Using PCA of latent distance between dyadic pairs to construct measure of strategic interest

After having first estimated the latent space and then subsequently calculating the latent distance between each dyadic pair for each of our three variables, dyadic alliances, UN voting and joint membership in intergovernmental organizations, we then combined these separate distances into one coherent measure.

To do so, we built off of the work of Chen et al. (2012). They developed a measure of relation strength similarity (RSS) which facilitates the discovery of relationships in complex networks. It allows for the combination of multiple-relationship networks (for the purposes of our paper, these are the latent distances between dyads as measured through alliances, UN voting scores and IGO membership), into a single weighted network (our measure of strategic interest). It does so using a principle components analysis (PCA) for each dyad. To that end, Chen et al. (2012) developed an R package *dils* to calculate the RSS.

We identified a number of issues with the original coding that we have adapted for our analysis. In particular, we adjusted the code : i) to scale and center the data as PCA analysis is sensitive to relative scaling of data ii) to sample with replacement as best practice with bootstrapping would seem to indicate that the sample size of each bootstrapped sample should be the same as the size of the original sample iii) such that the directions of the eigenvectors are consistent across the dyads. We then use the adapted version of this code to calculate the PCA for each dyad pair for a given year and then used the first principle component as our measure of strategic interest, which on average explains 42% of the variability across the three original measures.

Validating our measure of strategic interest

We further conduct a series of post-estimation validation tests for our resulting strategic variable. In particular, we (1) evaluate the relationship between our political strategic interest variable against S scores and Kendall's τ_b for alliances and (2) investigate how our measure of strategic interest describe well-known dyadic relationships.

First, we perform a simple bivariate OLS with and with year fixed effects to evaluate how our measures compare to S scores and Kendall's τ_b.[1] Note in order to make our strategic measures somewhat interpretable, for the validation we scale our strategic measures to be between 0 and 1 just as S scores and Kendall τ_b is scaled. The results are shown in Table A.1.

TABLE A.1: Validation of Political Strategic Interest Variable against S scores and Kendall's τ_b

	Unweighted S Scores	Unweighted S Scores	Weighted S Scores	Weighted S Scores	Tau-B	Tau-B
(Intercept)	0.97***	1.03***	1.01***	1.02***	0.29***	0.25***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Strategic Interest	−0.80***	−0.84***	−1.22***	−1.26***	−0.89***	−0.87***
	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
Year FE?	No	Yes	No	Yes	No	Yes

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

In brief, we find that our political strategic measure performs well against S scores and Kendall's τ_b for alliances with and without fixed effects. Note that because the PCA is of latent distances between any two dyads, dyads that are closer in space will have smaller values and therefore represent a stronger strategic relationship. Therefore the negative relationship we find between the political strategic measure and S scores and τ_b are interpreted to mean the greater the foreign policy similarity as measured by the S score or Kendal's τ_b, the smaller the latent distance or the greater the political strategic relationship between a dyad.

---

[1]Note for comparison that the bivariate relationship of S scores on Kendall's τ_b is statistically significant with a coefficient of 0.62 while the bivariate relationship of Kendall's τ_b on S Scores is statistically significant with a coefficient of 0.31.

Incorporating Lagged Dependent Variable

We evaluate the robustness of our key findings when including a lagged dependent variable in our model specifications:

$$
\begin{aligned}
Log(Aid)_{sr,t} = {}& \beta_1(Aid_{sr,t-1}) \\
& + \beta_2(Pol.\ Strat.\ Distance_{sr,t-1}) \\
& + \beta_3(Colony_{sr,t-1}) \ + \ \beta_4(Polity_{r,t-1}) \\
& + \beta_5 Log(GDP\ per\ capita_{r,t-1}) \ + \ \beta_6(Life\ Expect_{r,t-1}) \\
& + \beta_7(No.\ Disasters_{r,t-1}) \ + \ \beta_8(Civil\ War_{r,t-1}) \\
& + \beta_9(Pol.\ Strat.\ Interest_{sr,t-1} \times No.\ Disasters_{r,t-1}) \\
& + \ \delta_{s,r} \ + \ \rho_t
\end{aligned}
$$

When including the lagged dependent variable we see no change in the substantive implications of the model for humanitarian, civil society, or development aid. The results including the lagged dependent variable in the specification are shown in Figure A1 and those without are shown in Figure 4 of the paper.

