Supplementary Appendix

A Data Appendix

A.1 Summary Statistics for Main Variables in Analysis

	N	Mean	SD	Min	Max
Statement-Level					
Uncertainy	551	2.27	0.59	1.00	4.00
Number of References	548	6.28	7.97	0.00	78.00
Statement Length (characters)	574	695.75	478.80	88.00	5579.00
Country-Level					
UN Ideal Point (2007-2013)	85	−0.01	0.86	−1.57	2.60
Total GHG emissions, mt (2013)	84	455.49	1488.53	0.02	11,861.84
GDP per Capita, 000s (2013)	85	23.25	31.38	0.49	185.06
Log GDP (2013)	85	25.27	2.53	17.47	30.46
Dyad-Level					
UN Ideal Point Difference	3570	1.13	0.88	0.00	4.93
Emissions Difference	3486	0.75	1.96	0.00	11.81
log(Trade)	3570	1.85	2.92	0.00	12.85
log(Inter-Capital Distance)	3403	8.79	0.80	4.39	9.89
Military Alliance	3570	0.05	0.22	0.00	1.00
Contiguous	3403	0.02	0.15	0.00	1.00
Common Language	3403	0.17	0.38	0.00	1.00
Common Colonizer	3403	0.09	0.28	0.00	1.00
Dyad-Statement-Level					
Agreement	2097970	0.07	4.30	−100.00	100.00

A.2 Number of Statements Across Reports

Table 4: Number of Statements Included from Different IPCC Reports

		Working Group			
Report		WG1	WG2	WG3	**All**
AR5	N	81	81	95	257
	% row	31.5	31.5	37.0	100.0
AR6	N	83	90	144	317
	% row	26.2	28.4	45.4	100.0
All	N	164	171	239	574
	% row	28.6	29.8	41.6	100.0

A.3 Countries Included in the Analysis

Table 5 lists the countries and supranational entities included in the analysis. The list includes all countries that participated in at least one intervention in any of the plenary sessions included in the analysis. The EU, while a supranational entity, is included as a separate country in the analysis since it often participates as an independent voice at the IPCC plenary. Covariate values for the EU are averaged across its member states (including the UK, which left the EU in 2020[52]).

[52]Excluding the UK does not meaningfully change any results

Table 5: List of Countries and Supranational Entities Included in the Analysis

Angola	Argentina	Antigua and Barbuda
Australia	Austria	Belgium
Belize	Bahamas	Bolivia
Brazil	Bhutan	Botswana
Canada	Switzerland	Chile
China	Côte d'Ivoire	Congo - Brazzaville
Colombia	Costa Rica	Comoros
Cuba	Germany	Denmark
Algeria	Ecuador	Egypt
Spain	Estonia	Ethiopia
Finland	Fiji	France
United Kingdom	Ghana	Guinea
Gambia	Grenada	Indonesia
Hungary	India	Ireland
Iraq	Iceland	Italy
Jamaica	Japan	Kenya
St. Kitts and Nevis	South Korea	St. Lucia
Luxembourg	Madagascar	Monaco
Maldives	Mexico	Mali
Malawi	Malaysia	Netherlands
Norway	New Zealand	Panama
Peru	Philippines	Poland
Qatar	Russia	Saudi Arabia
Sudan	Senegal	Sierra Leone
South Sudan	Slovenia	Sweden
Trinidad and Tobago	Chad	Tanzania
Tuvalu	Ukraine	United States
Venezuela	South Africa	Zambia
European Union		

