Supplementary Appendix for: See No Evil, Speak No Evil? Morality, Evolutionary Psychology, and the Nature of International Relations

Brian C. Rathbun
University of Southern California
brathbun@usc.edu

Caleb Pomeroy
The Ohio State University
pomeroy.38@osu.edu

July 18, 2021

1 Text Analysis

1.1 Corpora & Preprocessing Details

For the United Nations (UN) speeches, we use the UN General Debate (GD) corpus (Baturo et al 2017) available on Harvard Dataverse at doi.org/10.7910/DVN/0TJX8Y. Each year at the UNGD, delegations deliver an address, often by the heads of state, to the General Assembly. Delegations use these speeches to outline their views and stances on whichever topics they deem important, unconstrained by explicit agenda setting dynamics. The corpus presently contains each statement by each country that participated in the UNGD from 1965 to 2018, for a total of 8,640 speeches.

For internal documents, we turn to the Foreign Relations of the United States (FRUS), a collection of diplomatic documents curated by the US Department of State's Office of the Historian.[1] Entries in the FRUS consist of intra-governmental memoranda, correspondence, and cables, typically previously classified and originating from cabinet settings and communications with US diplomatic missions (Katagiri and Min 2015: 3). These contemporaneous, more candid documents thus provide an invaluable internal complement to the public-oriented UNGD speeches. Although text analysts in IR are only at the earliest stages of collecting and processing the FRUS collection in a format amenable to quantitative text analysis, we found two previous projects with a sizable sample of FRUS documents available. First, we draw on Lauretig (2019), which includes all available FRUS entries from 1964 to 1966 (excluding editorial entries like lists of abbreviations), for a total of 7,446 documents. Second, we draw on Katagiri and Min (2015: 4), who made available 8,474 documents spanning 1952-1977, focusing on documents "centered on the Soviet Union and the Eastern Bloc during the Cold War." Thus, together, our corpus of FRUS documents provides coverage of a comprehensive set of FRUS entries for a limited period of time, as well as documents centered on the Soviet Union but spanning several decades.

[1] See the State Department's website for more information about the FRUS: history.state.gov/historicaldocuments/about-frus.

To obtain word embeddings, we locally fit GloVe models to these political corpora. We minimally preprocess the texts, converting all terms to lowercase and removing numbers and punctuation. We also stem the terms, because this appeared to improve the quality of the embeddings. For each corpus, we retain only those words that occur at least 20 times in the entire corpus. To fit the GloVe models, we use a context window of 10 (i.e. 10 words before and 10 words after the target feature), fit the model to 200 dimensions, and use a maximum term co-occurrence of 15 for the weighting function. The resulting "main" and "context" vectors are then averaged to locate the final embedding space, per the original GloVe paper's suggestion (Pennington et al 2014).

We compare these locally-trained embeddings from the political corpora to pre-trained embeddings from GloVe's website, available at nlp.stanford.edu/projects/glove/. The embeddings were trained on Wikipedia (in 2014) and Gigaword 5, with a total of 6 billion tokens and a 400,000 word vocabulary. We use the embeddings trained in 200 dimensions, the same dimension size as our locally-trained embeddings. See Pennington et al (2014) for more about GloVe.

1.2 Additional Text Analyses and Robustness Checks

1.2.1 Democracy and Threat Embeddings Analysis

In the main text, we present embedding results that show that positive and negative moral terms systematically vary on the dimensions of harm and threat. As a check on the quality of the embedding space, here we compare threat dimension projections for the UN and quotidian texts to a second theoretically important dimension: regime type. To construct a democracy versus non-democracy dimension, we contrast the vector space locations of different words that describe regime type (e.g. democracy-autocracy, democracy-dictator). We then project moral terms onto this regime type dimension using the same procedure as described in the main text. The intuition is as follows: if our embedding space captures meaningful semantic patterns (rather than simple noise), then we should also find that moral terms

sensibly vary along other important theoretical dimensions. This would suggest that the moral terms do not simply vary as an idiosyncratic function of our particular harm and threat dictionaries.

Figure A1 presents the results. Words further to the right on the democracy dimension tend to co-appear more often with terms describing democracies, and words further to left on the democracy dimension tend to co-appear with terms describing non-democracies. As expected, positive moral traits tend to cluster on the positive end of the dimension, and vice versa. The moral terms vary as much, if not more, with the regime type dimension compared to the threat dimension. Further, these results are similar across elite speeches and quotidian human texts. These results suggest that the embedding spaces capture meaningful semantic and theoretical patterns, beyond our immediate focus on harm and threat.

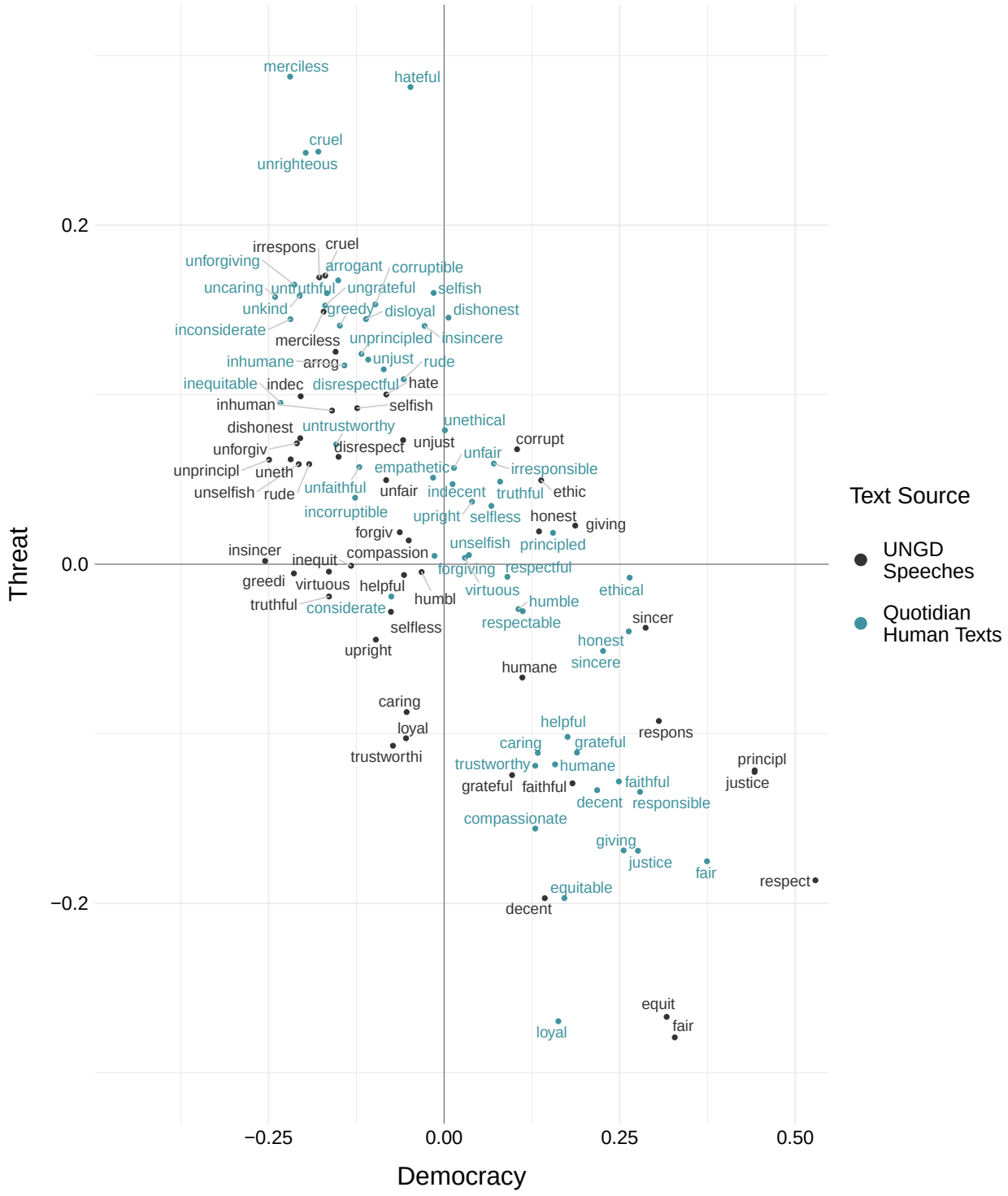

Figure A1: *Threat versus democracy embedding dimensions.* Projection of moral terms onto a regime type dimension (crossed with a threat dimension for comparison) shows that moral terms vary sensibly with the theoretically important dimension of regime type.