Within the paper we also analyzed the persistence of foreign aid allocations over time. We do so here as well when including a lagged dependent variable in the specification. The results for each of the dependent variables are shown in Figures A2, A3, and A4 below. The findings from this analysis match with those shown in the paper where a lagged dependent variable is not included in the specification.
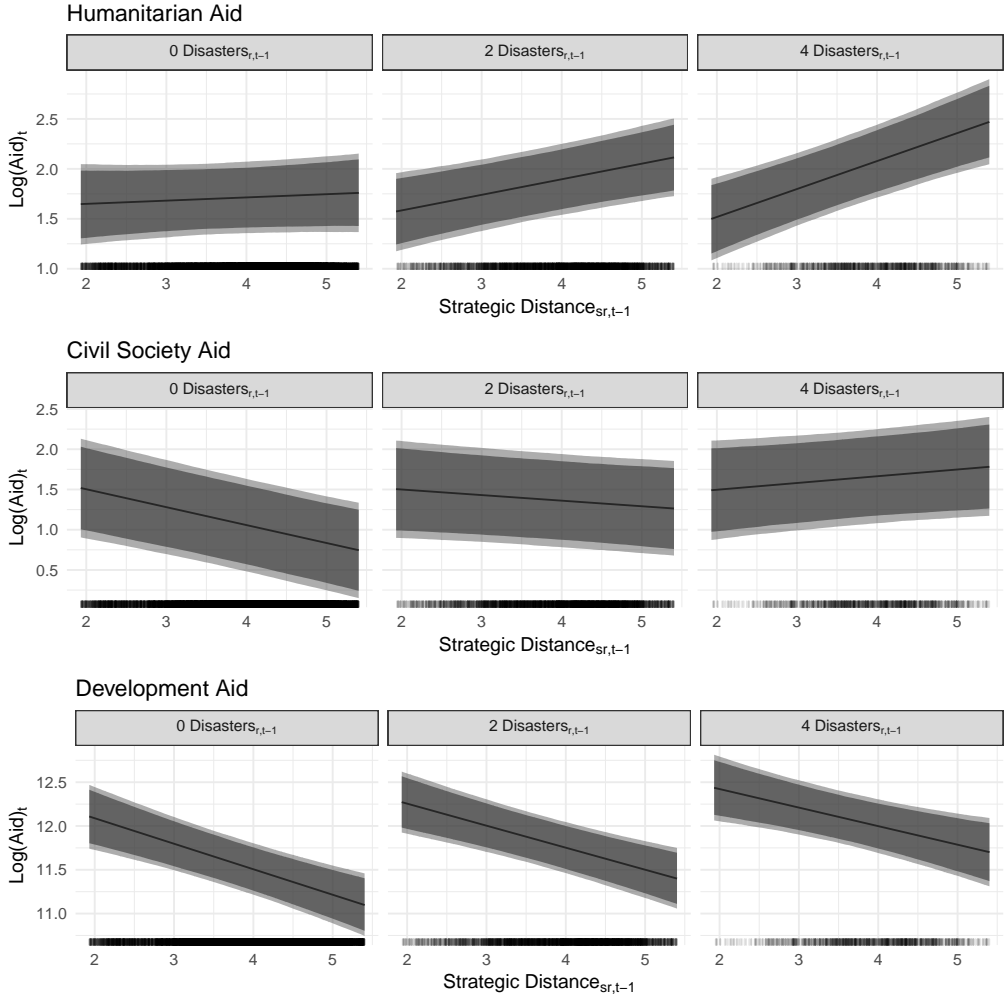
Humanitarian Aid

Civil Society Aid

Development Aid

Figure A1: Simulated substantive effect plots when including lagged dependent variable in the model for each dependent variable (humanitarian aid, civil society aid, and development aid) for different levels of natural disaster severity across the range of the strategic distance measure. A rug plot is provided below each panel.
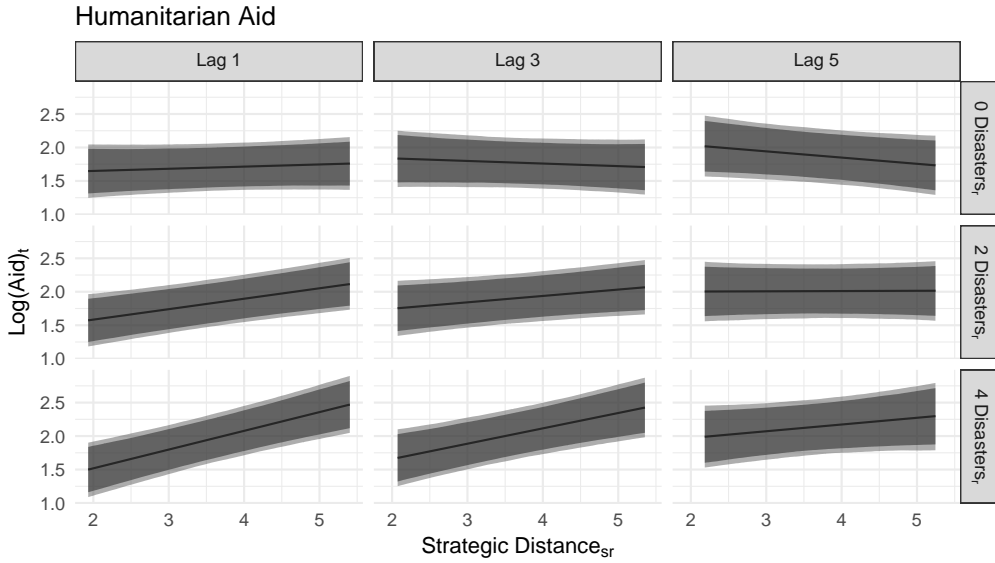
Humanitarian Aid

Figure A2: Simulated substantive effect plots when including lagged dependent variable in the model for humanitarian aid for varying lags of variables of interest and different levels of natural disaster severity across the range of the strategic distance measure.
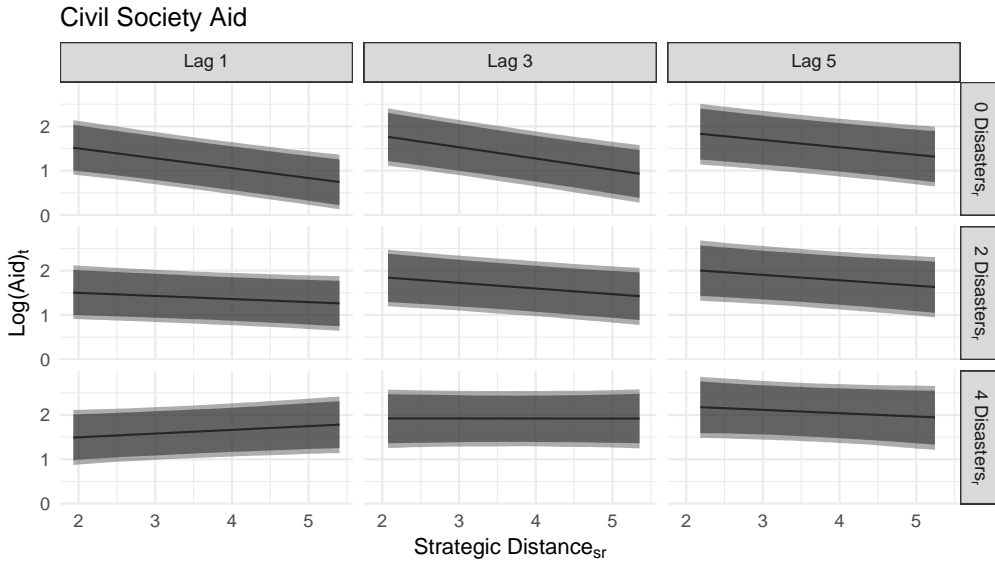
Civil Society Aid

