particularly the Kyoto Protocol frame, decrease in their use over time (Figure 2, right panel). Across all time periods, the P5 employ the language of securitization more than do SIDS (Figure 3)

Figure 1: Selecting Number of Topics



Note: Number of topics (k was chosen to maximize semantic coherence and exclusivity (left), to minimize correlation between topics (right), and based on best performance under manual examination.

Figure 2: Topic Proportions

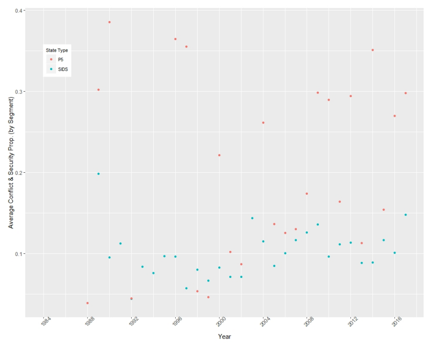


Note: Left panel shows expected topic proportions across the corpus, right panel shows linear estimation of changing topic proportions over time.

Robustness

One of the key benefits of topic modeling is its inductive nature, reducing the impact of the *ex ante* beliefs of the researcher. Rather than pre-specifying the words that I expect to be associated with securitization, as in a dictionary-based approach,

Figure 3: P5 Make Securitizing Moves More Than SIDS Across Time



I allow the model to identify the words associated with securitization. However, topic modeling can be sensitive to the specifications of the model, so I take several steps to ensure robustness, aligning with the recommendations on validation outlined in Ying et al. (2021).

I validate the results of the STM with a supervised analysis. I randomly selected 10% of the speech segments on climate change (448 segments) and read them to identify the topics that they employ, coding them as being security relevant or not. In this exercise, I was also able to test the keyword filtering approach, verifying that there were no false positives (that is, there were no speeches included in the data that were not actually about climate change). Based on my binary coding of the segments as security-relevant or not, I found a high level of agreement with the results of the STM. When the results of the STM are binarized at the level of the mean (segments with 15% or higher topical prevalence in the conflict and security topic are coded as security-relevant), my coding agreed at 58% with segments in the conflict and security topic. When the results of the STM are binarized as more than half of the paragraph identified as the conflict and security topic, my coding agreed at 67%. This provides confidence that the results of the unsupervised analysis are not artifact but are indeed identifying clusters of security language in the data. Based on the words found to be most representative of the different topics (words within each topic with the highest topic-word probability, β), I constructed the label for each.

These labels were validated with a crowdsourcing exercise whereby anonymous respondents were shown the set of words representing each topic, conducted with a convenience sample and expanded with snowball sampling in November 2019. Respondents were instructed, "For each of the following eight questions, you will be shown a set of 20 words. Please respond with one word that you think best summarizes the content of these words- that is, assign the best fitting (in your opinion) topic label to the words- there are no incorrect responses." For the conflict and security topic, the words presented in the task were "weapons, nuclear, problems, destruction, crime, war, mass, threats, hunger, conflict, today, diseases, live, terrorism, still, life, crises, values, increasingly, complex."

Out of 16 respondents (10 non-political scientists and 6 political scientists), 10 provided words that corresponded with my characterization of the topic ("security, crises, war, weapons of mass destruction, fighting, war and terrorism, conflict, international conflict, war, threat"), while the remaining 6 respondents provided more general words ("Middle East, international relations, human influence, changing society, foreign policy, fear"). No respondents assigned a security0related label to Topic 7 (sea-level rise). For Topic 7, the most common label applied was in fact was "climate change" generally (9/16 respondents), followed by "natural disasters" (4/16). There were no noticeable differences in responses across political scientists from others.

