Online Appendix/Supplemental Information—The Quality of District Representation in U.S. House Committees

Appendix Table of Contents

Appendix A: Descriptive Statistics

Table A1: Number of Legislators With an OC Score by Committee and Congress, 104th-114th Congresses

	Congress											
	104	105	106	107	108	109	110	111	112	113	114	Total
Agriculture	42	0	0	0	0	0	0	32	42	42	0	158
Appropriations	53	58	59	63	64	0	0	57	48	53	48	503
Armed Services	50	44	51	50	52	0	59	59	58	60	60	543
Budget	40	20	41	40	42	6	0	0	30	0	0	219
Education and the Workforce	40	41	46	48	45	45	42	43	38	38	33	459
Energy and Commerce	45	42	47	52	50	51	52	55	53	51	53	551
Financial Services	44	51	54	63	57	59	65	64	48	57	55	617
Homeland Security	0	0	0	8	0	28	0	24	28	0	28	116
House Administration	0	0	0	0	0	6	8	4	0	0	0	18
Intelligence (Permanent)	0	0	0	0	17	16	16	21	0	0	0	70
International Relations	38	0	0	35	42	42	0	0	37	0	0	194
Judiciary	32	33	34	35	37	39	39	38	37	35	31	390
Natural Resources	44	36	41	46	23	17	43	43	41	39	37	410
Oversight and Government Reform	46	0	0	0	35	0	0	0	36	0	34	151
Rules	13	12	12	13	12	11	11	13	12	12	13	134
Science, Space, and Technology	48	0	18	0	24	0	0	0	0	0	32	122
Small Business	0	0	0	0	0	0	0	0	20	0	0	20
Transportation and Infrastructure	55	0	0	0	0	0	0	0	0	46	0	101
Ways and Means	34	20	35	36	38	37	38	38	36	35	39	386

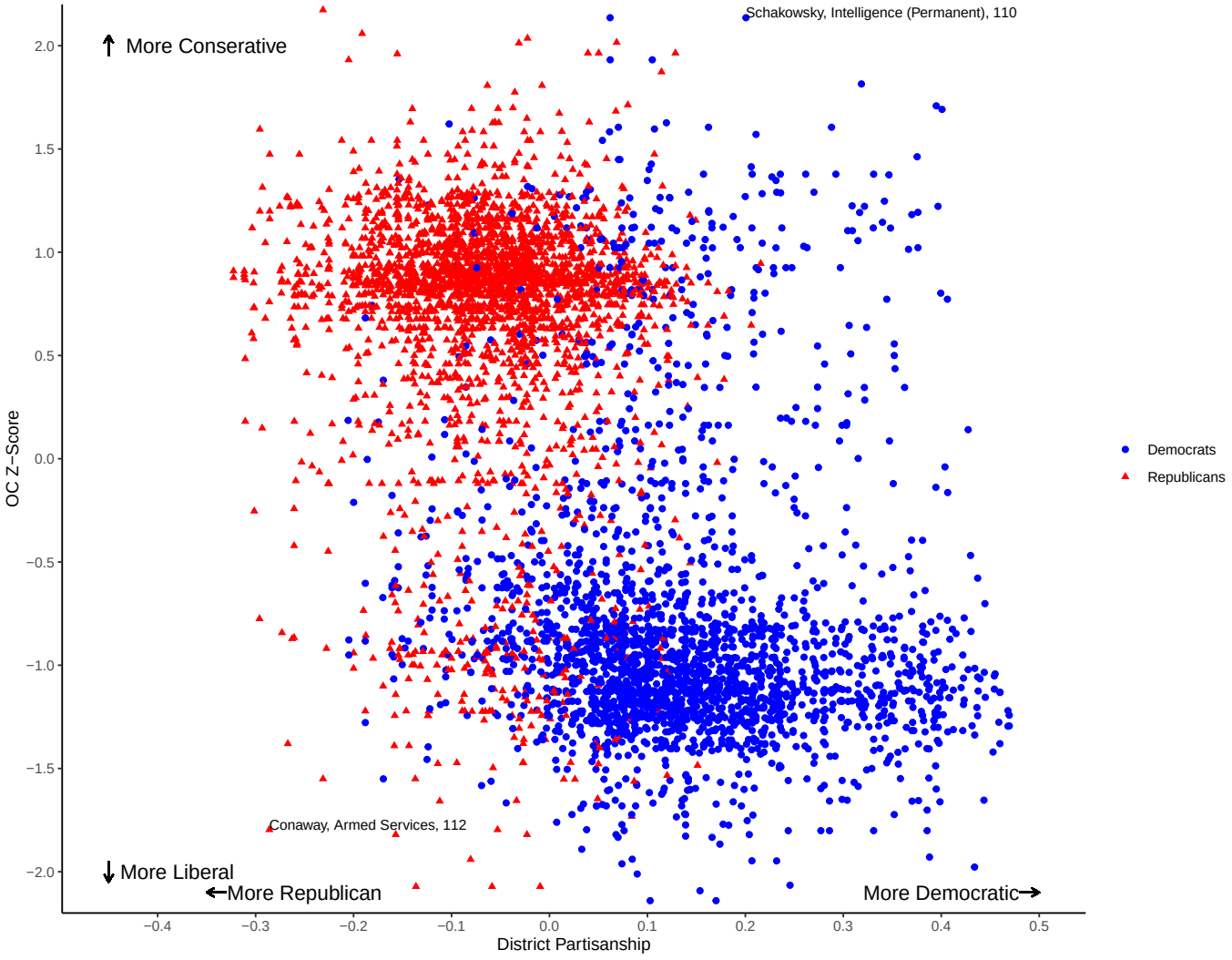
Members do not receive an OC score if they voted fewer than ten times on committee j in congress t. Some committees excluded if they have no observations across all congresses (e.g., Veterans' Affairs).