A.4 Examples for Agreement Measure

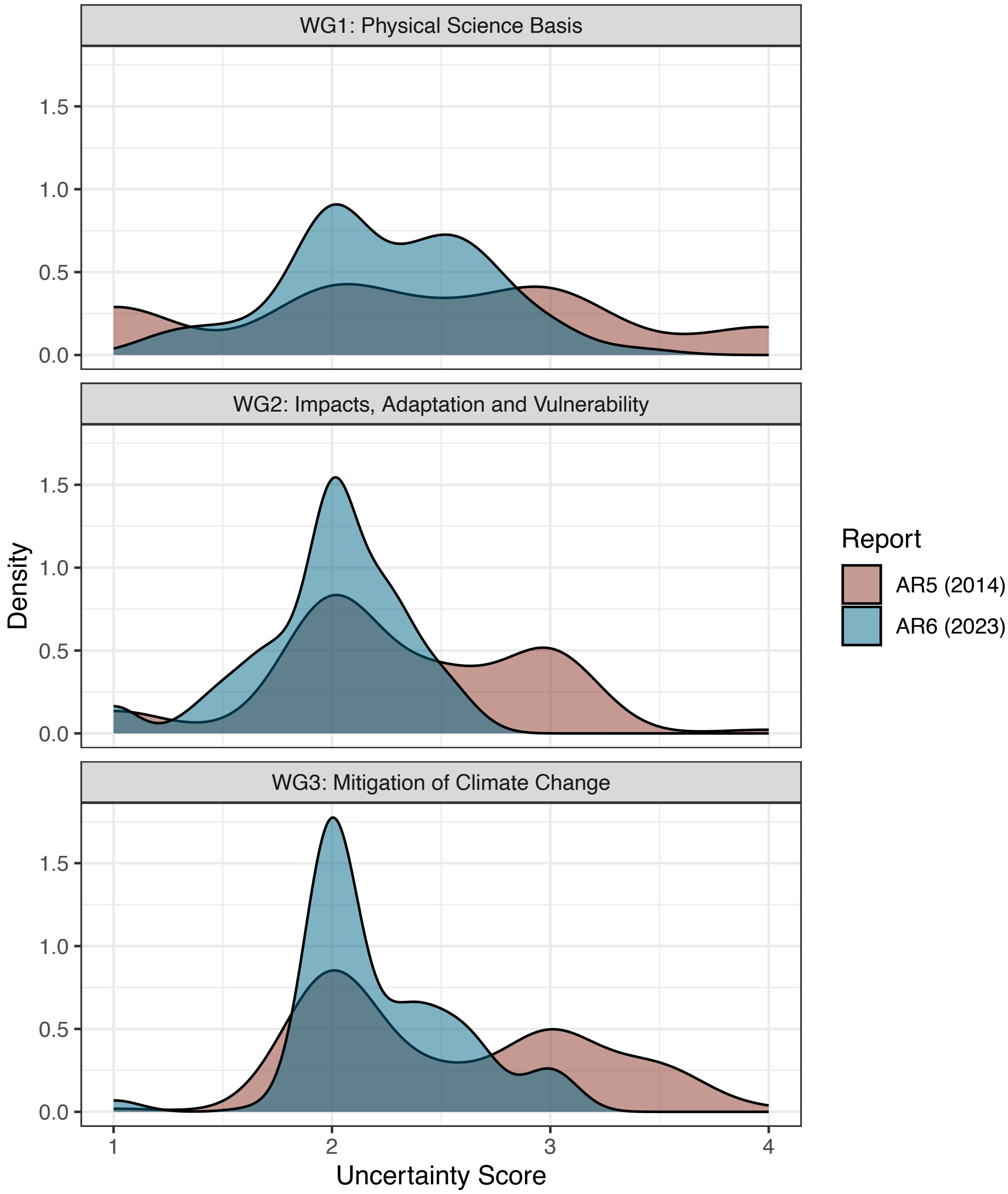
The dyadic agreement measure is scaled to range from -100 to 100 to make downstream results more interpretable. This is because, given that the measure is defined at the dyad-statement level, most dyad-statement combinations will not have any interaction. This does not mean, however, that political influence is rare at the IPCC.

For example, politicization by India and Saudi Arabia arguably plays a large role in shaping the IPCC's Summaries for Policymakers since these are some of the most active countries in plenary sessions. However, at the dyad level, these countries only interacted with each other on 56 out of 574 statements (all in agreement), yielding them an agreement score of 9.8 on a range between -100 and 100. Similarly, conflict between India and the US often features prominently in plenary sessions, but in terms of the agreement measure, these two countries only interacted on 43 statements (28 of which were in opposition), yielding them an agreement score of -2.26 on the same range. These examples illustrate the idea that the highly skewed nature of the agreement measure does not imply that political influence is rare at the IPCC.

A.5 Distribution of Uncertainty

This table plots the distribution of the four-point uncertainty measure by report and Working Group. Working Group 3 reports, which deal with mitigation, tend to have the highest levels of uncertainty on average. By comparison, Working Group 1 reports, which deal with the physical science of climate change, tend to have the lowest levels of uncertainty. Moreover, the 6th Assessment Report features fewer statements with high levels of uncertainty than the 5th Assessment Report. This is consistent with the idea that the underlying science behind climate change has become more certain over time.

Figure 7: Distribution of Uncertainty in Statements



Notes: The figure plots the distribution of statement-level uncertainty, as described in Table 1, by working group and Assessment Cycle.