1.2.2 Embeddings Robustness Checks

In the main text, we present the projections of moral terms onto harm and threat dimensions based on the average cosine similarity (from 20 separate embedding models) between these terms and the harm and threat dimensions. To assess the robustness of the results, we follow Kozlowski et al's (2019) approach to construct nonparametric confidence intervals for the political corpora results, as well report a permutation test for our quotidian results from the pretrained GloVe vectors.

For both the FRUS and UNGA corpora, we construct 20 separate corpora by sampling with replacement from the full corpora, with each newly sampled corpus containing the same number of documents as the original corpus. For each of these resampled corpora, we then conduct the same projection procedure described in the main text, calculating the cosine similarity between the moral terms and the harm and threat dimensions (which are constructed by taking the average vector space locations of harm-related and threat-related words, and subtracting off the average locations of harm and threat antonyms, respectively). This results in 20 separate cosine similarities for each moral term, on each dimension. To construct a 90% nonparametric confidence interval, then, we take the 2nd and 19th order statistics (i.e. the 2nd smallest and 19th largest cosine similarity) to span the 5th and 95th percentiles of a given cosine similarity estimate.

Kozlowski et al (2019: 935) summarize the intuition: "If a word occurs only rarely or is used in a diffuse set of very distinct contexts, the word's position in the vector space will be radically different between subsamples and therefore will produce larger confidence or credible intervals. On the other hand, words that are frequently used in consistent contexts will hold more stable positions across the subsamples and hence produce smaller confidence or credible intervals."

Figure A2 and A3 present the results. For the political corpora, 56.0% of the FRUS terms and 59.5% of the UNGA terms fall reliably far from zero. Interestingly, the similarity of these rates suggests that

moralization of harm and threat occurs similarly in both public speeches and private communications. Importantly, we only reliably misclassify three terms in the political corpora ("virtuous" on the UNGA's harm dimension, "ethic" on the UNGA's threat dimension, and "honest" on the FRUS's threat dimension). Given that these intervals are quite conservative, and given the relatively small size of the underlying corpora, these results increase our confidence that moral terms are surprisingly non-orthogonal to these dimensions.

Finally, because we use pretrained GloVe vectors for the quotidian texts, we do not have access to the underlying documents to conduct the above resampling procedure. Thus, we instead use a simple permutation test to construct null distributions of cosine similarities for each term and then compare each of our observed similarities to these null distributions. For each moral term, we sample without replacement from the term's estimated vector space coordinates to construct a new, "null" vector of the same length as the original vector (i.e. a new 1×200 vector). We repeat this process 2,000 times to construct a null distribution of cosine similarities for a given term. Then, we calculate whether the term's observed cosine similarity is larger than at least 90% of the null cosine similarities for moral terms expected to fall on the positive end of the dimensions or smaller than at least 90% of the null cosine similarities for moral terms expected to fall on the negative end of the dimensions. We find that 59.3% of the moral terms significantly diverge from a null distribution, and perhaps more importantly, we do not misclassify any of these more robust terms on the wrong side of the dimensions.

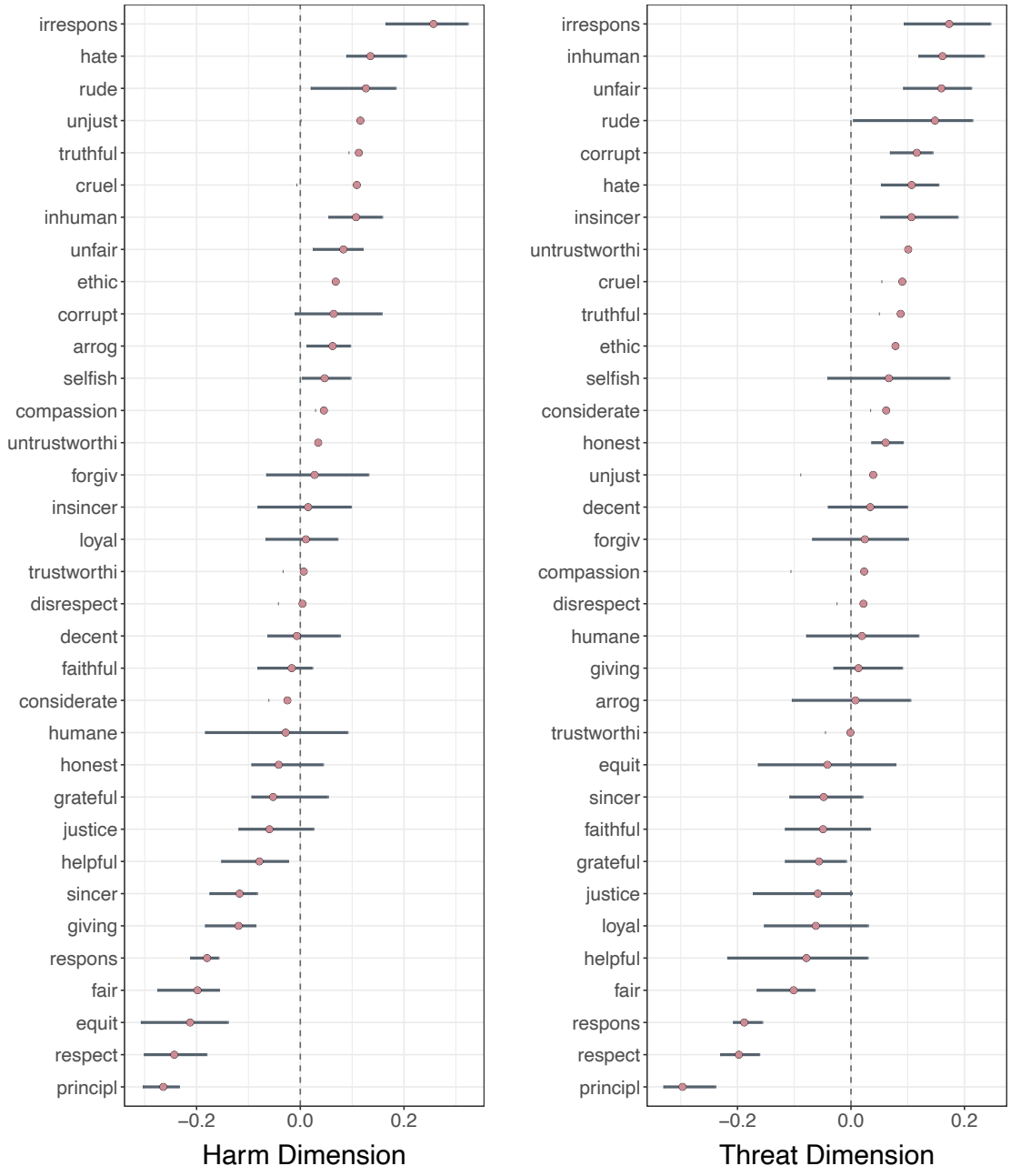

Figure A2: *Nonparametric confidence intervals (FRUS embeddings)*. Twenty separate embedding models fitted to resampled FRUS texts, with 90% nonparametric intervals formed by taking the 2nd smallest and 19th largest cosine similarity. Point estimates represent the mean cosine similarity for each term across the multiple models. Any term that does not include an interval was not present in each of twenty resamples (i.e. the term was too rare to reliably estimate a CI).