Figure A3: Simulated substantive effect plots when including lagged dependent variable in the model for civil society aid for varying lags of variables of interest and different levels of natural disaster severity across the range of the strategic distance measure.
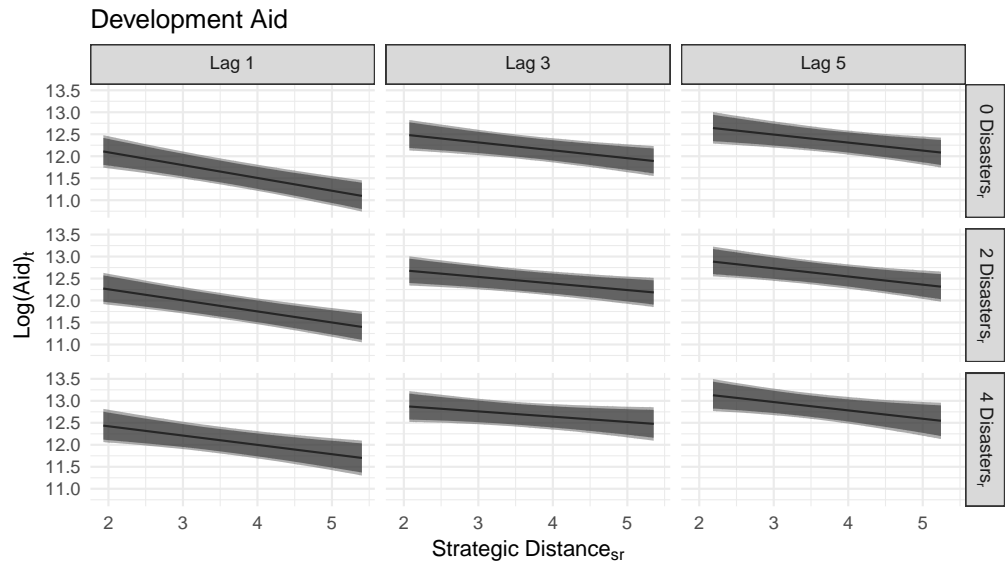
Figure A4: Simulated substantive effect plots when including lagged dependent variable in the model for development aid for varying lags of variables of interest and different levels of natural disaster severity across the range of the strategic distance measure.

Alternative Parameterization of Disaster Severity

We have also run our analysis using a dummy variable for whether a natural disaster occurred instead of a count. We show the substantive results of this analysis below in Figure A5. The findings from this analysis reflect those that we observe when we use the count variable. However, given the variation in relationships that we observe when using a count of the number of natural disasters, we choose to focus on that in the main portion of our paper.
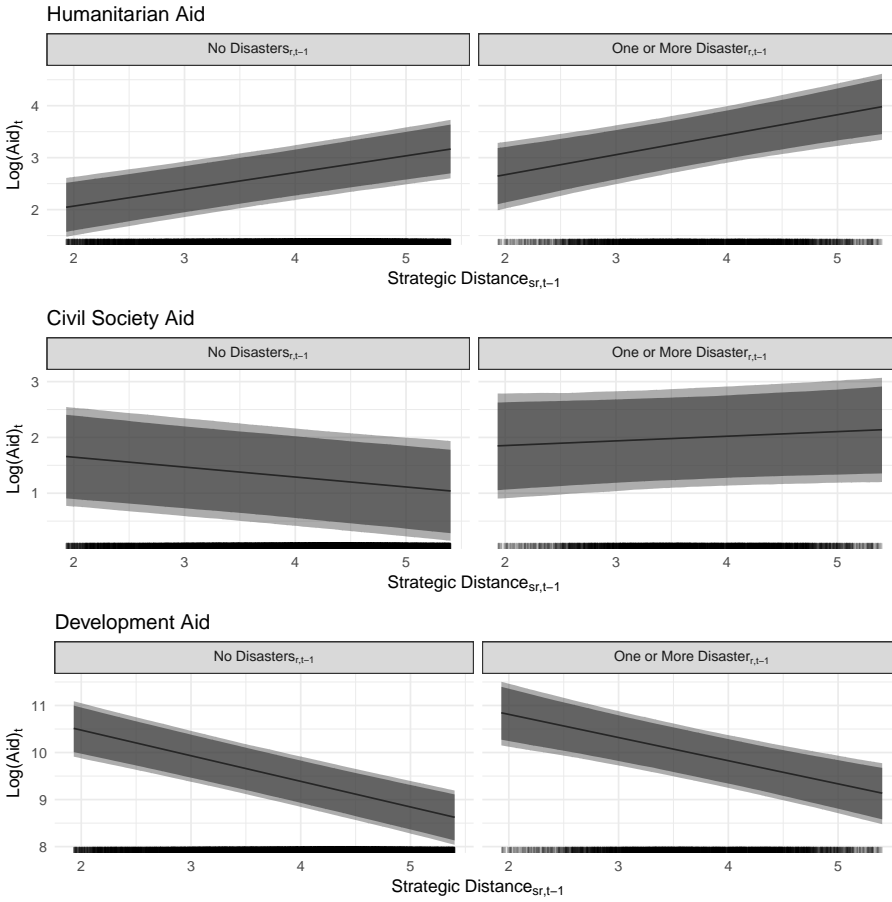

Figure A5: Simulated substantive effect plots for development aid for varying lags of variables of interest and whether or not a recipient country experienced a natural disaster across the range of the strategic distance measure.