B Ideal Points Details

B.1 Ideal Point Estimation Details

Model Specification To estimate ideal points, I model country choices as a three-category ordinal outcome defined at the country-intervention level. A given statement might be the subject of zero, one, or more interventions. Let k index such interventions and denote $s(k)$ as the statement that intervention k refers to. The choice of country i on an intervention k on a statement $s(k)$ is defined as follows:

$$y_{ik} = \begin{cases} 1 & \text{if } i \text{ supports an intervention } k \text{ on statement } s(k) \\ 0 & \text{if } i \text{ does not participate in intervention } k \text{ on statement } s(k) \\ -1 & \text{if } i \text{ opposes an intervention } k \text{ on statement } s(k) \end{cases}$$

I use a item-response theory (IRT) approach to model a country's choice of intervention as a function of (a) its ideal point in a latent ideological space, (b) the degree to which a statement activates the predominant cleavage in this ideological space, and (c) parameters governing a statement's tendency to be the subject of an intervention.

Specifically, I define a latent variable $\tilde{y}_{ik}$ that represents the tendency of a country i to intervene on a statement k as follows:

$$\tilde{y}_{ik} = x_i^T \beta_k + \epsilon_{ik}$$

Where the error term ϵ_{ik} follows a standard logistic distribution. The parameter $x_i \in \mathbb{R}^D$ represents the ideal point of country i in a $D-$dimensional latent ideological space. The discrimination parameter $\beta_k \in \mathbb{R}^D$ represents the degree to which statement k activates the predominant cleavage in this ideological space. Appendix Section B provides further details on the estimation of ideal points and the convergence of the model.

Given this latent variable formulation, I model the outcome of a country i intervening on a statement k as follows:

$$y_{ik} = \begin{cases} 1 & \text{if } \tilde{y}_{ik} \geq c_{1k} \\ 0 & \text{if } c_{0k} < \tilde{y}_{ik} < c_{1k} \\ -1 & \text{if } \tilde{y}_{ik} \leq c_{0k} \end{cases}$$

Where the cutpoints c_{0k} and c_{1k} are parameters that govern the thresholds at which a country intervenes to support or oppose a statement. These cutpoints account for differences in agenda between statements, specifically the idea that some statements may be more likely to be the subject of interventions than others.

I use a Bayesian approach to estimate the parameters of this model using `stan`. The parameter of interest is the estimated country-level ideal points x, which is a summary measure of the country's preferences over the interpretation of scientific statements at the IPCC.[53]

Estimation As mentioned in Section 3.1, I use a Bayesian approach to estimate the parameters of the ideal point model using `stan`. I assume the latent ideological space is one-dimensional, and that the error term ϵ_{ik} follows a standard logistic distribution.[54]

The model is estimated using 4 chains of 2000 iterations each, with a warmup of 1000 iterations. As is well-known, ideal point parameters $\{\boldsymbol{x}, \boldsymbol{\beta}, \boldsymbol{c_0}, \boldsymbol{c_1}\}$ are not identified without some form of normalization. I normalize the ideal points to have a mean of zero and a standard deviation of one. I also normalize the statement-level discrimination parameters β and cutpoints c_0 and c_1 to have a mean of zero and a standard deviation of one. To account for rotational invariance, I constrain the United States' ideal point to be 1 and that of Saudi Arabia to be -1. This is a common normalization in ideal point models, and it allows us to interpret the other ideal points as deviations from the US position and towards Saudi Arabia's position. I choose these two countries as anchors since these two countries are often prominently at odds at the IPCC. Other normalizations are possible, but yield poor convergence of the model. Not setting a normalization makes convergence of the model difficult, but recovers the same substantive results as the normalization chosen here.

In order to obtain convergence of the model, the data need to be pruned. Specifically, interventions on statements involved fewer than two countries are excluded from the analysis. This also leads to the exclusion of statements that did not lead to any intervention. Such subsetting of the data is common in ideal point estimation, as it leads to more stable estimates.[55]

I specify the following priors for the model parameters. The ideal point parameters x_i are given a normal prior with mean 0 and standard deviation 1. The discrimination parameters

[53] To maximize the likelihood of the model converging, the data used to estimate the ideal points pools together IPCC negotiations from two different assessment cycles. The estimated ideal points, therefore, do not speak to changes in country preferences over time but rather capture time-invariant features of a country's bargaining position on climate issues. Future research could study how IPCC ideal points change over time, which could be used to study how preferences change as the underlying scientific consensus solidifies.