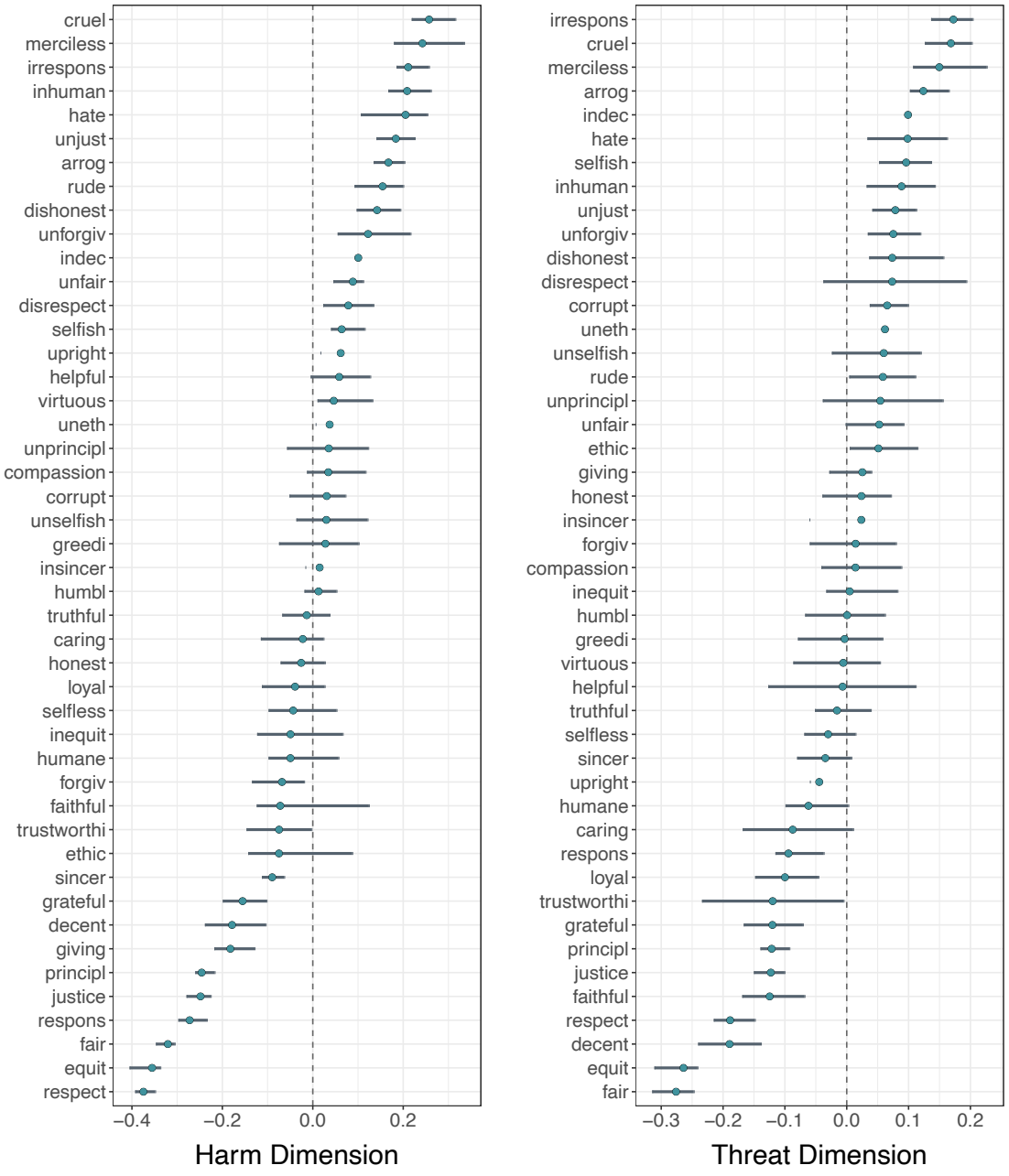

Figure A3: *Nonparametric confidence intervals (UNGA embeddings)*. Twenty separate embedding models fitted to resampled UNGA texts, with 90% nonparametric intervals formed by taking the 2nd smallest and 19th largest cosine similarity. Point estimates represent the mean cosine similarity for each term across the multiple models. Any term that does not include an interval was not present in each of twenty resamples (i.e. the term was too rare to reliably estimate a CI).

1.2.3 Uncombined FRUS Results

Figure 1 of the main text reports harm and threat projection results from embedding models fitted to a combined version of the FRUS corpora. That is, for the two separate collections of FRUS documents that we have access to (one focused on the Soviets from 1952-77, the other capturing the universe of FRUS documents available from 1964-66), we simply fit the embedding models to all available documents to increase the sample size of our underlying FRUS collection and to streamline presentation of the results. Here, we show that those findings are robust to fitting separate embedding models to the two different collections of FRUS documents, rather than combining the documents into a single corpus. Figure A4 plots the results, which show that our substantive results reported in the main text are unchanged: negatively-valenced moral terms appear on the positive ends of the harm and threat dimensions, suggesting that harm and threat speech contains negative moral content, whereas positively-valenced moral terms appear on the negative ends of these dimensions.

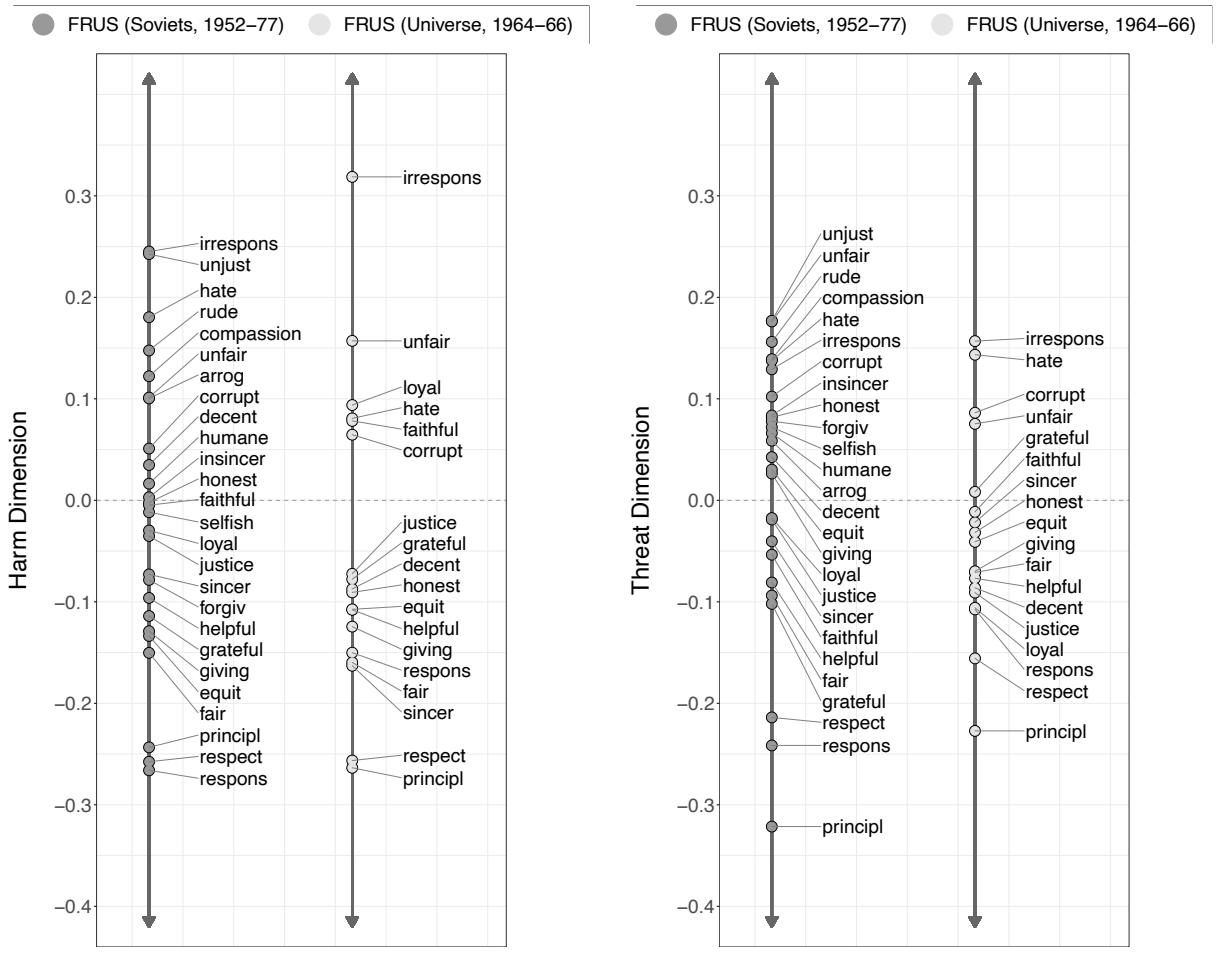

Figure A4: *Uncombined FRUS Results.* Moral terms projected onto "harm" and "threat" dimensions in vector space for the individual FRUS corpora. The results show the same substantive pattern as the combined results presented in the main text.