Table A2: Summary Statistics of OC Z-Scores by Congress and Committee

	Minimum	1st Quartile	Median	3rd Quartile	Maximum
Overall					
	-5.66	-1.01	0.28	0.88	4.05
By Committee					
Agriculture	-1.83	-0.92	0.27	0.89	2.06
Appropriations	-1.41	-1.07	0.58	0.91	1.40
Armed Services	-1.98	-0.97	0.25	0.91	2.25
Budget	-1.82	-1.03	-0.12	0.87	1.97
Education and the Workforce	-2.26	-1.07	0.46	0.88	1.61
Energy and Commerce	-2.07	-1.03	0.30	0.86	2.0
Financial Services	-2.07	-0.96	0.23	0.90	2.01
Homeland Security	-1.77	-0.85	-0.11	0.85	3.45
House Administration	-2.47	-0.60	0.36	0.50	0.91
Intelligence (Permanent)	-3.46	-0.79	-0.24	0.55	2.14
International Relations	-2.42	-0.97	-0.08	0.94	2.04
Judiciary	-1.55	-1.09	0.47	0.87	2.17
Natural Resources	-2.14	-1.00	0.42	0.87	1.67
Oversight and Government Reform	-5.66	-0.73	0.17	0.85	1.68
Rules	-2.09	-0.76	0.30	0.68	2.57
Science, Space, and Technology	-1.60	-0.92	0.31	0.79	4.05
Small Business	-1.24	-1.24	0.72	0.72	1.26
Transportation and Infrastructure	-1.94	-1.19	0.40	0.65	1.47
Ways and Means	-2.35	-0.94	0.51	0.89	1.70
By Congress					
104	-1.98	-1.03	0.45	0.85	1.96
105	-2.26	-1.03	0.37	0.87	2.02
106	-1.95	-1.06	0.29	0.89	1.67
107	-2.42	-1.04	0.34	0.90	2.25
108	-2.42	-1.02	0.29	0.88	4.05
109	-3.46	-1.10	0.50	0.85	1.78
110	-2.47	-0.92	-0.53	1.09	2.14
111	-1.96	-0.88	-0.56	1.16	2.06
112	-1.83	-1.13	0.60	0.84	1.93
113	-1.94	-1.12	0.46	0.80	2.57
114	-5.66	-1.09	0.18	0.84	3.45

OC z-scores shown for the entire sample, by committee, and by congress. Because these are z-scores, mean is approximately zero for all rows.