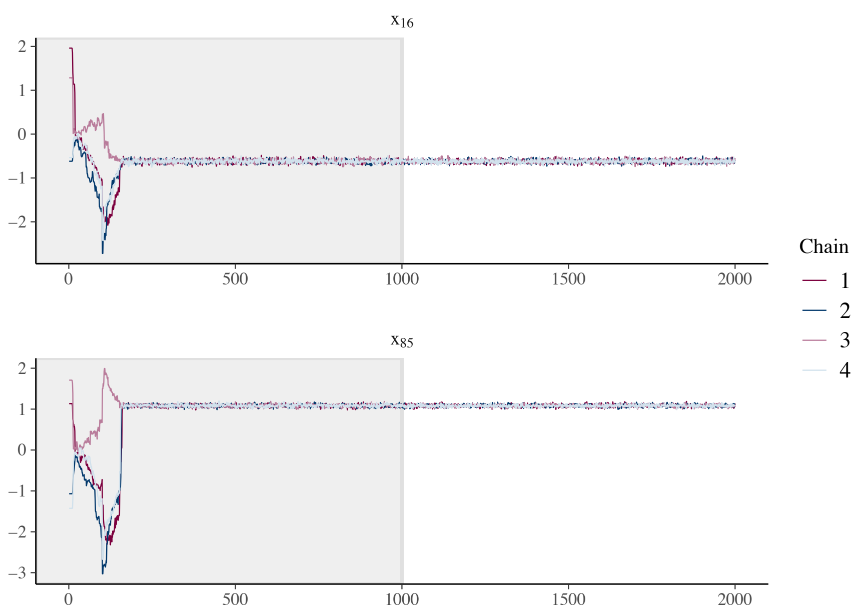
[54] Specifying multiple dimensions yields poor model convergence and returns second dimension ideal points that are highly correlated with the first dimension ideal points. The one dimensional ideal point model can thus be thought of as a parsimonious model that captures the predominant cleavage at the IPCC.

[55] See for example Crosson, Furnas and Lorenz (2020)

β are given a normal prior with mean 0 and standard deviation 5. The cutpoints c_0 and c_1 are given a normal prior with mean 0 and standard deviation 25. The wide prior distribution for the cutpoint parameters was chosen after iterations with narrower priors which led to poor model convergence. This is because the cutpoints need to be wide enough to capture the fact that most countries do not intervene on most statements with in support or against the statement.

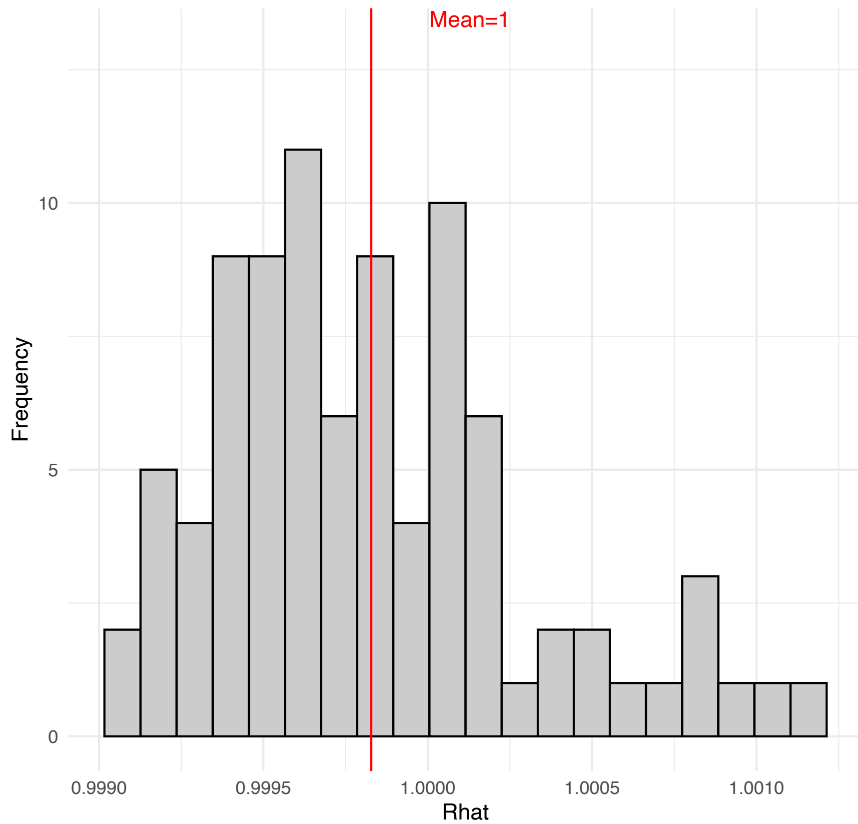
Ideal Point Diagnostics Below, I show two diagnostic plots for the ideal point estimation. First, Figure 8 plots the MCMC draws of the ideal point parameter (x_i) for two countries, China and the US. The chains seem to mix well before the burn-in iteration, suggesting that the MCMC algorithm has converged. Second, Figure 9 shows the distribution of the $\hat{R}$ statistic for the ideal point parameter. The mean of the distribution is 1, which provides further confidence that the MCMC algorithm has converged.

