A. APPENDIX A: COMPARISON OF OUR RESULTS WITH THE VALIDATION METRICS IN THE SOURCE PAPERS

In this section, we compare our results to the results reported in the original papers that we evaluate. Table 1 shows the percentage of the variation in incumbents' DW-Nominate scores explained by each model that we report in the main text. It also shows the percentage of the variation in DW-Nominate scores that each of the source papers reports that their model explains.

Table 1: Validation Statistics for Various Measurement Models Against Contemporaneous Nominate Scores in the U.S. House - Comparison with Results in Source Papers

Name	Survey Respond's	Experts	Twitter	CF-Score
Success in Explaining Within-Party Variation in DW-Nominate Scores for Democrats				
% Variation Explained (our analysis)	0.45	0.59	0.48	0.19
% Variation Explained (source paper)	0.57	0.52	0.52	0.31
Success in Explaining Within-Party Variation in DW-Nominate Scores for Republicans				
% Variation Explained (our analysis)	0.23	0.25	0.08	0.28
% Variation Explained (source paper)	0.29	0.31	0.12	0.44

Overall, our results regarding the amount of within-party variation explained by each model are very similar to the results reported in the source papers.[1] The only notable differences between our results and those in the source papers are for CF-Scores (Bonica, 2014) and the Survey-based scores (Ramey, 2016). This difference for CF-Scores likely stems from the fact that we focus on the 107-113 Congresses, while Bonica (2014, 370-371) focuses on the 96-112 Congresses. There appears to have been a tighter relationship between CF-Scores and DW-Nominate scores in earlier Congresses. Also, we use dynamic rather than static CF-Scores. The difference for the survey-based scores likely stems from the fact that Ramey (2016) uses W-Nominate rather than DW-Nominate to validate his estimates, and

---

[1]Note that the results from Barberá (2015) that we report in Table 1 were calculated from his replication data based on members of the U.S. House in the 112th Congress. The correlations reported in the paper are somewhat higher, but they include both members of the U.S. House and Senate.

there is a slightly higher correlation between the survey scores and W-Nominate than there is with DW-Nominate.

REFERENCES

Barberá, Pablo. 2015. "Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data." Political Analysis 23(1): 76–91.

Bonica, Adam. 2014. "Mapping the Ideological Marketplace." American Journal of Political Science 58(2): 367–386.

Ramey, Adam. 2016. "Vox Populi, Vox Dei? Crowdsourced Ideal Point Estimation." Journal of Politics 78(1).