2 Russian "Build a Threat"

The "build a threat" experiment – fielded in early March 2020 on a sample of the Russian mass public ($N = 1{,}245$) – randomly assigned subjects to one of two conditions: subjects selected traits useful for the determination of threat from a *state* or *individual*. The objective of this experiment was to assess whether those traits systematically vary according to the actor type in question, with the hypothesis that

subjects would select moral traits as the most important predictor of threat and that those traits would show few differences between interpersonal versus interstate comparisons.

2.1 Instrumentation

Here, we present the exact wording of the prompts in Russian, alongside English translations. The translations were conducted by native Russian speakers.

2.1.1 Interpersonal condition

Original Russian text: На протяжение всей жизни мы формируем мнение о других людях и о том, могут ли они причинить нам вред. Мы никогда не знаем наверняка, но, представьте, что вам необходимо сформировать свое мнение о человеке и вам дается надежная информация о ряде качеств этого человека. Какие качества кажутся вам наиболее важными и существенными для формирования вашего мнения об этом человеке? Другими словами, о каких качествах вы бы хотели узнать?

English translation: In our lives, we must form impressions of others and whether they might harm us or not. We never know for sure, but if you were asked to form a judgment about someone and were offered reliable information about the following traits, which would you find to be most relevant and most important? In other words, which would you want to know?

2.1.2 Interstate condition

Original Russian text: В вопросах внешней политики, лидеры наших стран формируют мнение о других странах и о том, могут ли они причинить вред нашей стране. Мы никогда не знаем наверняка, но, представьте, что вам необходимо сформировать свое мнение о другом

государстве и вам дается надежная информация о ряде качеств жителей этой страны и ее лидеров. Какие качества кажутся вам наиболее важными и существенными для формирования вашего мнения об этой стране? Другими словами, о каких качествах вы бы хотели узнать?

English translation: In foreign policy, our leaders must form impressions of other countries and whether they might harm us or not. We never know for sure, but if you were asked to form a judgment about another country and were offered reliable information about the following characteristics of that country's people and its leaders, which would you find to be most relevant and most important? In other words, which would you want to know?

2.2 Sample Characteristics

Our Russia sample, recruited through the survey firm Anketolog, yielded the following demographic profile, presented in Figure A5. Table A1 compares the sample profile to the broader Russian public, suggesting that the sample is relatively representative of the broader Russian public. Population data comes from the Russian Federation Federal State Statistical Service (Rosstat), accessed in May 2020: https://www.gks.ru/.

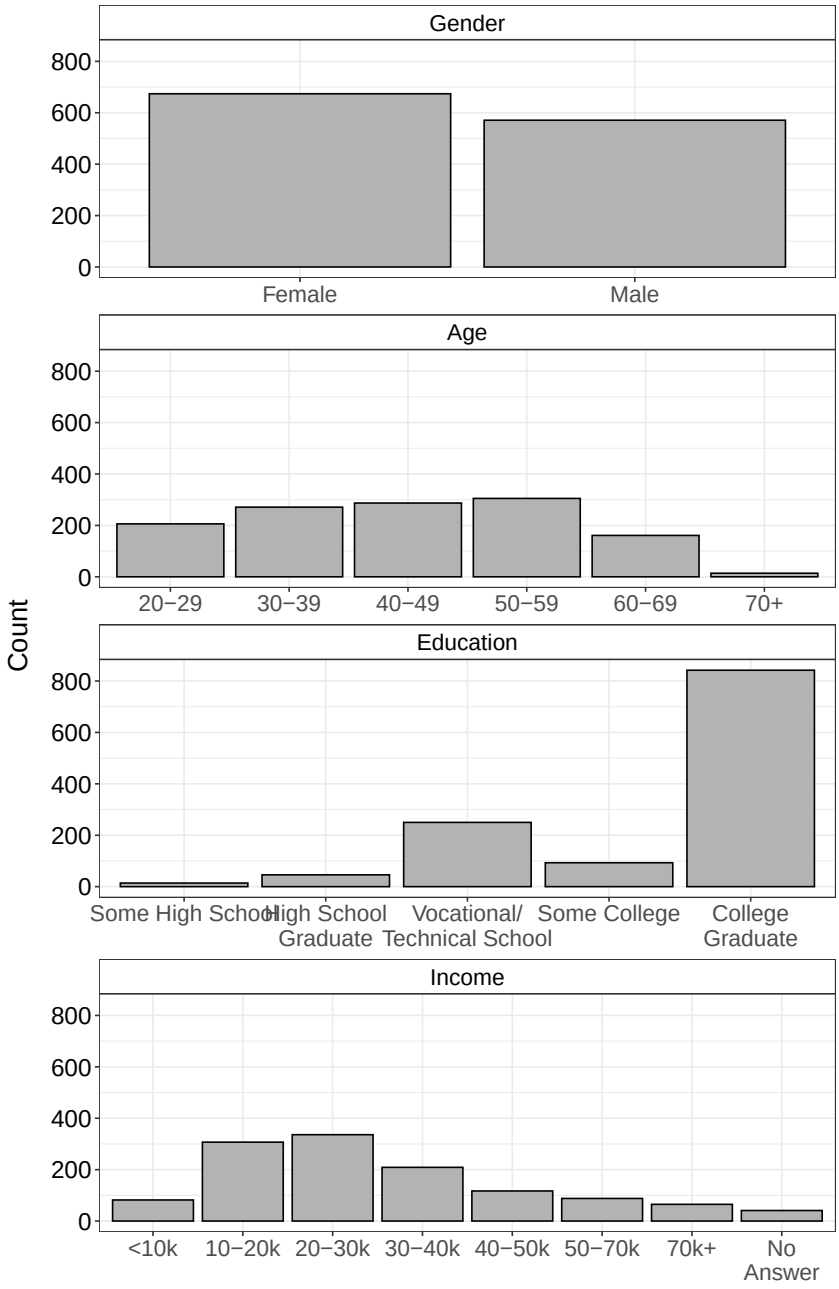

Figure A5: *Russian Sample Characteristics,* $N = 1{,}245$. Income is listed in rubles.

Table A1: Survey sample characteristics

	Adult population	Sample
Male	46%	46%
Age 18-30	20%	19.30%
Age 31-45	29%	29.80%
Age 46-55	27%	28.60%
Age 55+	24%	22.40%
Less than 10000 rub	12.0%	6.50%
10-20000 rub	25.0%	25.40%
20-30000 rub	22%	25.30%
30-40000 rub	19%	17%
40-50000 rub	10%	9.50%
50-70000 rub	7%	8.30%
Above 70000 rub	5%	5%

3 Chinese Security Perceptions

Our Russia survey provides the advantage of having respondents choose from a carefully selected set of potentially threat-relevant traits. However, in order to judge respondents' use of different traits at the interpersonal and international level, we posed the problem generally, without specific country names. In order to increase confidence in the external validity of the findings, here we present supplementary observational evidence of those same tendencies in a population of significance to IR researchers.