A second supervised task asked respondents to provide words based on the top segments in the topics, rather than the top words. Respondents were instructed, "You will be shown a series of 10 short paragraphs. After each paragraph, you will be asked to respond with one word/phrase that you think best summarizes the content of these paragraphs- that is, assign the best fitting (in your opinion) topic label to the paragraphs- there are no incorrect responses, but since all of the paragraphs are about climate change, please DO NOT say that "climate change" is the best label." 10 respondents were shown in a random order the 5 speech segments with the highest topic proportion in the sea-level rise topic, and the 5 speech segments with the highest topic proportion in the conflict and security topic. In this task, respondents were 4 times more likely to assign a security label to the conflict and security topic segments than to the sea-level segments. Topic labels for the sea-level rise segments highlighted natural disasters, the challenges for small island states, and inequality.

Predicting Securitization

Variables

- **P5:** Indicator for P5 Member status (United States, United Kingdom, China, Russia, and France).
- **E10:** Indicator for elected UNSC membership status during the years when a state served on the UNSC (Dreher et al. 2009).
- **SIDS:** Indicator for UN designation as a small island developing state (Cuba, Dominican Republic, Haiti, Singapore, Trinidad and Tobago, Fiji, Guyana, Jamaica, Mauritius, Barbados, Bahamas, Grenada, Comoros, Cabo Verde, Guinea-Bissau, Maldives, Papua New Guinea, Sao Tome and Principe, Suriname, Seychelles, Samoa, Dominica, Saint Lucia, Saint Vincent and the Grenadines, Antigua and Barbuda, Belize, Solomon Islands, Vanuatu, Saint Kitts and Nevis, Micronesia, Marshall Islands, Palau, Nauru, Tonga, Tuvalu, Kiribati, Timor-Leste) (United Nations Statistics Division, 2021).
- **Developing:** Indicator for UN designation as a developing state (144 countries) (United Nations Statistics Division, 2021).
- **Vulnerability:** Vulnerability is measured with the ND-GAIN index, a composite measure combining metrics of country-year level vulnerability to climate disruptions and readiness to leverage private and public sector investment for adaptive actions(University of Notre Dame, 2021). This data is available beginning in 1992, so for analyses utilizing this indicator, the time series is left-censored at that point. In addition to the continuous measure of ND-GAIN, four binary indicators are constructed with this indicator to test for robustness of the most vulnerable measure: countries with lower than mean (49.10) ND-GAIN composite scores, countries in the lowest quartile of ND-GAIN composite scores (less than 40.22), countries with higher than mean (0.46) vulnerability composite scores, and countries in the fourth quartile of vulnerability composite scores (higher than 0.53). Mean ND-GAIN composite is reported in the results below, results were robust across all specifications of the vulnerability variable (results with these specifications available upon request).

- **Public Concern**: This data comes from the 2010 Gallup World Poll conducted across 106 countries in 2010 that featured two questions on climate change (Gallup 2010). The first survey question measured *understanding*, asking respondents "How much do you know about global warming of climate change?" and coding understanding as the percentage of respondents saying they know something or a great deal. The second question measured *concern*, asking respondents "How serious of a threat is global warming to you and your family?" and coding concern as the percentage of respondents reporting that they view global warming as a very or somewhat serious threat. This question was only asked to respondents reporting familiarity with the topic of climate change. The panel included all of the P5 members except France, as well as 4 SIDS (Singapore, Dominican Republic, Haiti, and Comoros, out of 37 total SIDS).[3] I use the measure of concern rather than understanding in the model because it more closely capture the kind of electoral pressure that would be likely to move policymakers to advance the issue of climate change on the global agenda.

- **Democracy:** I use scores from the Polity IV project to indicate regime type (Center for Systemic Peace 2018). Polity scores can take on a possible range of -10 to 10. Countries with Polity scores greater than 6 are coded as democracies, while countries with Polity scores less than -6 are coded as autocracies.