Figure A1: Scatter Plot of District Partisanship and OC Z-Score by Legislator Party, 104th-114th Congresses



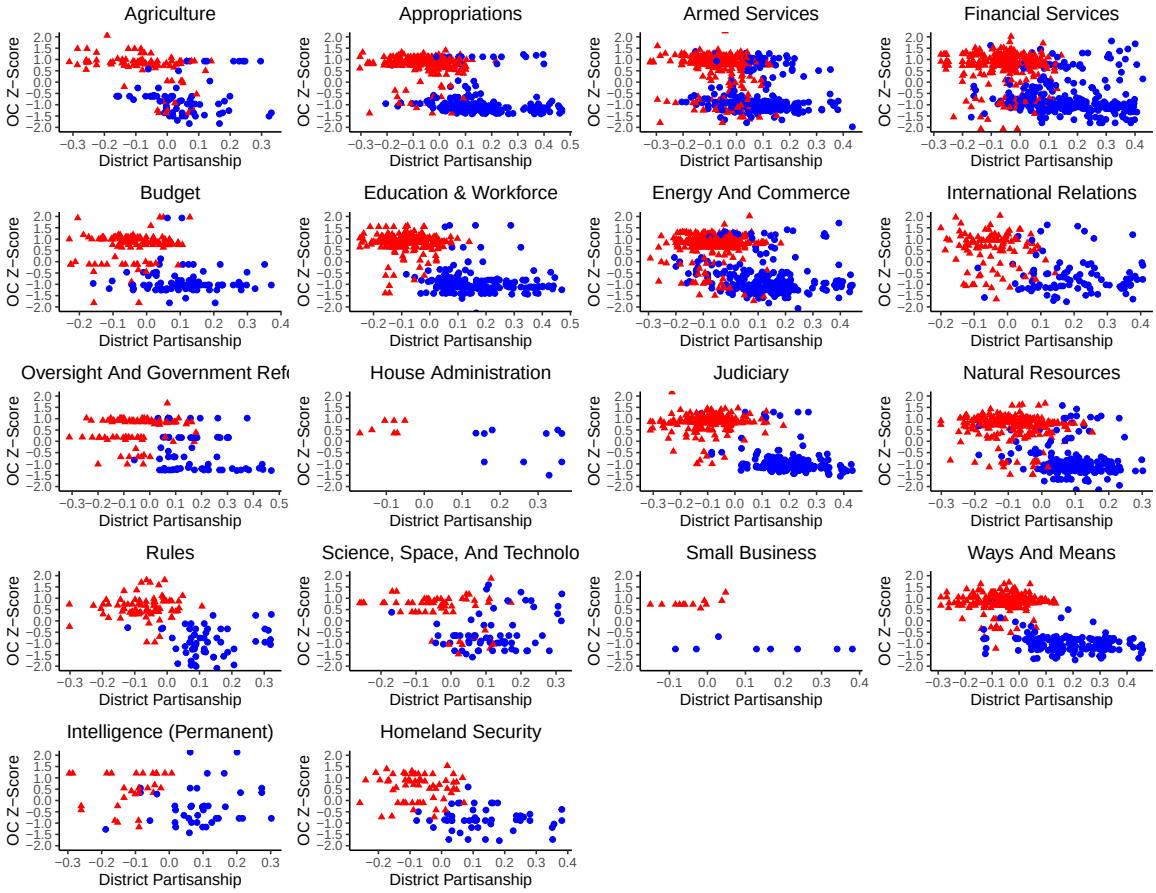
Points are legislator-committee-congress in 104th-114th Congresses. Some outlying district-assignments are omitted if their OC z-score is greater than two or less than -2.

Figure A1 shows the range for OC z-scores for each committee across all congresses (top panel) and for each Congress across all committees (bottom panel). The gray dot in each bar represents the median, and larger ranges indicate greater ideological dispersion within the committee/Congress. Because most House committee votes are partisan, a legislator casting even a small number of cross-party votes produces a more moderate OC z-score. Table A2 in Appendix A shows the minimum, quartiles, and maximums by committee and Congress.

Figure A1 shows a scatter plot of OC z-scores for all legislator-committee-congresses. The