We draw on two different surveys of the Chinese public: the Beijing Area Study (BAS) survey and the US-Chinese Security Perceptions Project of the Carnegie Foundation (USCSPP). The BAS is a random sample survey of opinion in the Beijing municipality administered by the Research Center for Contemporary China at Peking University ($N = 1,410$). The 2007 version of the BAS asks respondents to characterize the US, Japan, and China itself along continua that present a forced choice between contrasting end points central to our theoretical arguments: peace-loving or warlike, moral or immoral, modest or arrogant, sincere or insincere, and civilized or barbaric.[2] Further, the USCSPP survey was conducted May 2–July 5, 2012, among 2,597 adults in urban areas and used the same sampling technique as the BAS.

If the results from the Russia screening experiment extend to named countries, then perceptions of immorality should correlate positively with perceptions of threat. Figure A6(A) presents correlations between perceptions that a country is "warlike" and negative moral attributions. Each correlation is significant ($p < 0.001$) and substantively large (ranging between $\rho = 0.45$ and $\rho = 0.76$), including for the perception of China by the Chinese themselves.

[2]For previous IR work that employs data from the BAS, see e.g. Johnston (2017).

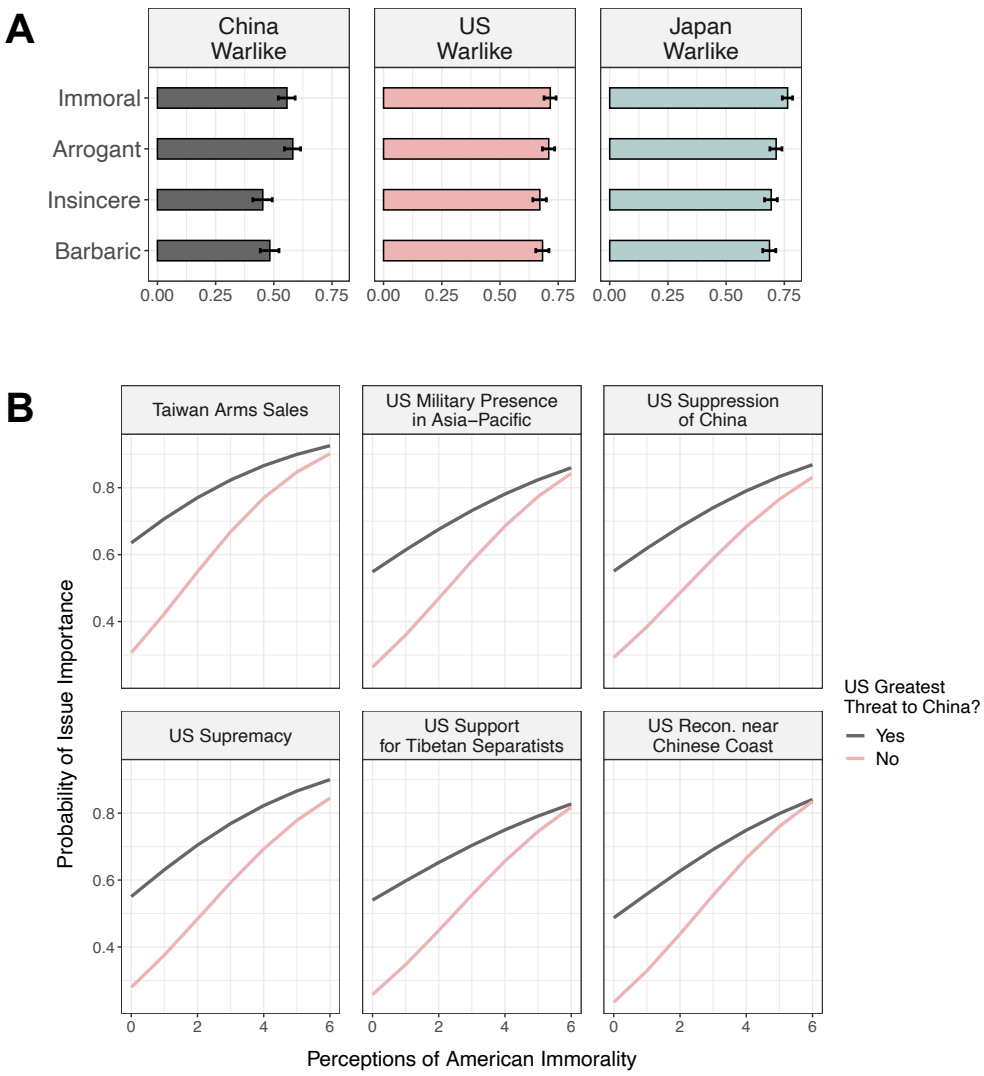

Figure A6: *Morality and Chinese security perceptions.* In (A), negative moral attributions correlate positively with the perception that a given country is "warlike." In (B), fitted probabilities from logistic regressions show that perceptions of American immorality and perceptions of American threat correlate with a higher likelihood that subjects attach importance to a host of US-China security issues. The slight negative immorality by threat interaction suggests that subjects concerned with threats other than the US, but that view Americans as immoral, more quickly attach importance to US-China security issues.

The USCSPP asks the Chinese mass public whether they associate a number of traits with Americans – greedy, selfish, deceitful, honest, generous, and tolerant – each gathered on a "yes/no" dichotomous scale. Furthermore, the survey collected respondents' perceptions of the importance of a host of security issues, gathered on 4-point scales that range from "very important" to "not at all important." Here, we use responses to these attribute questions to construct an immorality scale by summing the number of

negative ethical traits subjects associate with Americans, as well as the reverse-coded positive traits. We also control for whether respondents perceive the US to be China's largest security threat to establish whether, even if a respondent does not believe the US poses the greatest danger to China, immorality can provide an independent pathway to explain variance in security issue importance.

Figure A6(B) displays the results of logistic regressions that estimate the relationship between the importance respondents attach to a host of security issues (the DVs, with responses of "very serious problem" and "somewhat serious problem" coded as 1, and "not a serious problem" or "not a problem at all" coded as 0) and perceptions of US threat, immorality, and a threat-immorality interaction (the IVs). Across issues, we naturally find that respondents who believe the US presents the greatest danger to China also attach greater importance to security issues that involve the US. In addition, however, subjects who believe the US is not China's greatest security threat – but do believe the US is immoral – eventually settle at security positions nearly identical to those respondents who do believe that the US represents China's greatest threat. In fact, the slight negative interactive relationship suggests that those respondents concerned with threats other than the US are precisely the population most affected by the moderating effects of perceived American immorality: those subjects are quicker to attach importance to US security issues by virtue of perceived American immorality.

Because we also show that moral attributions undergird threat perception, we recognize the potential for post-treatment bias in these models. Thus, below, we present regressions that more straightforwardly model the effect of perceived American immorality on the importance respondents attach to each security issue.

3.1 Peking University Beijing Area Study

All Chinese survey data were obtained, with thanks, from Alastair Iain Johnston at Harvard University.

3.1.1 Instrumentation

Below, we present the relevant subset of questions from the Beijing Area Study (BAS) used in the above analysis. Each participant also responded to these same questions with Japanese and American people substituted for "Chinese people." Responses were gathered on a 7-point Likert scale with an option for "don't know." Translations were conducted by native Chinese speakers. For more on the BAS, see Johnston (2017). The sample's demographic profile closely mirrors Beijing's population on the dimensions of gender and per capita income, though BAS participants tend to be better educated than the average Beijing resident, and the sample may not be representative of China's national-level demographic profile (Johnston 2017: 16).

