Appendix: Democracy and Alliance Treaty Depth

This appendix contains a series of robustness checks for the findings in the manuscript. In the first section, I examine results with an alternative measure of democracy when the alliance formed: the proportion of democracies in the alliance. Then I look at results from two altnerative measures of treaty depth. Last, I consider how uncertainty in the latent treaty depth measure affects my inferences.

1 Proportion of Democracies

In the paper, I use the average polity score of all alliance members when the treaty as the key independent variable. The proportion of democracies in the alliance is an alternative measure of how much heft democracies have in negotiations Chiba, Johnson and Leeds (2015). Rather than average the POLITY scores of alliance members, this variable is the share of alliance members with a POLITY score greater than 5, which is a common threshold for democracy. My argument anticipates that the proportion measure should lead to similar inferences about the connection between democracy and alliance treaty depth.

First, the descriptive statistics for the proportion of democracies across treaty depth and unconditional military support are as expected. Figure 1 shows the average proportion of democracies in four types of alliances. Deep, conditional alliances have the highest average proportion of democratic members, although it is not much higher than the shallow and conditional average.

Results from a generalized joint regression model of treaty depth and unconditional military support match the findings in the paper. In Figure 2, there is a no substantive association between
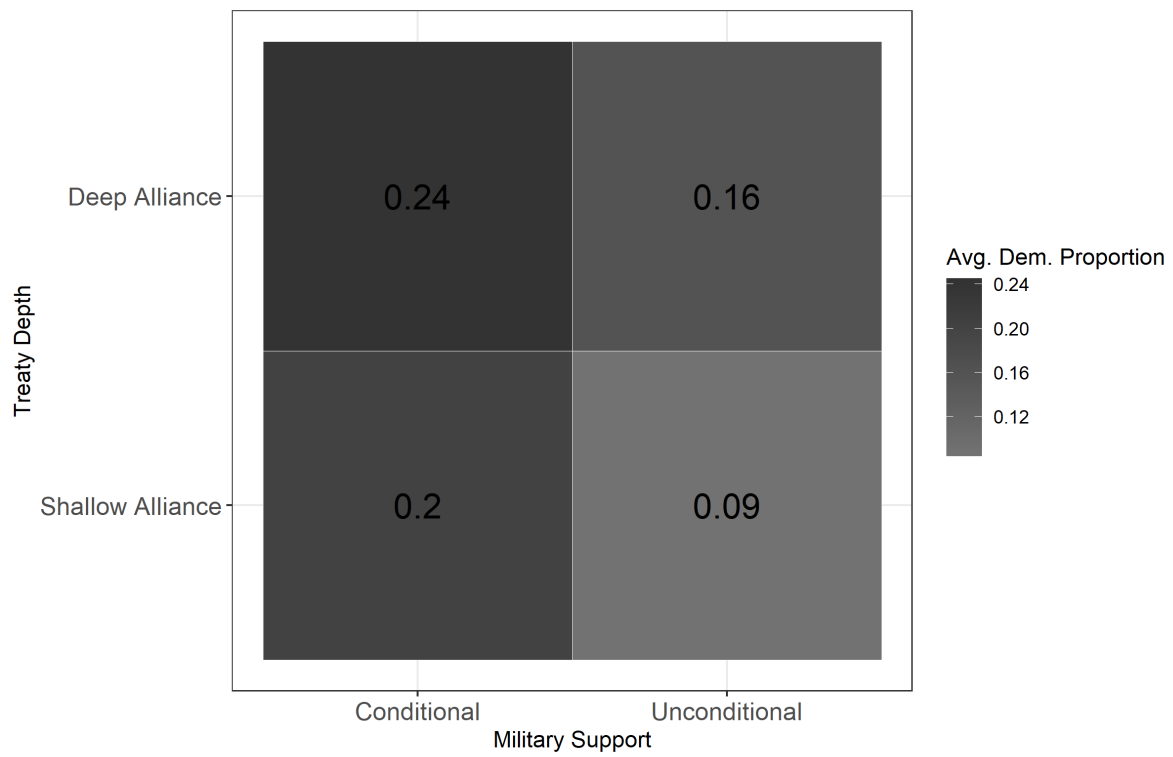
Figure 1: Average proportion of democratic members in four groups of alliances. Alliances are grouped based on treaty depth and unconditional military support.
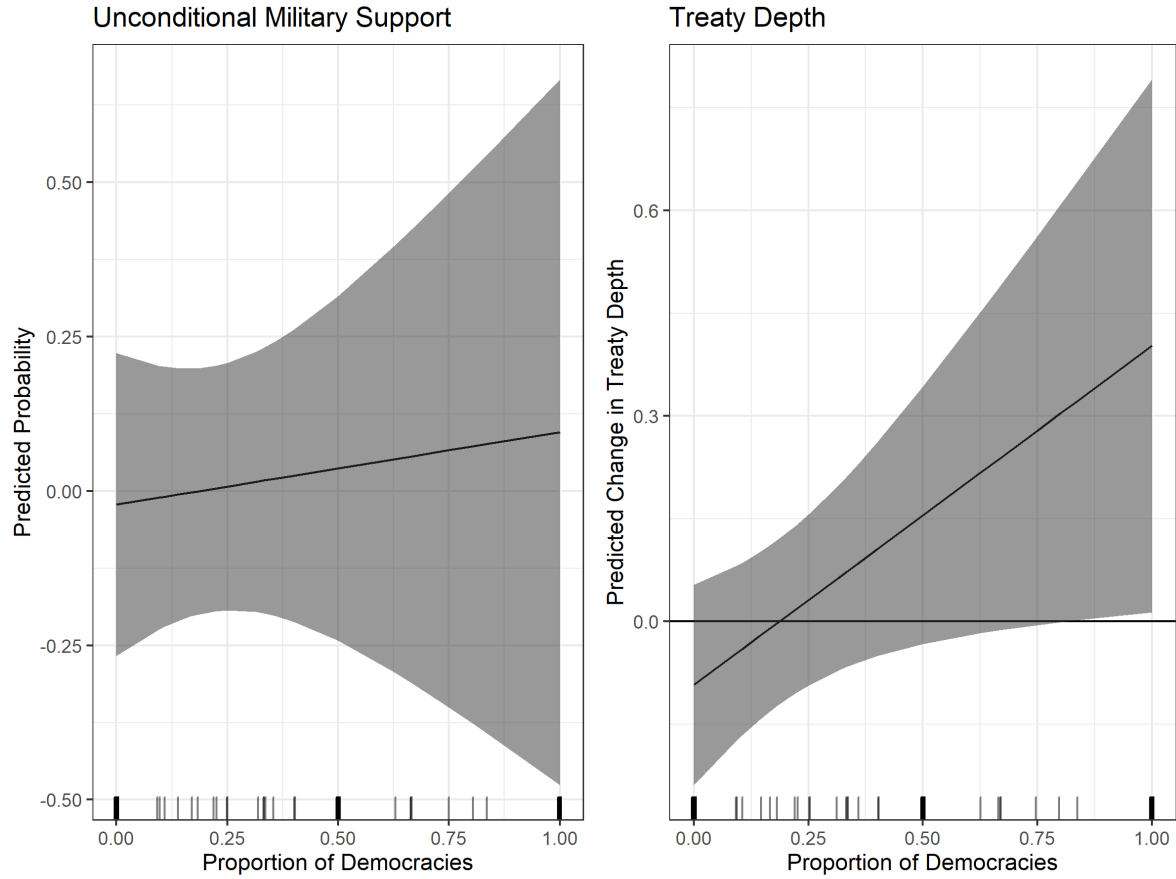
Figure 2: Predicted probabilities of unconditional military support and predicted changes in treaty depth by the the proportion of democratic alliance members. The line marks predicted values, and the shaded areas encapsulate the standard errors. The rug plot on the x-axis marks observed values of allied democracy. Predictions based on the smoothed terms from the joint generalized regression model.

the proportion of democracies in an alliance and the probability of unconditional military support.