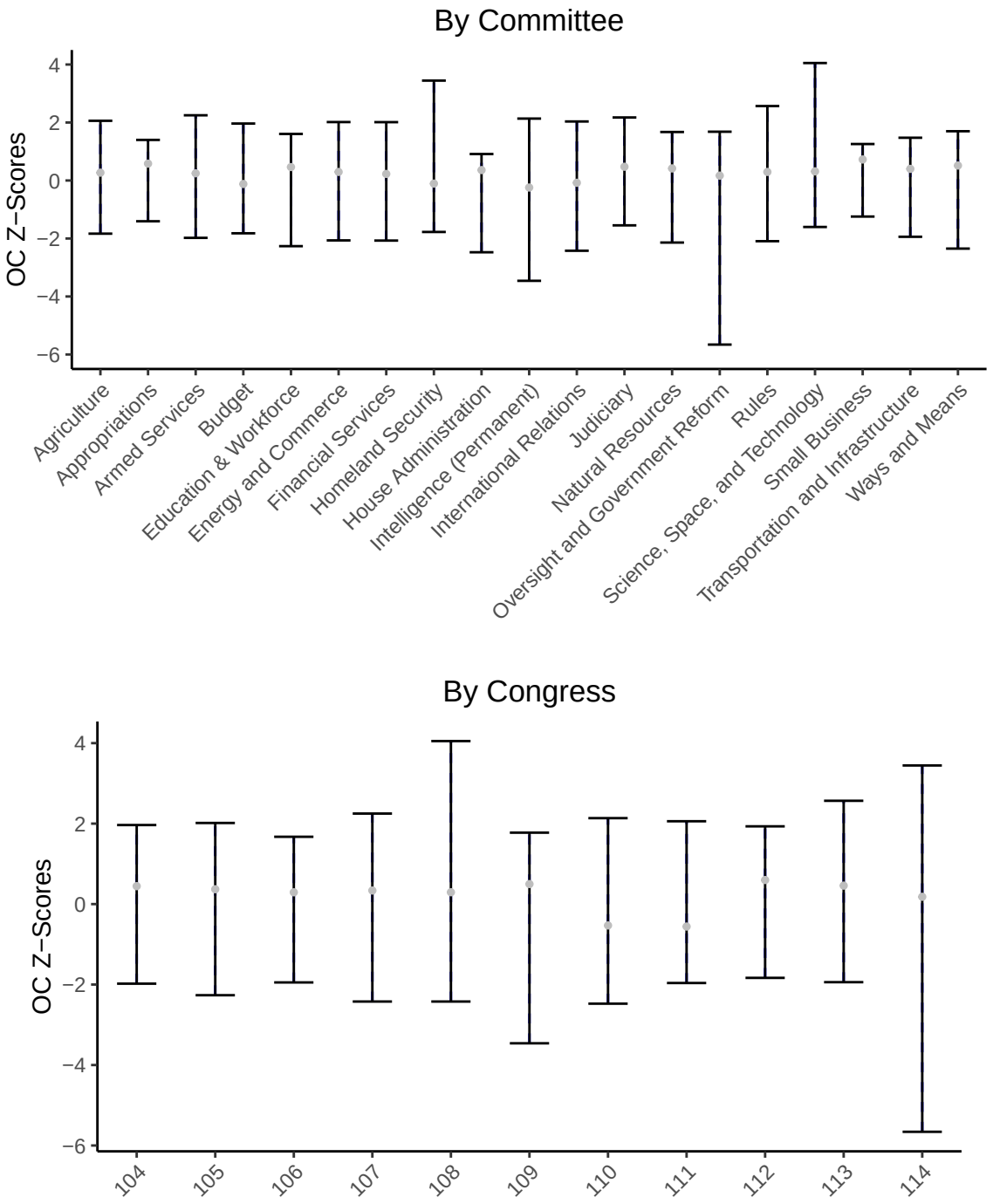
x-axis is district partisanship where lower values indicate a more Republican district, and the y-axis shows OC z-scores increasing from most liberal to most conservative. As the plot shows, Democrats and Republicans exhibit markedly different roll call voting behavior, as expected, with Republicans being far more conservative, on average. There is substantial variation within the parties, however, with some Democrats having conservative OC z-scores and some Republicans having liberal OC z-scores. For example, Michael Conaway's (TX-11) OC z-score on Armed Services in the 112th Congress is notable for being among the most liberal, despite having a very conservative district. Conversely, Jan Schakowsky compiled a very conservative voting record on Intelligence in the 110th Congress despite representing a relatively Democratic district. Figure A2 shows the same scatter plots separated by committee.

Figure A2: Scatter Plot of District Partisanship and OC Z-Score by Legislator Party and Committee, 104th-114th Congresses



Points are district-assignments by committee in 104th-114th Congresses. Some extreme district-assignments are omitted if their OC z-score is greater than two or less than -2.

Figure A3: OC Z-Score Summary Statistics by Committee and Congress



Bars show minimum and maximum OC z-scores by committee (top panel) and Congress (bottom panel). Gray dots represent median. Because these are z-scores, mean is approximately zero for all committees/Congresses.