Original Chinese text:

1. 您认为从根本上说，中国人的本性是爱好和平的，还是好战的？
2. 您认为从根本上说，中国人的本性是有道德感的，还是没有道德感的？
3. 您认为从根本上说，中国人的本性是谦和的，还是傲慢的？
4. 您认为从根本上说，中国人的本性是言行一致的，还是言行不一的？
5. 您认为从根本上说，中国人的本性是文明的，还是野蛮的？

English translation:

1. Do you basically believe that Chinese people are essentially peace-loving or warlike?

2. Do you basically believe that Chinese people are essentially moral or immoral?

3. Do you basically believe that Chinese people are essentially modest or arrogant?

4. Do you basically believe that Chinese people are essentially sincere or insincere?

5. Do you basically believe that Chinese people are essentially civilized or barbaric?

3.2 Carnegie Security Perceptions Project

3.2.1 Instrumentation

The second set of observational survey data derives from the U.S.-Chinese Security Perceptions Project (USCSPP) administered by the Carnegie Foundation and Peking University. In order to construct an index of perceived American immorality, we utilize the following questions from the USCSPP.[3] Each question starts with "Which of these characteristics do you associate with Americans?" followed by the following traits: greedy, selfish, deceitful, generous, honest, and tolerant.[4] The latter three are reverse coded in order to place all traits on a scale of immorality. Higher values correspond to higher levels of perceived American immorality. Responses are gathered with options for "yes, associate" or "no, do not associate" with an option for "don't know."

For the specific security issue DVs, we use the questions with the following wording. Each question begins with the phrase "Is the following a very serious problem, a somewhat serious problem, not a serious problem or not a problem at all," followed by six issues relevant to our purposes: "U.S. arms sales to Taiwan," "U.S. military presence in Asian-Pacific region," "U.S. suppression of China's potential," "U.S. supremacy," "U.S. support of Tibetan separatists," and "U.S. military recon activities near Chinese coastal areas."

[3] These questions correspond to survey questions b9ac, b9ae, b9ap, b9aa, b9ad, b9ag, respectively.
[4] "Deceitful" was worded as "One's acts belie one's words."

3.2.2 Sample Characteristics

Figure A7 displays the demographic profile associated with the USCSPP data utilized in our analyses.

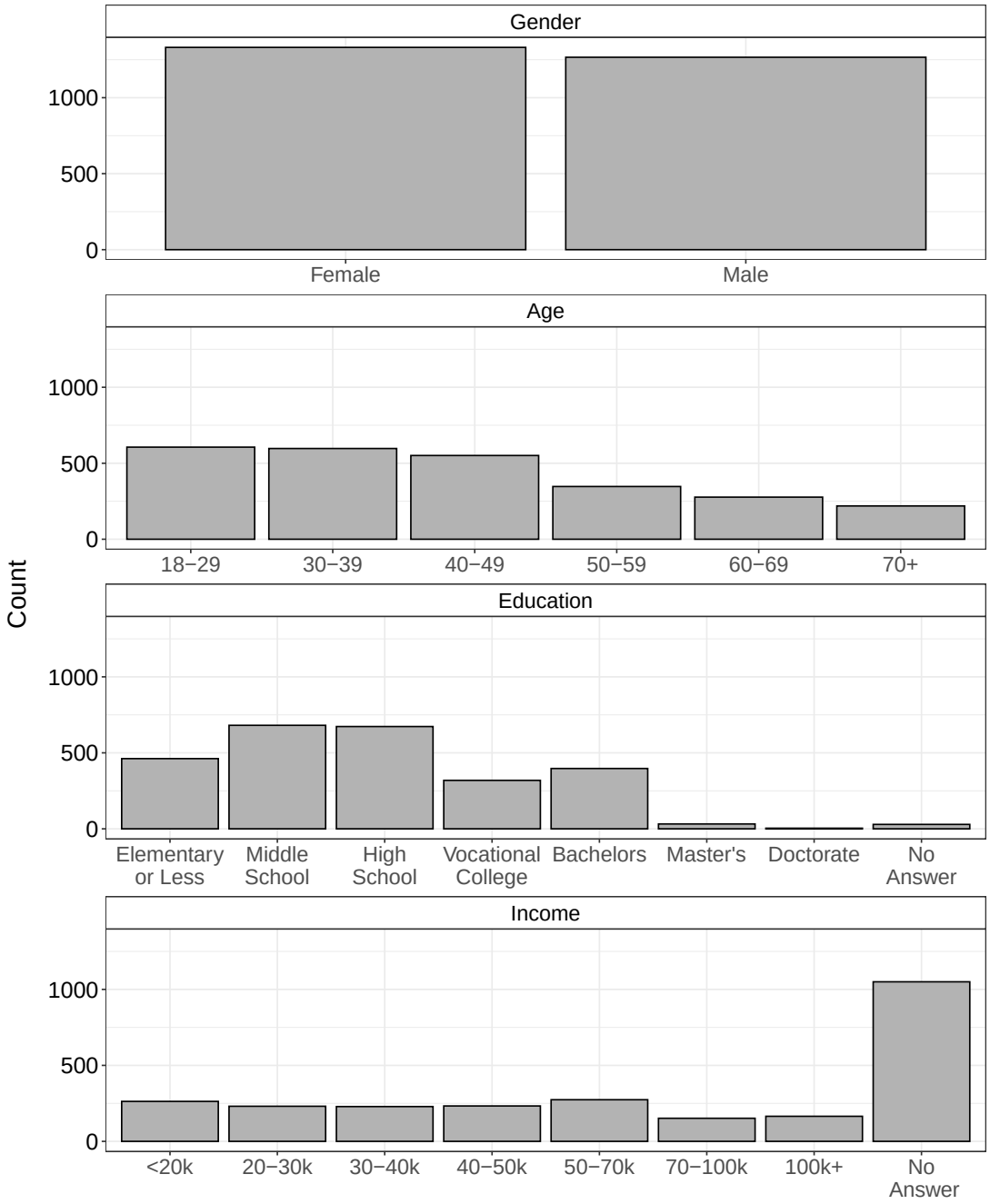

Figure A7: *USCSPP Sample Characteristics,* $N = 2{,}597$. Income is listed in yuan.

3.2.3 Security Issue Regression Tables

Above, we visually present the results of an interaction between perceptions of American immorality and beliefs that the US poses the greatest threat to China. We present this moderation relationship to highlight that perceptions of immorality correlate with increases in the importance that respondents attach to a host of security issues which involve the US. Here, Table A2 presents the full regression table, including control variables. With the addition of the control variables, one interactive relationship loses significance: worries about "US supremacy." However, beliefs about "US supremacy" and beliefs about the US posing the greatest threat to China likely capture very similar phenomena, so the instability in the moderating relationship here is somewhat unsurprising. Further, Table A3 presents the same models with the interaction term omitted in order to ease interpretation of the "greatest threat" and "immorality" main effects. Finally, because we also show that morality informs threat perception, modeling immorality and perceptions that the US poses the greatest threat to China together as IVs might lead to post-treatment bias issues. Therefore, Table A4 presents regressions that model the effect of immorality on security issue importance, without the "greatest threat" covariate. Again, the results suggest that perceptions of American immorality strongly explain the importance that respondents attach to each security issue.

Table A2: Logistic Regressions (with Interaction Relationship): Morality, Threat Perception, and Security Issue Importance