Alliances between democracies are more likely to have high treaty depth, however. Therefore, I

find little evidence that democracies are more likely to form alliances with conditional obligations,

but I do find a positive relationship between democratic alliance membership and treaty depth.

2 Alternative Measures of Treaty Depth

There are two measures of treaty depth that operationalize similar concepts. Leeds and Anac (2005) develop an ordinal measure of military institutionalization. Benson and Clinton (2016) use a latent variable model to model treaty depth. While my measure is a better fit for my purposes in this paper, I find somewhat similar results with these two models. The null finding about unconditional military support holds. The Leeds and Anac measure of military institutionalization produces similar results in a GRJM model. The Benson and Clinton measure is harder to compare, because it covers fewer alliances and has more outliers that may impact inferences.

Figure 3 plots predictions from the average democracy smooth term for the two alternative measures. For the military institutionalization measure, I could not find an ordinal model in GJRM. Therefore, I created a dummy variable which is equal to one if the alliance had an institutionalization score of 1 or 2. I then modeled the military institutionalization dummy with a probit distribution.

3 Uncertainty in Treaty Depth

In a separate model, I consider how measurement uncertainty shapes inferences about the connection between non-major power alliances and treaty depth. The latent measure of treaty depth has some uncertainty. This is a reasonable approximation of alliance politics, because alliance treaty depth is not observed with certainty. There are perceptible differences in treaty depth, especially once states add substantial depth to the treaty. Even so, the results from the analysis of mean treaty depth may overstate the effect of non-major power membership.

To incorporate uncertainty over treaty depth, I fit a modification of the joint model. First, I created 1,000 datasets, one for each draw of the posterior distribution of the latent measure. Then I fit the model of mean treaty depth to 500 randomly sampled datasets from those 1,000 to facilitate computation. For models of depth with uncertainty, I use BRMS (Bürkner, 2017). BRMS is
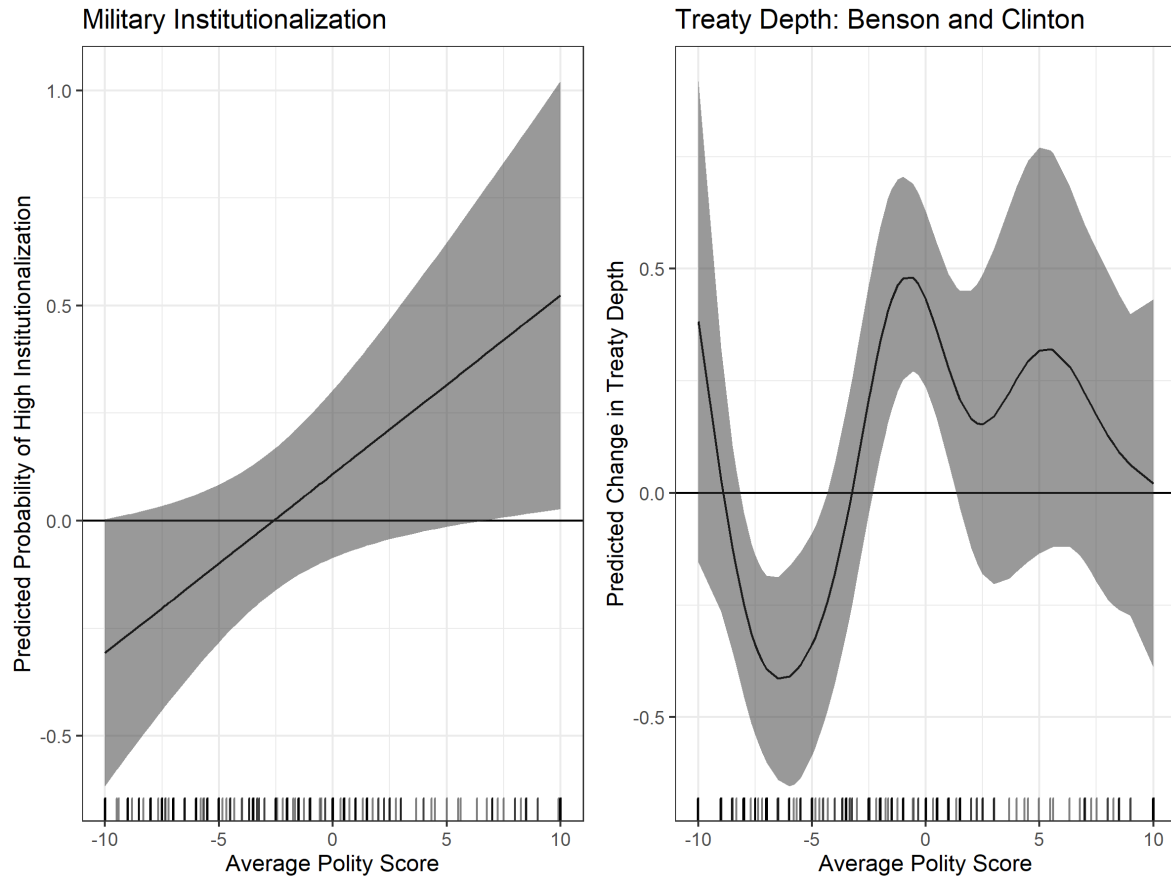
Figure 3: Predicted probabilities of unconditional military support and predicted changes in treaty depth by the the proportion of democratic alliance members. The line marks predicted values, and the shaded areas encapsulate the standard errors. The rug plot on the x-axis marks observed values of allied democracy. Predictions based on the smoothed terms from the joint generalized regression model.