Figure A4: Density of Democratic Presidential Support for Districts Used in Analysis

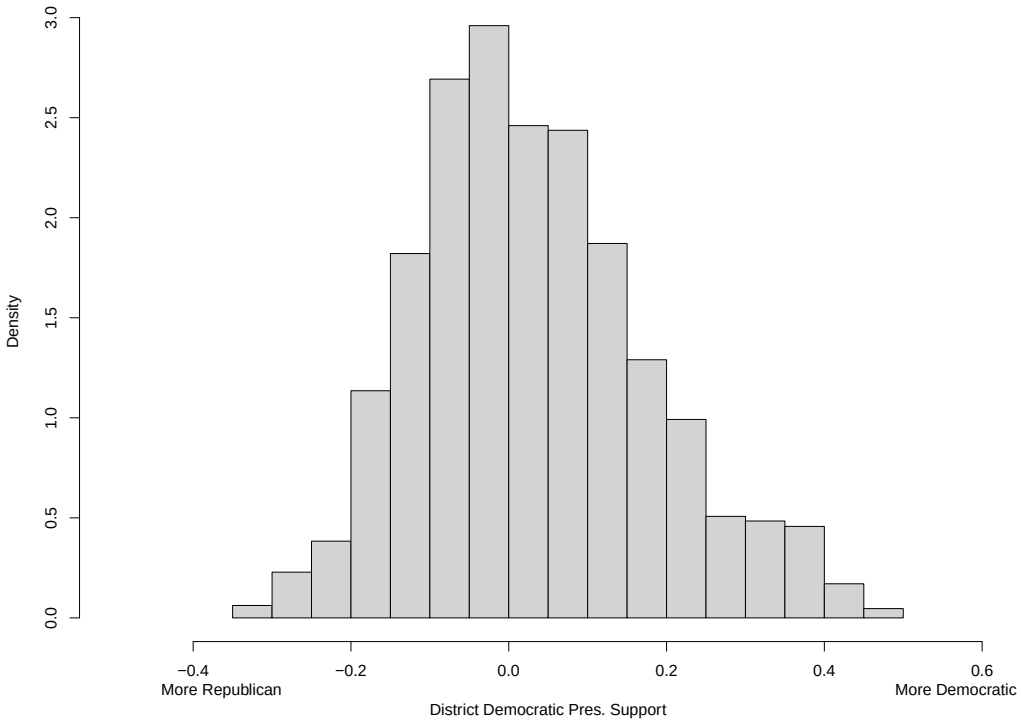


Figure A5: Density of Proportion of Committee Controlled by Majority Party Used in Analysis