	Taiwan Arms		**Military Presence**		**Suppression of China**		**US Supremacy**		**Tibetan Support**		**Maritime Reconnaissance**	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
(Intercept)	-0.81^{***}	-1.08^{***}	-1.03^{***}	-1.42^{***}	-0.88^{***}	-0.88^{***}	-0.95^{***}	-0.95^{***}	-1.06^{***}	-1.17^{***}	-1.18^{***}	-1.04^{***}
	(0.10)	(0.18)	(0.11)	(0.18)	(0.10)	(0.17)	(0.11)	(0.18)	(0.11)	(0.17)	(0.11)	(0.17)
US Greatest Threat	1.37^{***}	1.25^{***}	1.22^{***}	1.12^{***}	1.09^{***}	0.97^{***}	1.15^{***}	1.04^{***}	1.22^{***}	1.12^{***}	1.13^{***}	1.03^{***}
	(0.15)	(0.15)	(0.14)	(0.15)	(0.14)	(0.15)	(0.14)	(0.15)	(0.14)	(0.15)	(0.14)	(0.15)
US Immorality	0.51^{***}	0.47^{***}	0.45^{***}	0.42^{***}	0.41^{***}	0.38^{***}	0.44^{***}	0.41^{***}	0.43^{***}	0.40^{***}	0.47^{***}	0.43^{***}
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
US Greatest Threat × US Immorality	-0.18^{***}	-0.15^{**}	-0.18^{***}	-0.15^{**}	-0.13^{**}	-0.10^{*}	-0.11^{*}	-0.08	-0.19^{***}	-0.17^{***}	-0.18^{***}	-0.15^{**}
	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.05)	(0.04)	(0.05)	(0.05)	(0.05)
Male		0.66^{***}		0.66^{***}		0.41^{***}		0.37^{***}		0.39^{***}		0.43^{***}
		(0.10)		(0.09)		(0.09)		(0.09)		(0.09)		(0.09)
Age: 30-50		-0.24		-0.11		-0.14		-0.12		-0.25^{*}		-0.14
		(0.13)		(0.12)		(0.12)		(0.12)		(0.12)		(0.12)
Age: 50+		-0.63^{***}		-0.39^{**}		-0.65^{***}		-0.67^{***}		-0.59^{***}		-0.61^{***}
		(0.14)		(0.13)		(0.13)		(0.13)		(0.13)		(0.13)
Vocational College		0.22		0.50^{**}		0.41^{**}		0.35^{*}		0.52^{***}		0.44^{**}
		(0.17)		(0.15)		(0.15)		(0.16)		(0.15)		(0.15)
Bachelors		0.45^{**}		0.81^{***}		0.48^{***}		0.49^{***}		0.54^{***}		0.48^{***}
		(0.16)		(0.15)		(0.14)		(0.15)		(0.14)		(0.14)
Graduate Education		1.05		0.82		0.52		0.45		0.58		0.10
		(0.58)		(0.46)		(0.44)		(0.45)		(0.42)		(0.39)
No Education Answer		-0.13		-0.55		-0.22		-0.09		0.03		-0.14
		(0.43)		(0.42)		(0.41)		(0.42)		(0.41)		(0.41)
Income: 30-50k		0.41^{*}		0.31^{*}		0.25		0.37^{*}		0.31^{*}		-0.07
		(0.16)		(0.15)		(0.15)		(0.16)		(0.15)		(0.15)
Income: 50-80k		0.68^{***}		0.35^{*}		0.31		0.36^{*}		0.33^{*}		0.04
		(0.19)		(0.17)		(0.17)		(0.17)		(0.16)		(0.16)
Income: 80k+		1.00^{***}		0.56^{**}		0.44^{*}		0.43^{*}		0.37^{*}		0.30
		(0.24)		(0.20)		(0.20)		(0.20)		(0.19)		(0.19)
No Income Answer		0.14		0.11		-0.12		-0.12		0.10		-0.31^{*}
		(0.13)		(0.13)		(0.12)		(0.13)		(0.12)		(0.12)
Log Likelihood	-1294.08	-1225.87	-1486.92	-1412.87	-1484.62	-1427.99	-1419.78	-1364.79	-1541.01	-1489.12	-1526.39	-1471.54
Deviance	2588.15	2451.73	2973.84	2825.74	2969.24	2855.98	2839.56	2729.57	3082.02	2978.25	3052.78	2943.07
Num. obs.	2597	2597	2597	2597	2597	2597	2597	2597	2597	2597	2597	2597

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Note: for gender, female serves as the baseline. For age, 18-29 serves as the baseline. For education, high school or less serves as the baseline. For income, less than 20k yuan serves as the baseline. We include covariates for "no income answer" and "no education answer" because a large number of respondents did not report these answers, and omitting these observations would drastically decrease the sample size.

Table A3: Logistic Regressions (without Interaction Relationship): Morality, Threat Perception, and Security Issue Importance

	Taiwan Arms	Military Presence	Suppression of China	US Supremacy	Tibetan Support	Maritime Recon.
(Intercept)	-0.91^{***}	-1.21^{***}	-0.74^{***}	-0.85^{***}	-0.94^{***}	-0.83^{***}
	(0.17)	(0.16)	(0.16)	(0.16)	(0.16)	(0.16)
US Greatest Threat	0.94^{***}	0.75^{***}	0.73^{***}	0.86^{***}	0.71^{***}	0.65^{***}
	(0.10)	(0.09)	(0.09)	(0.10)	(0.09)	(0.09)
US Immorality	0.39^{***}	0.33^{***}	0.32^{***}	0.36^{***}	0.30^{***}	0.34^{***}
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Male	0.67^{***}	0.66^{***}	0.41^{***}	0.38^{***}	0.41^{***}	0.44^{***}
	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Age: 30-50	-0.24	-0.12	-0.15	-0.12	-0.25^{*}	-0.14
	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Age: 50+	-0.63^{***}	-0.39^{**}	-0.65^{***}	-0.67^{***}	-0.59^{***}	-0.61^{***}
	(0.14)	(0.13)	(0.13)	(0.13)	(0.12)	(0.13)
Vocational College	0.23	0.51^{***}	0.41^{**}	0.36^{*}	0.53^{***}	0.45^{**}
	(0.17)	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)
Bachelors	0.45^{**}	0.80^{***}	0.48^{***}	0.49^{***}	0.54^{***}	0.48^{***}
	(0.16)	(0.15)	(0.14)	(0.15)	(0.14)	(0.14)
Graduate Education	1.04	0.81	0.52	0.45	0.58	0.11
	(0.58)	(0.46)	(0.44)	(0.44)	(0.41)	(0.39)
No Education Answer	-0.11	-0.52	-0.21	-0.08	0.05	-0.12
	(0.42)	(0.42)	(0.41)	(0.41)	(0.40)	(0.41)
Income: 30-50k	0.41^{*}	0.30^{*}	0.24	0.37^{*}	0.31^{*}	-0.07
	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)
Income: 50-80k	0.69^{***}	0.37^{*}	0.33	0.37^{*}	0.35^{*}	0.06
	(0.19)	(0.17)	(0.17)	(0.17)	(0.16)	(0.16)
Income: 80k+	1.02^{***}	0.58^{**}	0.46^{*}	0.44^{*}	0.39^{*}	0.32
	(0.24)	(0.20)	(0.20)	(0.20)	(0.19)	(0.19)
No Income Answer	0.13	0.10	-0.12	-0.13	0.08	-0.32^{**}
	(0.13)	(0.13)	(0.12)	(0.13)	(0.12)	(0.12)
Log Likelihood	-1229.81	-1418.22	-1430.39	-1366.15	-1495.96	-1476.85
Deviance	2459.63	2836.43	2860.78	2732.31	2991.93	2953.71
Num. obs.	2597	2597	2597	2597	2597	2597

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Note: for gender, female serves as the baseline. For age, 18-29 serves as the baseline. For education, high school or less serves as the baseline. For income, less than 20k yuan serves as the baseline. We include covariates for "no income answer" and "no education answer" because a large number of respondents did not report these answers, and omitting these observations would drastically decrease the sample size.

Table A4: Logistic Regressions (without "greatest threat" predictor): Morality and Security Issue Importance