- **Climate Disasters:** I construct a measure of climate disaster occurrences to capture the experienced effects of climate change across states, utilizing data from the International Disaster Database from 1984-2018 (Guha-Sapir 2015). I collect data on climate-related disasters, including climatological (drought and wildfire), hydrological (flood and landslide), and meteorological (extreme temperature and storm). The disaster database captures the number of deaths, injured, affected, homeless, and costs for many of these events. To maximize data availability and reduce the effects of income-dependence, I follow Roberts and Parks (2007) and use a smoothed measure of total persons affected. For each country, I sum the number of total persons affected by climate disasters and divide by the country population in each year and then take the log.

- **Amount Warming:** I measure warming as a change in national average air temperature over land since 1960 in degrees Celsius, measured by Berkeley Earth (2020). This measure ranges from 0.65 degrees to 3.6 degrees, with a mean of 1.88 degrees.

- **Agreement:** The measure of affinity is constructed by Bailey, Strezhnev and Voeten (2017). These measures of state preference similarity are constructed using voting records in the UNGA. The voting similarity index compares one country's voting record in a given year with another, ranging from 0-1. The main model includes the measure of voting agreement with the US. I also specify models where this measure is replaced for vote similarity with Brazil and with India to determine

[3]Respondents in P5 states show higher understanding than either other states or SIDS (0.85 for the P5, 0.54 for SIDS, and 0.68 for other states). For both the P5 and SIDS the level of concern is less than other states (0.43 and 0.44 for the P5 and SIDS compared to 0.50 for other states). On both measures, China is substantially lower than other P5 members, while Singapore is substantially higher than other SIDS.

whether coalitions of developing countries are influential in setting the patterns of securitization discourse.

- **Military Expenditures**: Military expenditures as a proportion of GDP, measured by the World Bank Development Indicators (World Bank 2020).
- **Aircraft Carriers:** Indicator for countries with aircraft carriers. Countries included are Argentina, Australia, Brazil, Canada, China, France, Germany, India, Italy, Japan, Netherlands, Russia, Spain, Turkey, Thailand, United Kingdom, and United States. This binary measure specifies countries that have operated aircraft carriers at some point in history. All P5 members are included as naval powers.
- **GDPPC:** Annual measures of country level gross domestic product per capita (logged), measured by the World Bank Development Indicators (World Bank 2020).
- **Population**: Population (logged), measured by the World Bank Development Indicators (World Bank 2020).

Robustness

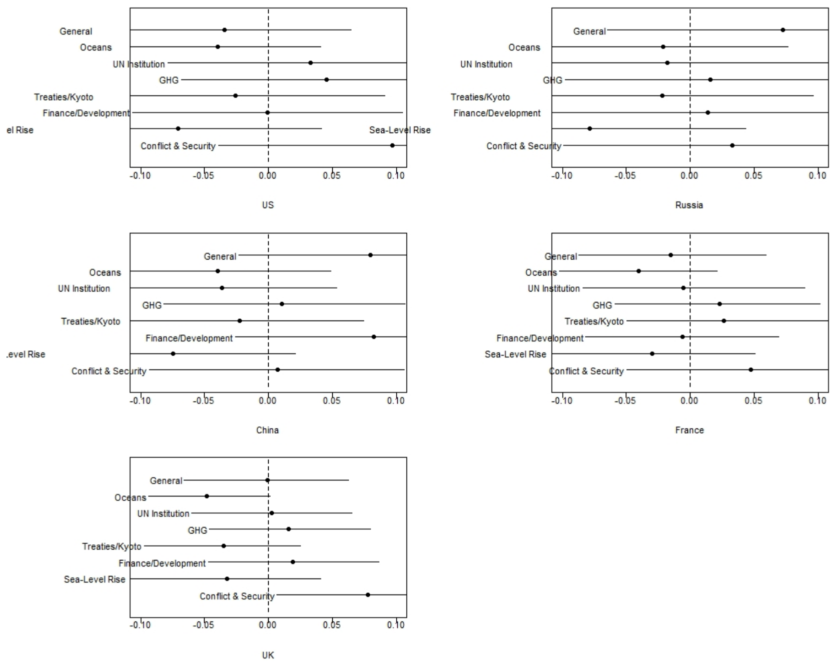
The results in the main model show linear regression estimates with standard errors averaged from 5 imputations of missing data, conducted with Amelia. These results are also observed in simulated first differences (Table 4). The main result that the P5 securitize more than other states holds for each P5 member directionally, although because of the reduced amount of data, the results do not hold statistical significance at the individual level (Figure 4). To further test robustness, I also re-estimate the model on a restricted sample of the climate change corpus to only states with high and medium levels of development, and find that even in the restricted sample, the P5 are more likely to securitize than non-P5 states (Figure 6, right panel), demonstrating that the over-representation of developing states in the original sample does not drive results, though the statistical significance of the finding is attenuated due to reduced sample size. As discussed in the main paper and shown in Table 2 of the main text, I find that the results are robust to many different specifications.