We have also run our analysis using the number killed from a natural disaster instead of a count of the number of natural disasters. We show the substantive results of this analysis in Figure A6.
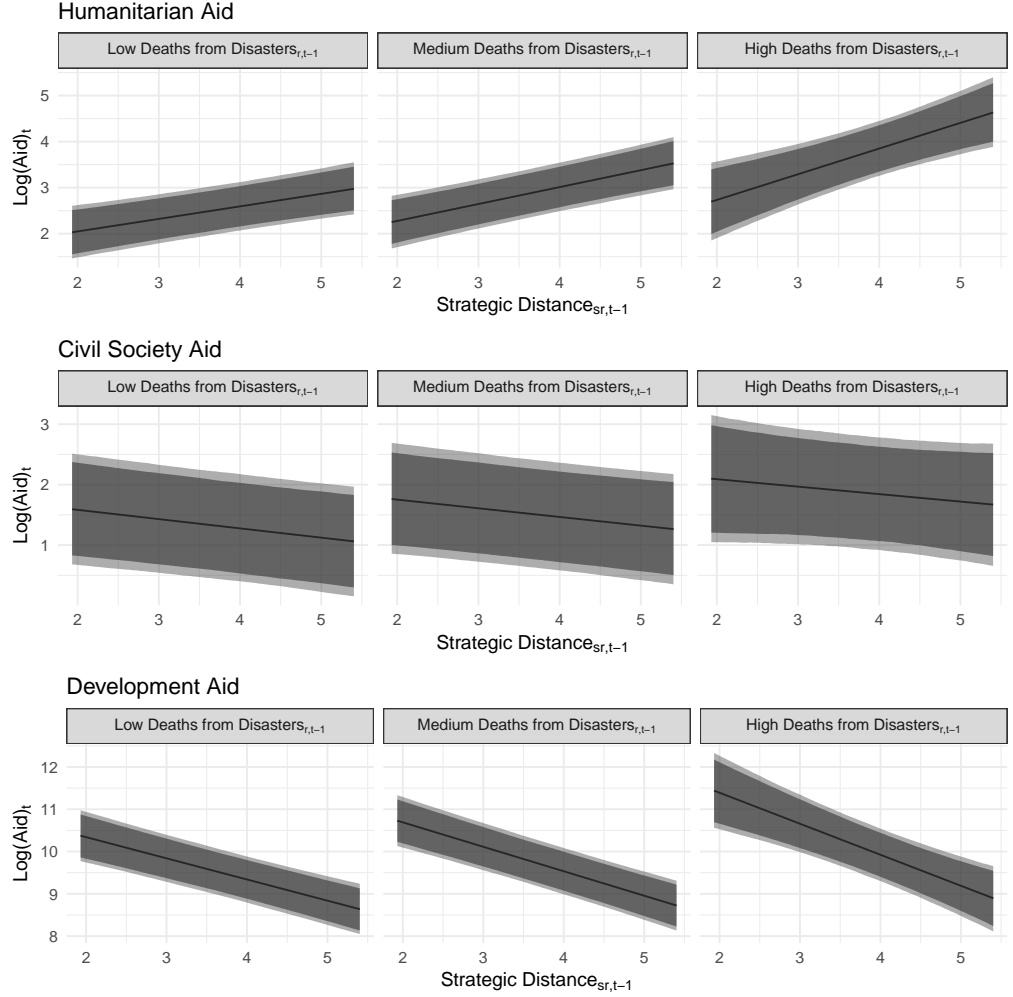

Figure A6: Simulated substantive effect plots for development aid for varying lags of variables of interest and different levels of natural disaster severity (specifically, the log of the number killed) across the range of the strategic distance measure.

The substantive trends with respect to humanitarian aid and development aid are notably similar to results that rely on a count of the number of natural disasters. There is a difference, however, with respect to the finding for the civil society aid dependent variable. In our analysis

with the count of the number of natural disasters we saw that at higher counts of natural disasters the slope between the amount of civil society aid given and strategic distance became positive. Here we see a less pronounced change in the slope between strategic distance when there are a higher number of deaths. This is perhaps explained by the fact that this measure has a missingness rate of 10.8%.

With regards to other potential measures, the EM-DAT database provides the data on number people injured, homeless, or affected and the dollar amount of the disaster. However such data has a high degree of missingness and, by their own admission, frequently imprecise or under-reported. For instance, there is 79% missingness for the number of injured, 36% missingness for the total number of homeless and 33% for the total damages. The number of affected has comparatively less missingness, with 9.6%. However the EM-DAT Guidelines note that, "The indicator affected is often reported and is widely used by different actors to convey the extent, impact, or severity of a disaster in non-spatial terms. The ambiguity in the definitions and the different criteria and methods of estimation produce vastly different numbers, which are rarely comparable." Generally all the indicators have varying degrees of imprecision. For instance, the guidelines further state, "Any related word like 'hospitalized' is considered as injured. If there is no precise number is given, such as 'hundreds of injured', 200 injured will be entered (although it is probably underestimated)." Given these problems with these other potential measures, we decided to focus on the number of disasters as our measure of disaster intensity.