Figure 8: Trace Plots for China (top) and USA (bottom) Ideal Points



Notes: The figure trace plots of the MCMC draws of ideal point parameters for two countries, China and USA.

Figure 9: $\hat{R}$ Statistic



Notes: The figure plots the distribution of $\hat{R}$ statistics from the MCMC estimation.

B.2 Stan Code

```
// Stan code for ideal points model
data {
    int<lower=2> K; // Number of response categories
    int<lower=0> N; // Number of observations
    int<lower=1> D; // Dimensionality of ideal points
    int<lower=1> I; //Num Countries
    int<lower=1> J; //Num Interventions
    int<lower=1, upper=I> ii[N]; //Country for observation n
    int<lower=1, upper=J> jj[N]; //Intervention for observation n
    int<lower=1, upper=K> y[N];
    int<lower=1> Hidx; // Index of country with high ideal point
    int<lower=1> Lodx; // Index of country with low ideal point
```

```
}
parameters {
    row_vector[D] beta[J]; // discrimination parameter of item j
    row_vector[D] x[I]; // ideal point of country i
    ordered[K - 1] c[J]; // item-specific cutpoints
}

model {
    for (d in 1:D) {
        to_vector(x[,d]) ~ normal(0, 1);
        x[Hidx,d] ~ normal(1, 0.0001); // Prior for country with high ideal point
        to_vector(beta[,d]) ~ normal(0, 5);
    }
    x[Lodx,1] ~ normal(-1, 0.0001); // Prior for country with low ideal point
    for (m in 1:J) {
        c[m] ~ normal(0, 25);
    }
    // Logit model
    for (n in 1:N) {
        y[n] ~ ordered_logistic(dot_product(x[ii[n]], beta[jj[n]]) , c[jj[n]]);
    }
    }
```

B.3 Predictors of Ideal Points

Table 6: IPCC Ideal Points Predictors

Dependent Variable:	IPCC Ideal Point				
Model:	(1)	(2)	(3)	(4)	(5)
UN Ideal Point		0.236***		0.251***	0.197***
		(0.026)		(0.029)	(0.045)
Total GHG emissions			-0.053	-0.070***	-0.049**
			(0.043)	(0.022)	(0.020)
GDP per capita					0.135***
					(0.045)
Log GDP					-0.061
					(0.039)
Constant	0.240***	0.220***	0.253***	0.237***	0.256***
	(0.040)	(0.031)	(0.037)	(0.028)	(0.026)
Fit statistics					
Observations	86	85	84	84	84
R^2		0.417	0.045	0.504	0.578
Adjusted R^2		0.410	0.033	0.491	0.556

Heteroskedasticity-robust standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C Additional Analyses and Robustness Checks

C.1 Alternative Explanations

The results presented in Section 4 are also consistent with several alternative hypotheses. First, uncertain statements may be less valuable to states. In this interpretation, interventions over uncertain statements feature less distributional conflict because the value of imposing a government's preferred interpretation of science is lower on topics where science is uncertain. However, Appendix Tables 10 and 11 below show that uncertain statements are no less likely to be contested at the IPCC plenary, nor do they differ in a statistically detectable way in the number of interventions. This is likely because, conditional on the topic, uncertain statements do not differ meaningfully from certain statements in the importance of the issues they deal with.

Second, the analysis above treats a statement's stated uncertainty as exogenous. However, states may attempt to influence the IPCC's designation of uncertainty. To be valid, the main analyses require a statement's designated level of uncertainty to be a function of underlying scientific evidence. This is echoed in the IPCC's guidance to authors, which considers a statement to be most certain when backed up by "multiple, consistent independent lines of high-quality evidence" [56].[57] The IPCC further requires that uncertainty designations be standardized across working groups and reports. However, it is still possible that authors ignore this guidance from the IPCC. This makes it important to test whether statements' degrees of stated uncertainty reflect the presence of multiple lines of evidence underlying them.