In addition to these primary robustness tests, I also conduct a placebo test to assess whether the P5 simply try to securitize across all topic areas, and the observed securitizing moves on climate by the P5 are not a unique feature of that discourse. The theoretical model of agenda control securitization contends that because of resource and legitimacy constraints, the P5 attempt to securitize issues strategically, and would not be likely to securitize across all issues. The evidence supports this expectation. To verify that the P5 are not simply making securitizing moves across all topics, I conduct a placebo test, fitting an additional topic model with the same specifications to the full General Debate corpus. Comparing the STMs on the climate discourse and the full discourse reveals that conflict security matters comprise a distinct topic in the General Debate, and that this topic particularly focuses on 'hard' security issues such as nuclear weapons, disarmament, and terrorism. There is not evidence that the overall discourse is securitizing: the prevalence of the conflict and security topic is declining over time, whereas the use of security language in the climate discourse specifically was increasing over time (Figure 5). In the overall discourse, the conflict and security topic is not used generally by many states, with only 6% of states using this as their most frequent topic, the second least common of all topics. The security language is also not prevalent across topics in the overall debate, but is confined to the particular conflict and security topic. In the general discourse, the P5 are no more likely to employ the conflict and security topic than other states (Figure 6, left panel). These findings support the analysis of climate discourse as a distinct case of securitization in UN discourse, and provide further evidence that the P5 are expected to securitize strategically, selecting particular issue areas where there securitization might be more likely to be accepted. Security words are not characteristic of other topics.

Table 4: First Differences on Changing P5 and SIDS Status

	P5 Ind	SIDS Ind
Mean	0.0512	-0.0800
St. Dev.	0.0243	0.0117
2.5 CI	0.0044	-0.1028
97.5 CI	0.1014	-0.0571

Figure 4: P5 States Individually Securitize

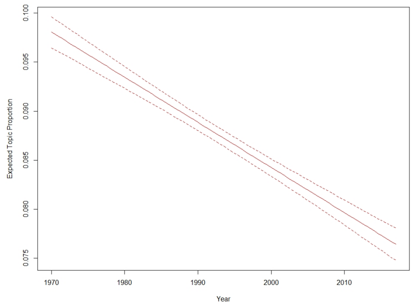


Note: Figures show expected speech segment proportion in the conflict and security topic estimated by STM, comparing each P5 states to other states. Uncertainty calculated from the STM by composition with 95% confidence intervals.