an interface to STAN, a probabilistic programming language for Bayesian estimation (Carpenter et al., 2016). Joint Bayesian estimation has the flexibility to incorporate the probit and beta models and can be easily extended to account for uncertainty in the depth measure, but it does not allow correlated errors. Fitting the model sequentially to each dataset produces 500 separate models, which I combine into a single model by aggregating the posterior draws in a single posterior that accounts for uncertainty in the treaty depth measure.[1] This approach is analogous to common techniques for analyzing missing data, where multiple imputation generates uncertainty about the missing values (Hollenbach et al., 2018). After multiple imputation, researchers fit a separate model to each imputed dataset and then combine the results.

After accounting for uncertainty over treaty depth, I find a similar pattern. Allied democracy increases treaty depth, but decreases the probability of unconditional military support.

References

Benson, Brett V and Joshua D Clinton. 2016. "Assessing the Variation of Formal Military Alliances." *Journal of Conflict Resolution* 60(5):866–898.

Bürkner, Paul-Christian. 2017. "brms: An R package for Bayesian multilevel models using Stan." *Journal of Statistical Software* 80(1):1–28.

Carpenter, Bob, Andrew Gelman, Matt Hoffman, Daniel Lee, Ben Goodrich, Michael Betancourt, Michael A Brubaker, Jiqiang Guo, Peter Li and Allen Riddell. 2016. "Stan: A probabilistic programming language." *Journal of Statistical Software* 20:1–37.

Chiba, Daina, Jesse C Johnson and Brett Ashley Leeds. 2015. "Careful Commitments: Democratic States and Alliance Design." *The Journal of Politics* 77(4):968–982.

Hollenbach, Florian, Iavor Bojinov, Shahryar Minhas, Nils W. Metternich, Michael D. Ward and Alexander Volfovsky. 2018. "Principled Imputation Made Simple: Multiple Imputation Using Gaussian Copulas." *Sociological Methods & Research* pp. 1–29.

Leeds, Brett Ashley and Sezi Anac. 2005. "Alliance Institutionalization and Alliance Performance." *International Interactions* 31(3):183–202.

---

[1]Standard convergence diagnostics indicate convergence in all 500 models. Diagnostics like $\hat{r}$ are less useful for the full posterior, because some of the chains in the submodels do not overlap.