Substantive Results when using Fixed Effects

In Figure A7 below we present the results of our analysis when using fixed effects. The results remain broadly the same as those shown in Figure 4 of the paper. Additionally, when running a Hausman specification test for our models we fail to reject the null hypothesis at both the 90 and 95% confidence intervals, providing at least some initial evidence that we are justified in our choice (Greene, 2008).
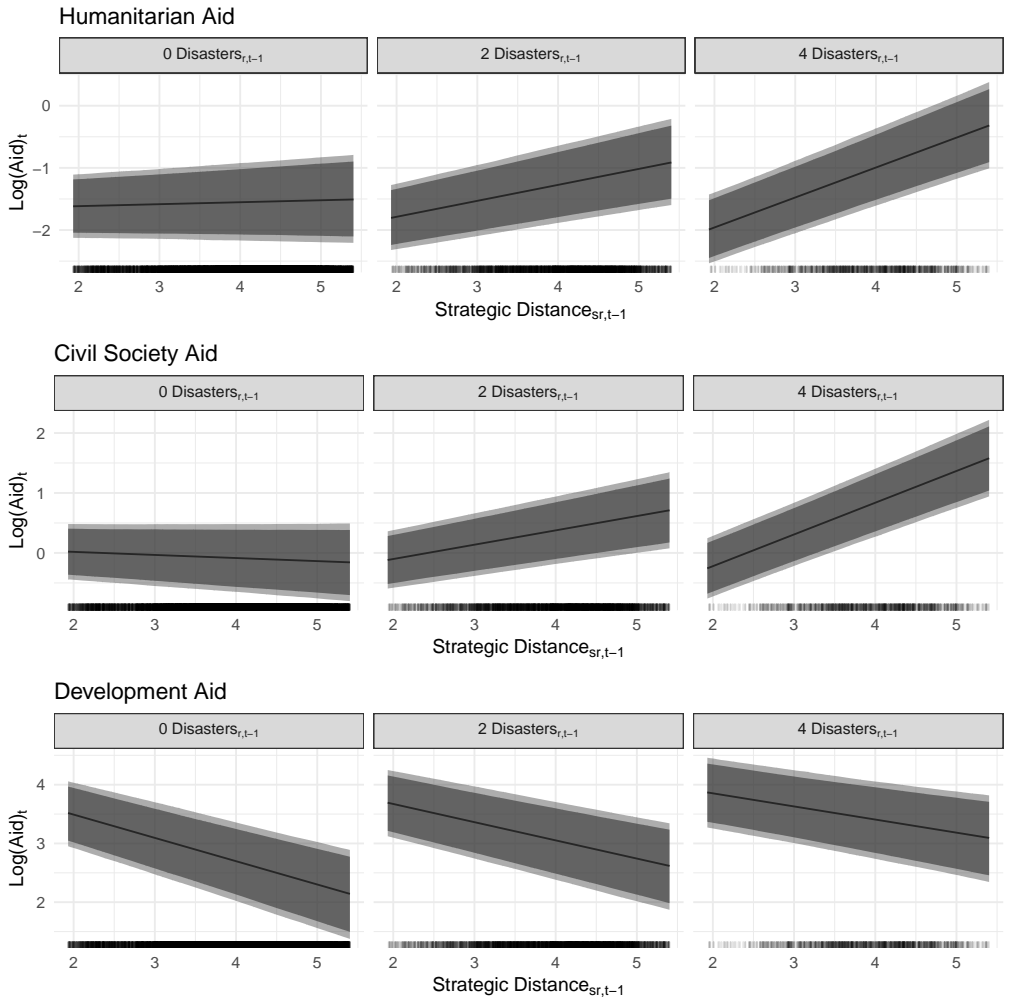

Figure A7: Simulated substantive effect plots when estimating using a fixed effects framework for each dependent variable (humanitarian aid, civil society aid, and development aid) for different levels of natural disaster severity across the range of the strategic distance measure. A rug plot is provided below each panel.

Temporal variation in patterns of aid

A limitation of our study is that it ends in 2005 because we face the constraint that the IGO data, an important component of our strategic interest measure, is simply not available past 2005. However, to show the potential relevance of our findings for more recent periods we have run our models using only data from the post Cold War period. The results are presented in Figure A8 below and mirror the findings presented in the paper.
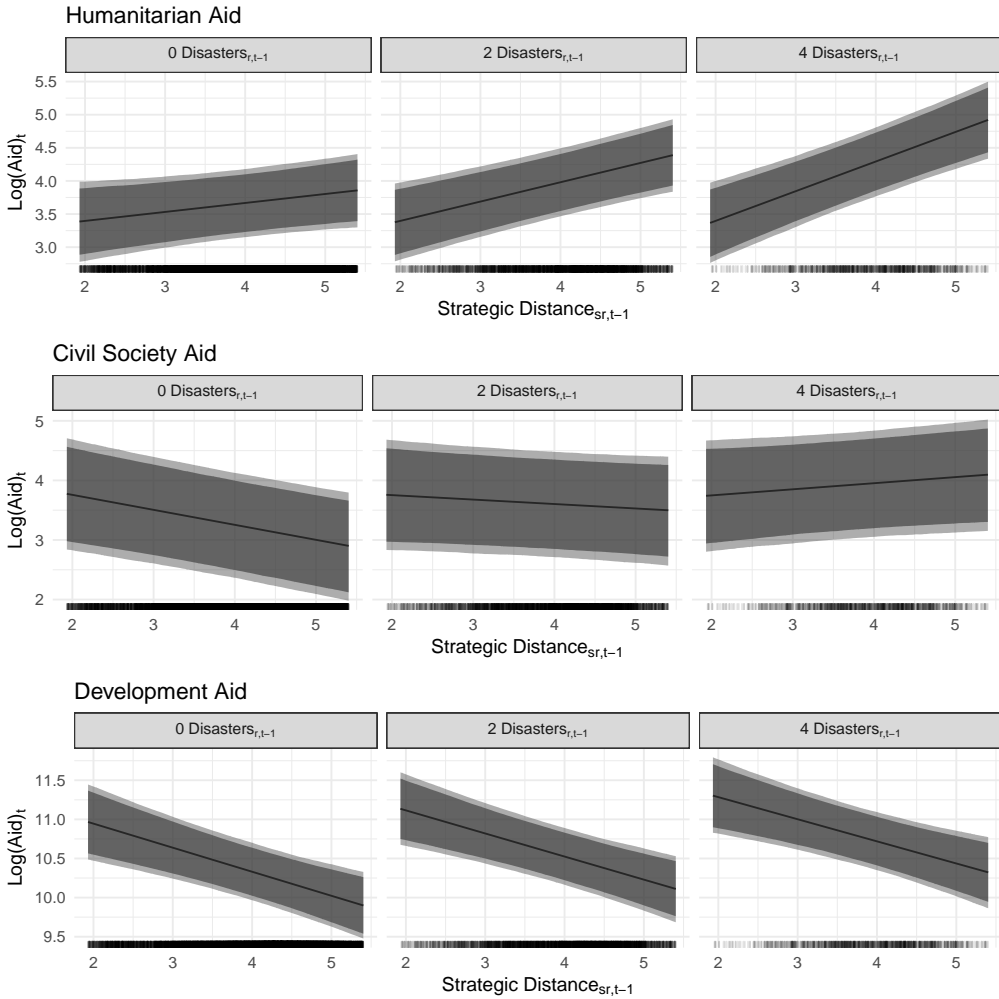