I also estimate the main model using alternative measures to SIDS status to capture vulnerability to the effects of climate change, and thus potential likelihood to make securitizing moves. I find that like SIDS, developing states are less likely to use securitizing language, indicating that concerns about increasing UNSC strength (as well as interest in the multidimensional aspects of climate change) appear to be at work for developing states as a larger category, as seen in the right panel of Figure 7. A similar finding holds in observing states with lower than mean vulnerability in the ND-GAIN index, shown in the left panel of Figure 7.[4] My key prediction – that SIDS are the *least* likely group of

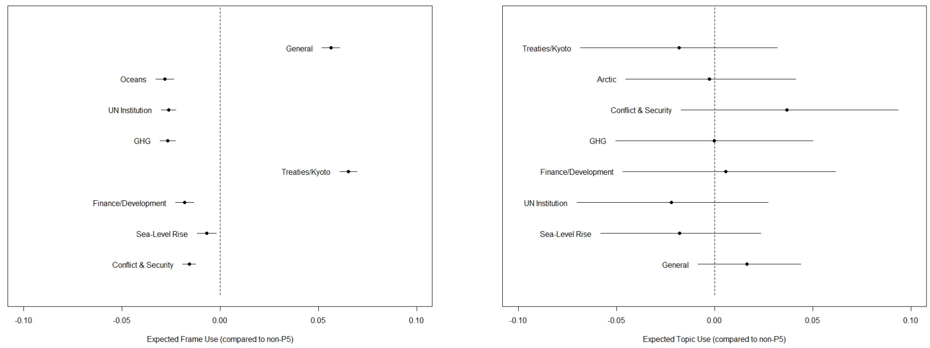
[4]Because the number of observations changes in the ND-GAIN analysis, the estimation of

Figure 5: Security in Placebo Test



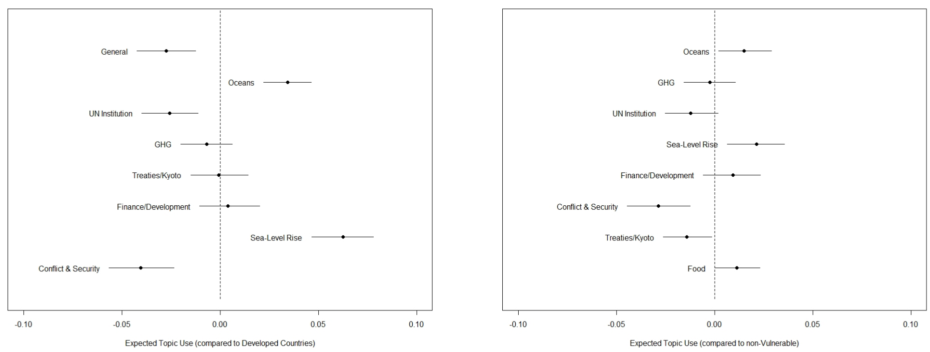
Note: Linear estimation of changing conflict and security topic proportion over time. Uncertainty calculated from the STM by composition with 99% confidence intervals.

Figure 6: P5 Robustness



Note: Left panel shows expected topic proportions in the placebo test of the full corpus, right panel shows expected topic proportions in the restricted sample of the climate change corpus to only states with high and medium levels of development. Uncertainty calculated from the STM by composition with 95% confidence intervals.

Figure 7: Developing states also less likely to securitize, but less so than SIDS



Note: Figures show expected speech segment proportion in the conflict and security topic estimated by STM. Left panel shows developing states compared to others, right panel shows vulnerable states (by ND-GAIN) compared to others. Uncertainty calculated from the STM by composition with 95% confidence intervals.

states to make securitizing moves – also bears out, as SIDS are 1.59 times less likely to use the language of securitization than are developing states. Model 7 (shown in Table 2 of the main text) shows that this result holds for developing states, even when controlling for other confounders: like SIDS, developing states are less likely to use the language of securitization, but the magnitude of this effect is much greater for SIDS. Table 5 shows similar results hold for the vulnerability measure, operationalizing the ND-GAIN measure as a binary (Model 1) and continuous (Model 2) measure. More vulnerable states are less likely to securitize to a substantively identical degree as developing states, although vulnerability is not statistically significant as a continuous measure. Note that the number of observations included in these models is smaller than the main models estimated, as the ND-GAIN measure is only captured beginning in 1992.