	Taiwan Arms	Military Presence	Suppression of China	US Supremacy	Tibetan Support	Maritime Recon.
(Intercept)	-0.52^{**}	-0.87^{***}	-0.43^{**}	-0.48^{**}	-0.62^{***}	-0.54^{***}
	(0.16)	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)
US Immorality	0.42^{***}	0.35^{***}	0.34^{***}	0.39^{***}	0.32^{***}	0.36^{***}
	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Male	0.71^{***}	0.69^{***}	0.45^{***}	0.42^{***}	0.43^{***}	0.47^{***}
	(0.10)	(0.09)	(0.09)	(0.09)	(0.09)	(0.09)
Age: 30-50	-0.21	-0.09	-0.13	-0.10	-0.23^{*}	-0.12
	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Age: 50+	-0.63^{***}	-0.40^{**}	-0.65^{***}	-0.66^{***}	-0.60^{***}	-0.61^{***}
	(0.14)	(0.13)	(0.13)	(0.13)	(0.12)	(0.12)
Vocational College	0.27	0.53^{***}	0.44^{**}	0.39^{*}	0.56^{***}	0.47^{**}
	(0.17)	(0.15)	(0.15)	(0.16)	(0.15)	(0.15)
Bachelors	0.42^{**}	0.77^{***}	0.45^{**}	0.46^{**}	0.52^{***}	0.46^{***}
	(0.16)	(0.15)	(0.14)	(0.14)	(0.14)	(0.14)
Graduate Education	0.96	0.77	0.48	0.41	0.54	0.08
	(0.57)	(0.45)	(0.43)	(0.44)	(0.41)	(0.39)
No Education Answer	-0.11	-0.50	-0.20	-0.07	0.04	-0.11
	(0.41)	(0.41)	(0.40)	(0.40)	(0.39)	(0.40)
Income: 30-50k	0.49^{**}	0.37^{*}	0.31^{*}	0.45^{**}	0.37^{**}	-0.00
	(0.16)	(0.15)	(0.15)	(0.15)	(0.14)	(0.15)
Income: 50-80k	0.81^{***}	0.45^{**}	0.41^{*}	0.47^{**}	0.43^{**}	0.14
	(0.19)	(0.17)	(0.17)	(0.17)	(0.16)	(0.16)
Income: 80k+	1.09^{***}	0.63^{**}	0.51^{**}	0.51^{*}	0.45^{*}	0.38^{*}
	(0.24)	(0.20)	(0.20)	(0.20)	(0.19)	(0.19)
No Income Answer	0.16	0.12	-0.09	-0.08	0.11	-0.28^{*}
	(0.13)	(0.12)	(0.12)	(0.12)	(0.12)	(0.12)
Log Likelihood	-1273.49	-1450.16	-1460.90	-1406.76	-1525.87	-1501.87
Deviance	2546.98	2900.32	2921.81	2813.53	3051.73	3003.74
Num. obs.	2597	2597	2597	2597	2597	2597

$^{***}p < 0.001$, $^{**}p < 0.01$, $^{*}p < 0.05$. Note: for gender, female serves as the baseline. For age, 18-29 serves as the baseline. For education, high school or less serves as the baseline. For income, less than 20k yuan serves as the baseline. We include covariates for "no income answer" and "no education answer" because a large number of respondents did not report these answers, and omitting these observations would drastically decrease the sample size.

4 Informed Consent

All participants in our Russia survey were asked to consent to participate in research. The study received relevant Institutional Review Board approval or exemption. Below, we provide the English version of the statement, which was translated into Russian for our Russian respondents, who were recruited by the survey firm Anketolog. Our respondents were compensated directly by Anketolog at market rates. All respondents had to consent in order to move forward with the survey. No identifying information was collected other than basic demographic data. Anketolog merely recruited subjects, who were directed to our Qualtrics platform. The firm has no access to the data. The Russian survey instrument contained no politically sensitive questions such as party identification or support for political leaders, only questions on hypothetical foreign policy scenarios.

4.1 Consent Form

Thank you for participating in this study! We would like to ask you some questions about foreign policy as well as some other basic things about yourself. Please answer the following questions to the best of your ability. There are no right or wrong answers. We're simply interested in what you think. Before we start, please read through the following.

CONSENT TO PARTICIPATE IN RESEARCH

Opinions on Foreign Policy

You are invited to participate in a research study conducted by xxxxx. You must be at least 18 years of age to participate. A total of N subjects who voluntarily agree on [survey platform] will participate. Your participation is voluntary. You should read the information below, and ask questions about anything you do not understand, before deciding whether or not to participate. Please take as much time as you need to read the consent form.

PURPOSE OF THE STUDY The purpose of the study is to investigate how individuals form foreign policy opinions.

PROCEDURES If you volunteer to participate in the study, we will ask you to do the following things: 1. Answer some questions about your preferences and personal attributes; 2. Answer some questions about your attitudes about foreign affairs. Answer follow-up questions about your experience. Participation will last approximately 10-15 minutes.

POTENTIAL RISKS AND DISCOMFORTS There are no physical risks during this survey.

POTENTIAL BENEFITS TO SUBJECTS AND/OR SOCIETY You will not benefit from the research study.

The researchers conducting the study may learn how personal attributes impact foreign policy behavior. This may help policymakers to understand how the mass public feels about foreign policy.

PAYMENT/COMPENSATION FOR PARTICIPATION You will be paid [amount] for your participation in this survey.

CONFIDENTIALITY Any information that is obtained in connection with this study and that can be identified with you will remain confidential and will be disclosed only with your permission or as required by law. The data used in the actual analysis and made public will have also personal identification removed and a unique identification number will be attached to each subject.

The members of the research team and the xxxxxxx's Human Subjects Protection Program (HSPP) may access the data. The HSPP reviews and monitors research studies to protect the rights and welfare of research subjects.

The data will be stored in the investigator's office on a password protected computer. The signed consent document will be maintained separately and will be destroyed three years after the study has been completed. The remaining data will be maintained indefinitely. The data may be used for other studies. If used, your consent will not be obtained because the data will not carry personal identifiers.

When the results of the research are published or discussed in conferences, no information will be included that would reveal your identity.

PARTICIPATION AND WITHDRAWAL You can choose whether to be in this study or not. If you volunteer to be in this study, you may withdraw at any time without consequences of any kind. You may also refuse to answer any questions you don't want to answer and still remain in the study.

ALTERNATIVES TO PARTICIPATION Your alternative is to not participate.

RIGHTS OF RESEARCH SUBJECTS You may withdraw your consent at any time and discontinue participation without penalty. You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you have any questions about your rights as a study subject or you would like to speak with someone independent of the research team to obtain answers to questions about the research, or in the event the research staff cannot be reached, please contact xxxxxx.

IDENTIFICATION OF RESEARCHERS If you have any questions or concerns about the research, please feel free to contact xxxxxxx.

AGE OF PARTICIPATION By signing this form below you agree that you are at least 18 years of age.

5 Dictionary for the Hitler Case

The Hitler case in the main text points out a marked shift in the leader's speeches as he rose to power, suggesting that Hitler was aware that domestic and international audiences were morally screening. The following terms comprise the dictionary we used when searching for those utterances, indicative of his evolutionary views: Selbserhaltung, biologisch, Auslese, Menschenwert, Lebenskraft, Volkswertes, Lebensbed[ü]rfnisse, Ernährung, Hochwert, Menschenmaterials, Blutsvermischung, Rassensenkung, Blutswertes, Lebenswiderständen, minderwertig. Further, we used the following terms as indicative of his amoralism: Macht des Stärkeren, Gesetz des Stärkeren, Recht des Stärkeren, Kampf des Stärkeren. We note that all major speeches based on his crude evolutionary thought, in which he dismissed

the very existence of humanitarian ethics, were made before the 1930 elections. After the 1930 elections, Hitler never made a public speech outlining his racial and crude Darwinian views, even after he took power. Only behind closed doors in his secret speech before leading industrialists on 26 January 1932 did he return to these themes. We see no mention of natural selection in any public comments until 1943, as the war was turning against the Nazis.

References

Baturo, Alexander, Niheer Dasandi and Slava J Mikhaylov. 2017. "Understanding state preferences with text as data: Introducing the UN General Debate corpus." *Research & Politics* 4(2):2053168017712821.

Johnston, Alastair Iain. 2017. "Is Chinese Nationalism Rising? Evidence from Beijing." *International Security* 41(3):7–43.

Katagiri, Azusa and Eric Min. 2015. Identifying threats: Using machine learning in international relations. In *Annual Meeting of the American Political Science Association*.

Kozlowski, Austin C, Matt Taddy and James A Evans. 2019. "The geometry of culture: Analyzing the meanings of class through word embeddings." *American Sociological Review* 84(5):905–949.

Lauretig, Adam M. 2019. Identification, interpretability, and Bayesian word embeddings. In *Proceedings of the Third Workshop on NLP and Computational Social Science*.

Pennington, Jeffrey, Richard Socher and Christopher D Manning. 2014. Glove: Global vectors for word representation. In *Proceedings of the 2014 conference on empirical methods in natural language processing (EMNLP)*. pp. 1532–1543.