Figure A8: Simulated substantive effect plots for development aid for varying lags of variables of interest and different levels of natural disaster severity across the range of the strategic distance measure for the post Cold War period.

Additionally, we also run our models using only data from 2002-2005 (post-2001 period in our sample). The results are presented in Figure A9 below and mirror the findings presented in the paper.
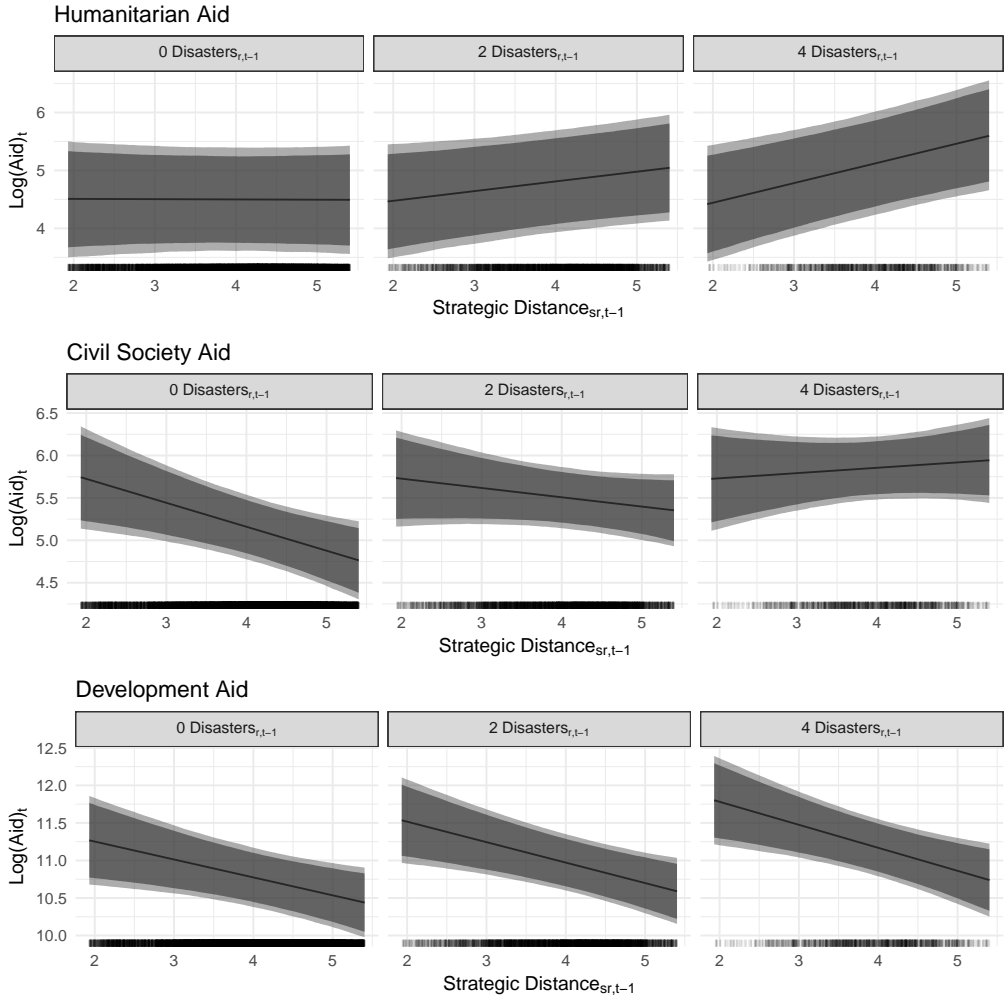

Figure A9: Simulated substantive effect plots for development aid for varying lags of variables of interest and different levels of natural disaster severity across the range of the strategic distance measure for 2001-2005.