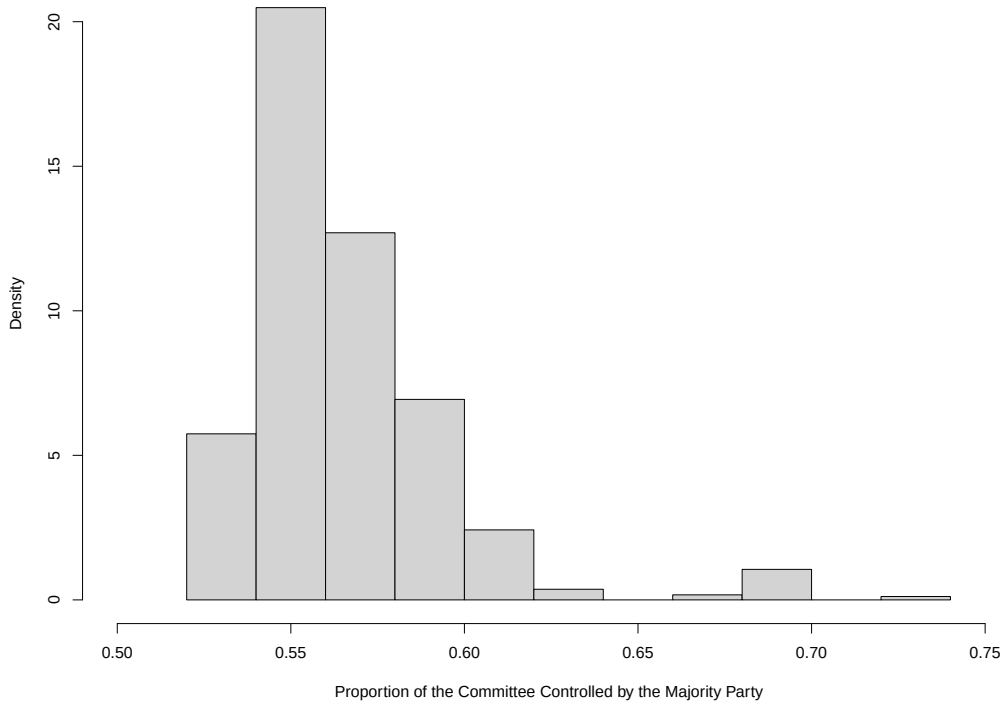


Figure A6: Density of Committee Staff Used in Analysis

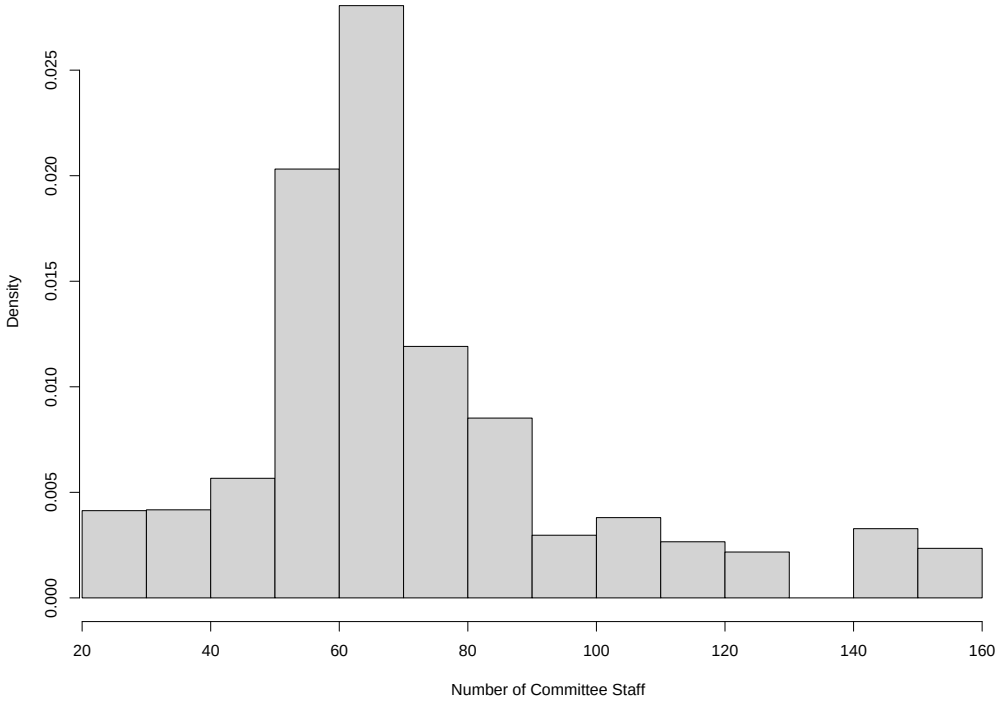


Figure A7: Density of Incumbent Vote Share Used in Analysis

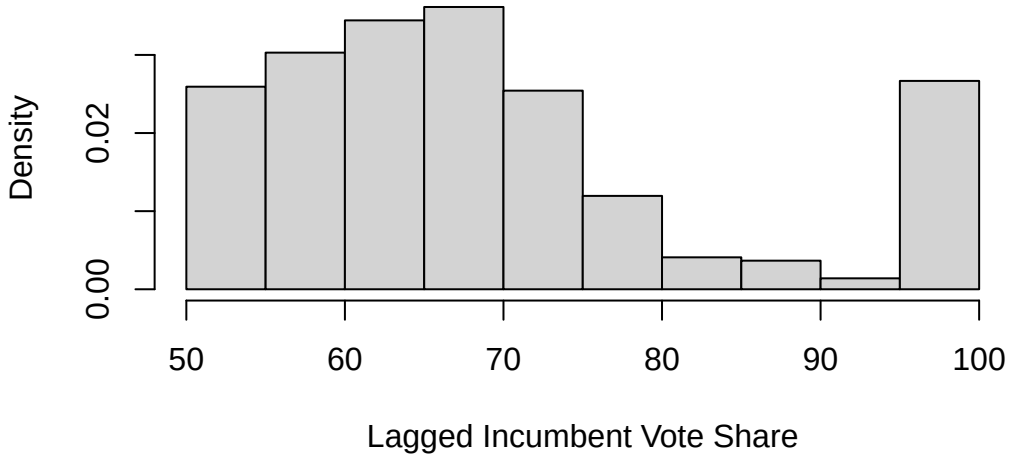


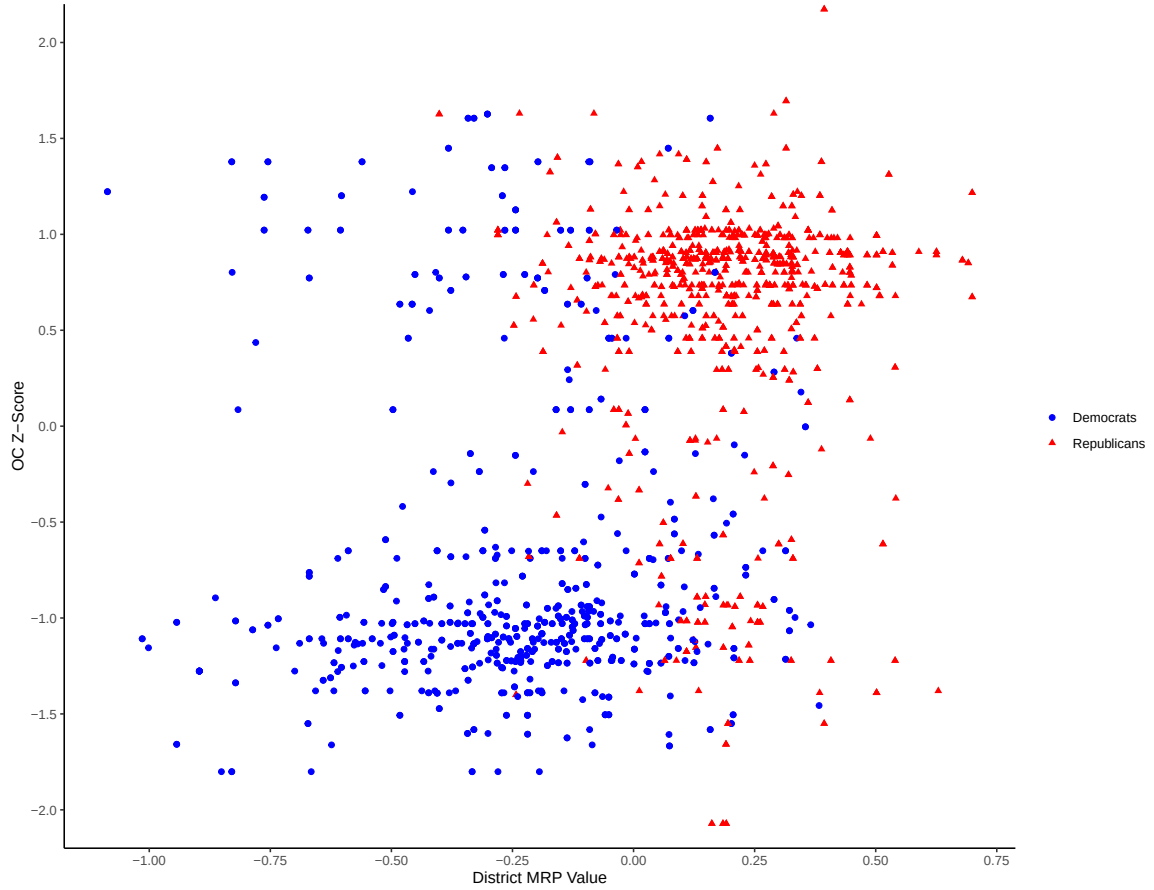
Figure A8 shows a scatter plot with DW-NOMINATE scores on the x-axis and OC z-scores on the y-axis. If the two scores were highly correlated, one would expect Democrats, with low DW-NOMINATE scores indicating greater liberalness to also have low OC z-scores, indicating greater liberal extremity relative to other committee members. Conversely, Republicans with high DW-NOMINATE scores should have high outlier OC scores on committees. Both claims are true in the figure. The correlation between OC z-scores and DW-NOMINATE scores is .72.

Figure A8: Scatter Plot of DW-NOMINATE and OC Z-Score by Legislator Party and Committee, 104th-114th Congresses

Points are district-assignments by committee in 104th-114th Congresses. Some extreme district-assignments are omitted if their OC z-score is greater than two or less than -2.