To test this, I estimate a statement-level regression of a statement's uncertainty score on the number of unique references it makes to underlying sections of the report. Each statement in the Summaries for Policymakers contains bracketed references to underlying chapters and subsections of chapters of the respective working group report. While imperfect, the number of unique references a statement contains is a rough proxy for the number of underlying lines of evidence that support it. The results, shown in Appendix Table 13 below, show that statements with more references are indeed less uncertain. The coefficient on the number of references is negative and statistically significant at the 1% level. This suggests that report authors follow the IPCC's guidance on uncertainty.

C.2 Illustrative Examples of Mechanism

Having conducted a quantitative analysis in the note showing that uncertain statements feature less distributional conflict, this Appendix section presents some examples that illustrate the mechanism. Specifically, the theory predicts that uncertain science provides large polluters (who may otherwise disagree on distributional issues) to unite on weakening and questioning science.

This is shown clearly in negotiations over IPCC statements that relate, for example, to the relationship between violent conflict and climate change. In the plenary session that negotiated over the Working Group 2 Summary for Policymakers in the 5th Assessment Report, the draft version of a statement stated, with medium confidence, that "Climate change indirectly increases risks from violent conflict in the form of civil war, inter-group violence, and violent protests by exacerbating well-established drivers of these conflicts such as poverty and economic shocks." [58]. This then engendered an intervention by a coalition of large polluters including Saudi Arabia, the United States, and the United Kingdom, who questioned the underlying scientific basis for this statement [59]. Eventually, the published

[56] Mastrandrea et al. (2011)

[57] As (Hughes, 2024, Ch.7) argues, while government input can alter the degree of ambiguity in summary statements, the authors of the summary reports hold final say and judgement over whether such requests are scientifically legitimate.

[58] Field et al. (2014*b*)

[59] ENB (2014)

version contained qualifiers that weakened the statement. For example, "Climate change indirectly increases risks..." was qualified to "Climate change *can* indirectly increases risks..." (emphasis mine). Similarly, the phrase "well-established drivers of these conflicts..." was changed to "*well-documented* drivers of these conflicts" (emphasis mine), suggesting a lower level of consensus in the scientific community [60].

On the other hand, most disagreements in the plenary are explicitly distributional, where countries ideologically opposed to each other in broader areas of international politics disagree with each other. In the same document, for example, conflict took on a more distributional character over the high confidence statement: "Some unique and threatened systems, such as ecosystems and cultures, are at risk from climate change at recent temperatures." [61]. Disagreement revolved around using the term 'recent temperatures,' which referred to the temperature increase from 1986-2005 [62]. While the US and Austria supported the use of the term, Saudi Arabia interjected and preferred to use pre-industrial temperatures as the baseline, which would call attention to the disproportionately larger role of the global North in total emissions since the Industrial Revolution. Saudi Arabia eventually prevailed, and a reference to warming over 1850-1900 was included.

These are only two examples and are not meant to test the hypothesis but rather illustrate the mechanism connecting scientific uncertainty with the type of coalitions and disagreements that arise over climate science. Specifically, they show that ideologically diverse coalitions between large polluters, which are more likely on uncertain science, often attempt to weaken the scientific basis for action in a way that benefits large polluters. On the other hand, coalitions dividing countries along ideological lines feature disagreement on interpretations that have consequences for how emissions burdens should be distributed *among* large polluters.

[60] Field et al. (2014*a*)
[61] Field et al. (2014*b*)
[62] ENB (2014)

C.3 Main Results with Country-Statement Fixed Effects

Table 7: Main Result is Robust to Country-Statement Fixed Effects

Dependent Variable:	Agreement			
Model:	(1)	(2)	(3)	(4)
Uncertainty × UN Ideal Point Difference			0.056*	0.061**
			(0.030)	(0.031)
Uncertainty × Emissions Difference			0.021	0.019
			(0.052)	(0.052)
UN Ideal Point Difference	-0.206***	-0.206***	-0.333***	-0.336***
	(0.028)	(0.028)	(0.084)	(0.087)
Emissions Difference	-0.044	-0.041	-0.087	-0.079
	(0.107)	(0.108)	(0.216)	(0.214)
Fixed Effects and Controls				
Country x Statement FE	✓	✓	✓	✓
Report Section FE	✓	✓	✓	✓
Dyadic Controls	x	x	x	✓
Outcome Mean	0.069	0.069	0.069	0.069
Observations	2,000,964	1,920,786	1,920,786	1,829,871

Clustered (Statement & Dyad) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The table shows the results of a dyad-statement level regression of the dyadic agreement measure on the dyad's divergence in UN ideal points and total emissions, as well as the uncertainty in the statement. The model is identical to the one estimated in Table 3, except that it includes country-statement fixed effects for both countries in a dyad instead of country fixed effects.