Accounting for Uncertainty in Strategic Interest Measure

One methodological concern about our strategic interest measure is that since it is estimated from a model it comes with uncertainty. In Figure A10, we show results when taking into account uncertainty in the latent variable compared with our original estimates. We do this by simulating 1000 values of each latent variable estimate from the underlying distribution. From this we create 1000 versions of our dataset in which for each dataset we have a different sampled value of the strategic interest variable. We then run each of our models on those 1000 datasets and combine the parameter estimates using Rubin's rules (Rubin, 1987). We present the results of this analysis juxtaposed against our original model where we just use the average value of the strategic uncertainty variable.
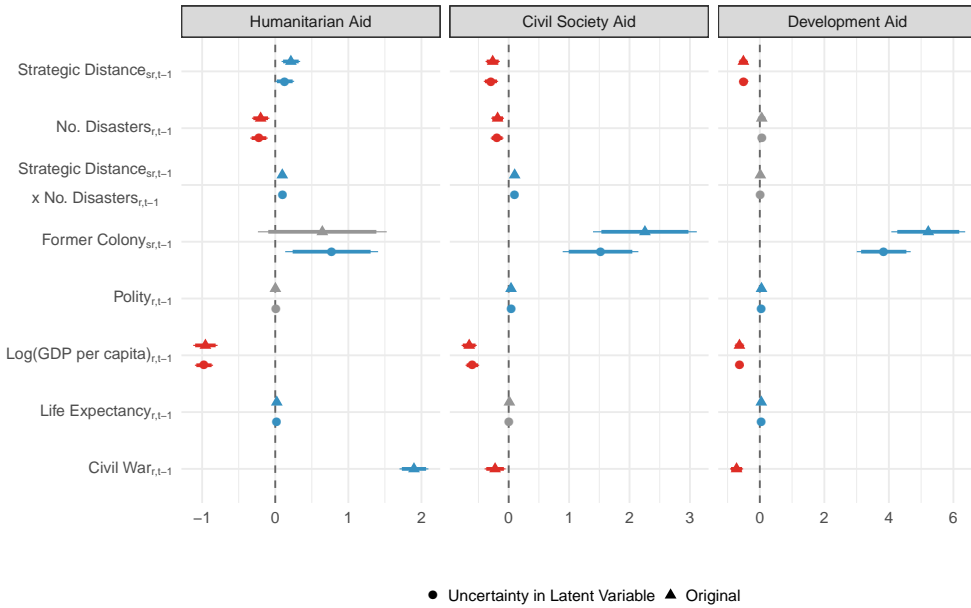

Figure A10: Effect of accounting for uncertainty in latent variable.

Investigating Whether Donor Labeling of Aid Channels is Driving Model Results

One point of concern is that donors may find it easier to channel aid to strategic allies through pre-existing development channels in the wake of natural disasters. Since donors cannot rely on such pre-existing channels to funnel aid to strategic opponents, our finding that donors give more humanitarian aid to strategic opponents in the wake of natural disasters may simply be an artifact. That is, it is possible that strategic allies and opponents both get more aid following natural disasters; strategic allies just get tend to get it through development aid and strategic opponents tend to get it through humanitarian aid.

We explore the possibility first through Figure A11. Here we break down how much development aid was given to countries experiencing that experienced either 0 disasters or 1-3 disasters across different types of development aid. The comparison between 0 or 1-3 disasters was used to maximized comparability, as around 40 percent of the country-years in the dataset had 0 disasters, while 43 percent experienced 1-3 disasters. This figure suggests that i) countries that are strategic allies (located at low levels of strategic distance) are more likely to get more aid related to economic infrastructure and services while ii) countries that experience disasters are much more likely to get debt relief when the they are strategic opponents (that is at high levels of strategic distance), compared to countries that experience 0 natural disasters. However, while Figure A11 does seem to be consistent the hunch that it may be easier to distribute different types of aid depending on whether one is a strategic opponent or strategic ally, there does not seem to be much in the way of strategic labeling going on. That is, the additional aid for economic infrastructure services to strategic allies and the additional aid for debt relief appear to be given in on top of existing levels of aid; neither appear to be offsetting other types of development aid.

We also investigate whether strategic allies are more likely to get a greater total amount of aid following a natural disaster. For instance, it is possible that strategic allies may receive more non-humanitarian or non-civil society aid, like budget aid following a natural disaster. To find evidence against this possibility and in support of the main argument in the paper, that natural disasters are indeed prompting more giving in strategic opponents compared to allies, we would expect a null effect of the interaction term between natural disasters and strategic distance on

total aid (that is, subtracting out humanitarian aid and civil society aid from total aid). We indeed find that there is a null effect on the interaction between natural disasters and our measure of strategic interest on total aid (as defined as all aid minus total humanitarian aid and minus total civil society aid; see Figure A12). When we plot the interaction effects, we verify this null finding (see Figure A13).
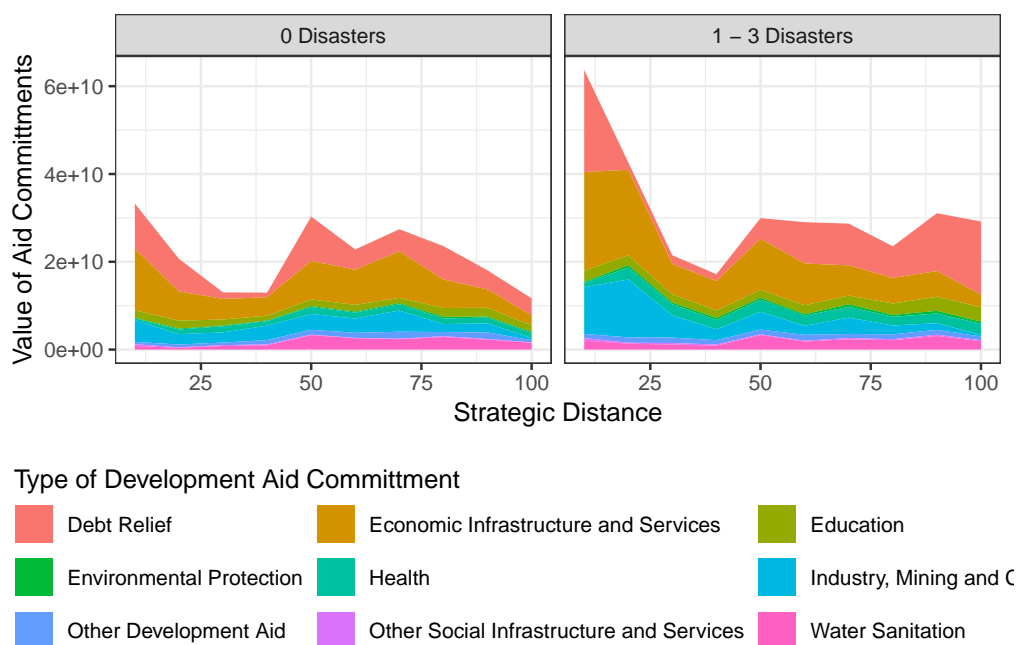