Table 5: Linear regression model of securitization

	(1)	(2)
P5	0.00	0.00
	(0.99)	(0.89)
Vulnerability (Binary)	-0.04**	
	(0.00)	
Vulnerability (Continuous)		0.00
		(0.49)
Pubic Concern	0.01	0.02
	(0.91)	(0.70)
Democracy	0.01	0.01
	(0.40)	(0.42)
Climate Disasters	-0.01**	-0.01**
	(0.00)	(0.01)
Amt. Warming	0.00	0.01
	(0.97)	(0.35)
Agree with US	-0.04	-0.04
	(0.23)	(0.25)
Military Expenditures	1.21**	1.19**
	(0.01)	(0.01)
GDPPC (Log)	-0.01	0.00
	(0.36)	(0.77)
Population (Log)	0.01***	0.01***
	(0.00)	(0.00)
Year	0.00	0.00
	(0.89)	(0.86)
Year Sq	0.00	0.00
	(0.56)	(0.55)
Constant	0.13	-0.02
	(0.09)	(0.69)

Note: *p<0.05; **p<0.01; ***p<0.001
Standard errors averaged from 5 imputations of missing data. 1,694 observations in all models.

the STM changes slightly: the numerical ordering of the topics shifts, and the 'general' topic is replaced by a specific 'food' topic.

References for Supplementary Materials

1. Bailey, Michael A., Anton Strezhnev, and Erik Voeten. 2017. "Estimating Dynamic State Preferences from United Nations Voting Data." *Journal of Conflict Resolution* 61(2), 430–456.

2. Baturo, Alexander, Niheer Dasandi and Slava J. Mikhaylov. 2017. "Understanding state preferences with text as data: Introducing the UN General Debate corpus." *Research and Politics* 4(2), 1-9.

3. Berkeley Earth. 2020. "Global Warming by Country."

4. Center for Systemic Peace. 2019. "Polity IV Annual Time-Series, 1800-2018."

5. Dreher, Axel, Jan-Egbert Sturm and James Vreeland. 2009. "Development Aid and International Politics: Does membership on the UN Security Council influence World Bank decisions?" *Journal of Development Economics* 88: 1-18 2009.

6. Gallup. 2010. "Gallup World Poll."

7. Grimmer, Justin, and Brandon M. Stewart. 2013. "Text as Data: The Promise and Pitfalls of Automatic Content Analysis Methods for Political Texts." *Political Analysis* 21.3: 267-297.

8. Guha-Sapir, D. 2015. "EM-DAT: The Emergency Events Database."

9. Hearst, Marti A. 1997. "TextTiling: Segmenting Text into Multi-Paragraph Subtopic Passages." *Computational linguistics* 23.1: 33-64.

10. Nicholas, H. G. 1971. *The United Nations as a Political Institution.* Oxford University Press

11. Roberts, J. Timmons, and Bradley Parks. *A Climate of Injustice: Global Inequality, North-South Politics, and Climate Policy.* MIT Press.

12. Roberts, Margaret E., Brandon M. Stewart, and Edoardo M. Airoldi. 2016. "A Model of Text for Experimentation in the Social Sciences." *Journal of the American Statistical Association* 111.515: 988-1003.

13. Smith, Courtney B. 2006. *Politics and Process at the United Nations: the Global Dance.* Oxford University Press

14. United Nations Statistics Division. 2021. "Standard country or area codes for statistical use."

15. University of Notre Dame. 2021. "Notre Dame Global Adaptation Initiative."

16. Voeten, Erik. 1997. "Data and analyses of voting in the United Nations General Assembly." In *Routledge Handbook of International Organization*, (pp. 80–92). Routledge

17. World Bank. 2020. "World Development Indicators."

18. Ying, Luwei, Jacob M. Montgomery and Jacob M. Montgomery. 2021. "Topics, Concepts, and Measurement: A Crowdsourced Procedure for Validating Topics as Measures." *Political Analysis* 1-20.