C.4 Main Result with Genovese (2014) Country Scores

Table 8: Dyads with Divergent Country Scores less Likely to Conflict over Uncertain Statements

Dependent Variable:	Agreement			
Model:	(1)	(2)	(3)	(4)
Uncertainty × Genovese (2014) Country Score Difference			0.067**	0.067**
			(0.029)	(0.029)
Genovese (2014) Country Score Difference	-0.357***	-0.349***	-0.502***	-0.466***
	(0.059)	(0.060)	(0.106)	(0.102)
Uncertainty		-0.002	-0.083	-0.083
		(0.035)	(0.054)	(0.054)
Fixed Effects and Controls				
Country FE	✓	✓	✓	✓
Report Section FE	✓	✓	✓	✓
Dyadic Controls	x	x	x	✓
Outcome Mean	0.069	0.069	0.069	0.069
Observations	594,090	570,285	570,285	570,285

Clustered (Statement & Dyad) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

Notes: The table shows the results of a dyad-statement level regression of the dyadic agreement measure on the dyad's divergence in country scores as measured by Genovese (2014), as well as the uncertainty in the statement. The unit of analysis is the undirected dyad-statement combination. Dyadic controls include the logged value of trade between dyad members, logged inter-capital distance, and indicators for shared language, shared colonial legacy, common official language, shared land border, and mutual membership in a military alliance. Standard errors are two-way clustered at the statement and dyad level.

C.5 Sample Fit of Main Results

This section calculates measures of in-sample goodness-of-fit for the main results in Table 3. Specifically, I estimate and compare two models. The first model is the main model estimated in Column 4 of Table 3, which regresses the dyadic agreement measure on the divergence in UN ideal points, total emissions, and the uncertainty in the statement. The second model is a reference model that includes only the fixed effects and controls. I then calculate predicted value of the agreement variable and create a three category ordinal variable that takes the value 100 if the predicted value is greater than 50, -100 if the predicted value is less than -50, and 0 otherwise. These correspond to agreement, disagreement, and non-interaction, respectively. I then calculate, for each outcome class, the specificity (proportion of a class correctly classified), the sensitivity (proportion of non-members of a class correctly classified), and the balanced accuracy (the average of the specificity and sensitivity). The results are shown in Table 9 below.

Table 9: Goodness of Fit of Results in Table 3

	Specificity			**Sensitivity**			**Balanced Accuracy**		
Outcome Level	-100	0	100	-100	0	100	-100	0	100
Main Model	1.00	0.36	0.94	0.002	0.94	0.46	0.50	0.65	0.70
Reference Model (fixed effects and controls only)	1.00	0.28	0.97	0.00	0.97	0.36	0.50	0.63	0.67

As the results show, the main model improves on the reference model on most metrics, with the largest gains in model fit coming from the ability of the model to correctly classify instances of agreement. Nevertheless, the improvement is small, suggesting that uncertainty is one among many different factors that affects the prevalence of agreement and disagreement at the IPCC plenary. Nevertheless, the statistically significant and substantially large coefficients in the main regression suggest a strong relationship between uncertainty and distributional conflict at the IPCC.

C.6 Salience of Uncertain Statements

C.6.1 More Uncertain Statements Are No Less Likely to be Contested

Table 10: More Uncertain Statements Are No Less Likely to be Contested

Dependent Variable:	$\mathbb{1}(AnyIntervention)$		
Model:	(1)	(2)	(3)
Uncertainty	-0.030	0.026	0.037
	(0.034)	(0.029)	(0.030)
log(Statement Length)	0.356***	0.066*	0.070*
	(0.027)	(0.035)	(0.038)
Fixed Effects and Controls			
Report FE	x	✓	x
Working Group FE	x	✓	x
Report Section FE	x	x	✓
Outcome Mean	3.18	3.18	3.18
Observations	551	551	551
R^2	0.283	0.488	0.536