Figure A11: Value of Aid Commitments categorized by type of development aid and by the number of disasters
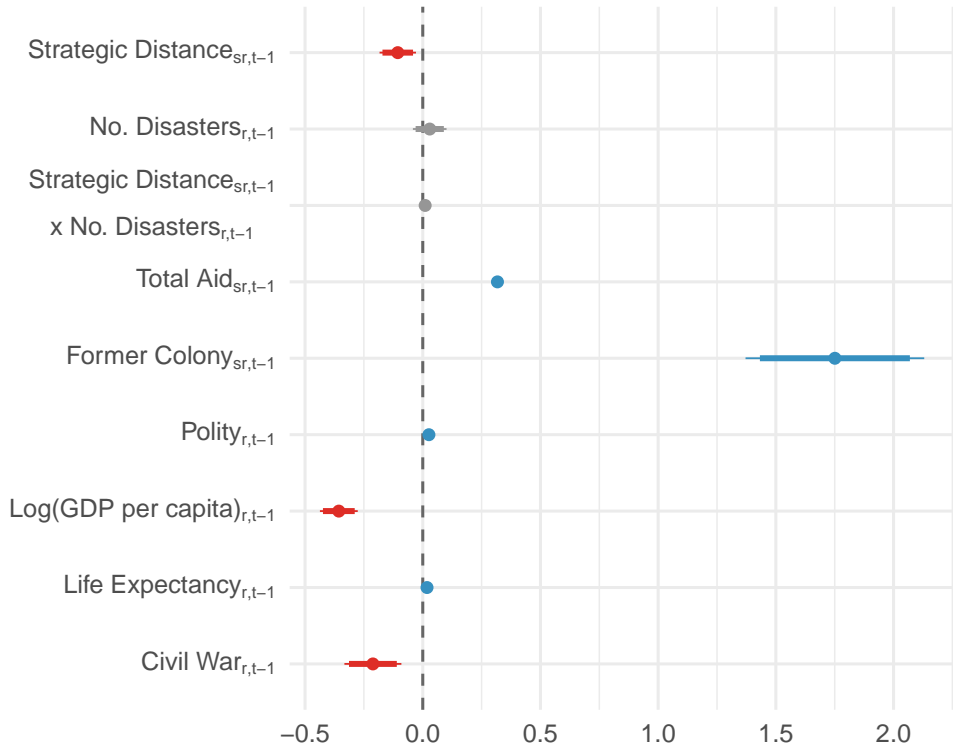
Figure A12: Coefficient plot for total aid model. Coefficients that are significant at the 5% level are shaded in blue if the coefficient is positive and red if the coefficient is negative. Coefficients that are not significant at the 5% level are shaded in gray.
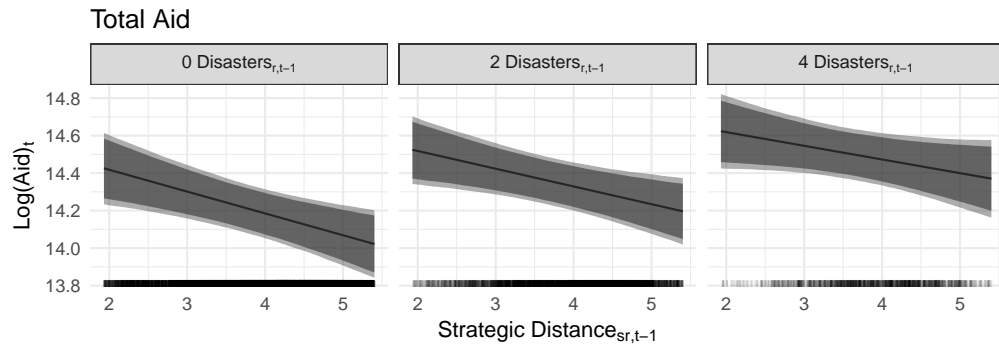

Figure A13: Simulated substantive effect plots for total aid for different levels of natural disaster severity across the range of the strategic distance measure. A rug plot is provided below each panel.

Rᴇꜰᴇʀᴇɴᴄᴇꜱ

Chen, H.-H., L. Gou, X. L. Zhang, and C. L. Giles (2012). Discovering missing links in networks using vertex similarity measures. In *Proceedings of the 27th annual ACM symposium on applied computing*, pp. 138–143. ACM.

Greene, W. H. (2008). The econometric approach to efficiency analysis. *The measurement of productive efficiency and productivity growth 1*(1), 92–250.

Rubin, D. B. (1987). The calculation of posterior distributions by data augmentation: Comment: A noniterative sampling/importance resampling alternative to the data augmentation algorithm for creating a few imputations when fractions of missing information are modest: The sir algorithm. *Journal of the American Statistical Association 82*(398), 543–546.