Clustered (Statement) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

C.6.2 More Uncertain Statements Have No Fewer Interventions

Table 11: More Uncertain Statements Have No Fewer Interventions

Dependent Variable:	Number of Interventions		
Model:	(1)	(2)	(3)
Uncertainty	-0.110	0.218	0.295
	(0.223)	(0.187)	(0.197)
log(Statement Length)	3.03***	1.22***	1.66***
	(0.286)	(0.354)	(0.431)
Fixed Effects and Controls			
Report FE	x	✓	x
Working Group FE	x	✓	x
Report Section FE	x	x	✓
Outcome Mean	3.18	3.18	3.18
Observations	551	551	551
R^2	0.229	0.319	0.347

Clustered (Statement) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

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

C.6.3 Is Uncertainty Exogenous?

In order to test whether the uncertainty score of a statement is related to the strength of the underlying scientific literature, I regress a statement's uncertainty score on the number of references it makes to the chapters of the underlying report. While imperfect, the number of references a statement makes to the larger underlying report is often a rough proxy for the number of underlying lines of evidence (the term officially used by the IPCC to denote confidence) that support it. As an illustrative example, the draft SPM for Working Group II in the 5th Assessment Report contained four statements under the heading 'vulnerability and exposure', two of which were designated 'very high confidence', one 'high confidence', and one 'medium evidence, high agreement'. These are show in Table 12 below. The number of references for the high or very high confidence statements ranges from 7 to 20. By contrast, the statement with medium evidence and high agreement, relating to the impact of violent conflict on vulnerability, has only 3 references.

Statement	Uncertainty	# References
"Differences in vulnerability and exposure arise from non-climatic stressors and multidimensional inequalities, which shape differential risks from climate change."	very high confidence	7
"Impacts from recent extreme climatic events, such as heat waves, droughts, floods, and wildfires, demonstrate significant vulnerability and exposure of some ecosystems and many human systems to climate variability."	very high confidence	20
"Climate-related hazards constitute an additional burden to people living in poverty, acting as a threat multiplier often with negative outcomes for livelihoods."	High Confidence	7
"Violent conflict strongly influences vulnerability to climate change impacts for people living in affected places"	Medium Evidence, High Agreement	3

Table 12: Example Draft Statements from the 5th Assessment Report

This suggests that the number of references is an imperfect but useful proxy for the strength of the underlying scientific evidence. Table 13 shows the results of regressing a statement's uncertainty score on the number of references it makes to the chapters of the underlying report. Column 3 shows that after controlling for a statement's length and looking at variation within a narrow section of the report, statements with more references are less uncertain. The coefficient on the number of references is negative and statistically significant at the 1% level. The R-squared of the regression is low, underscoring that the number of references is an imperfect and noisy proxy.

However, substantively the coefficient is large, suggesting that a standard deviation increase in the number of references in a statement (around 8 more references) is associated with a one unit decrease in the four-point uncertainty score. This suggests that the strength of the underlying scientific evidence is a strong predictor of a statement's uncertainty score.

Table 13: Statements with More References are Less Uncertain

Dependent Variable:	Uncertainty		
Model:	(1)	(2)	(3)
Number of References	-0.017***	-0.011***	-0.012***
	(0.003)	(0.003)	(0.004)
log(Statement Length)	-0.086**	-0.013	-0.006
	(0.041)	(0.052)	(0.058)
Fixed Effects and Controls			
Report FE	x	✓	x
Working Group FE	x	✓	x
Report Section FE	x	x	✓
Outcome Mean	2.27	2.27	2.27
Observations	528	528	528
R^2	0.069	0.086	0.189

Clustered (Statement) standard-errors in parentheses
*Signif. Codes: ***: 0.01, **: 0.05, *: 0.1*

It is important to note that there may be many instances where very low or very high confidence statements in the underlying report for a working group are not reflected in the uncertainty designations in the Summary for Policymakers (SPM). There are at least three distinct reasons for this. First, since the SPM uncertainty designations are summaries of the underlying report's uncertainty, it is possible that aggregating the uncertainty in a given chapter or section averages out the uncertainty in statements in the underlying report. Second, it may be that highly uncertain (or highly certain) statements are strategically removed (or their uncertainty designations changed) from the SPM due to pressure from governments or self-censoring by report authors. While this cannot be ruled out within the current empirical setting, it is unlikely to be a systematic issue since omission of statements because of their uncertainty is bound to engender opposition at the plenary.