Figure A9 shows a scatter plot with Tausanovitch and Warshaw district MRP Values on the x-axis and OC z-scores on the y-axis. If the two scores were highly correlated, one would expect Democrats, with low MRP values, indicating they represent more liberal districts, to also have low OC z-scores. Conversely, Republicans with high MRP scores should have high outlier OC scores on committees. Both claims are true in the figure. The correlation between OC z-scores and MRP values is .45. MRP values are calculated after redistricting, so they are correlated with the first Congress after a redistricting in the dataset (108th and 113th).

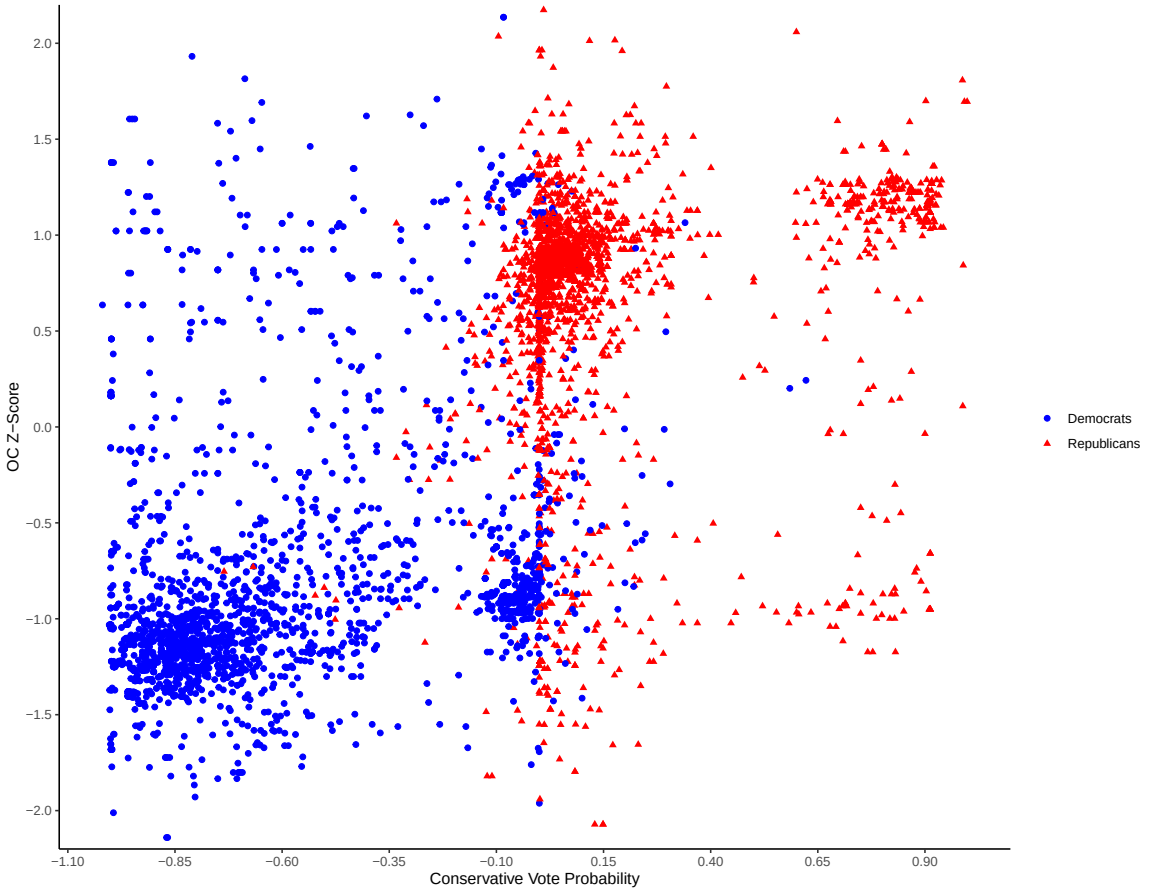
Figure A9: Tausanovitch and Warshaw MRP Values by Legislator Party, 108th and 113th Congresses



Points are district-assignments by committee in 108th and 113th Congresses. Some extreme district-assignments are omitted if their OC z-score is greater than two or less than -2.

Figure A10 shows a scatter plot with Folwer and Hall Conservative Vote Probabilities on the x-axis and OC z-scores on the y-axis. If the two scores were highly correlated, one would expect Democrats, with low conservative vote probability values, indicating they represent more liberal districts, to also have low OC z-scores. Conversely, Republicans with high conservative vote probability scores should have high outlier OC scores on committees. Both claims are true in the figure.

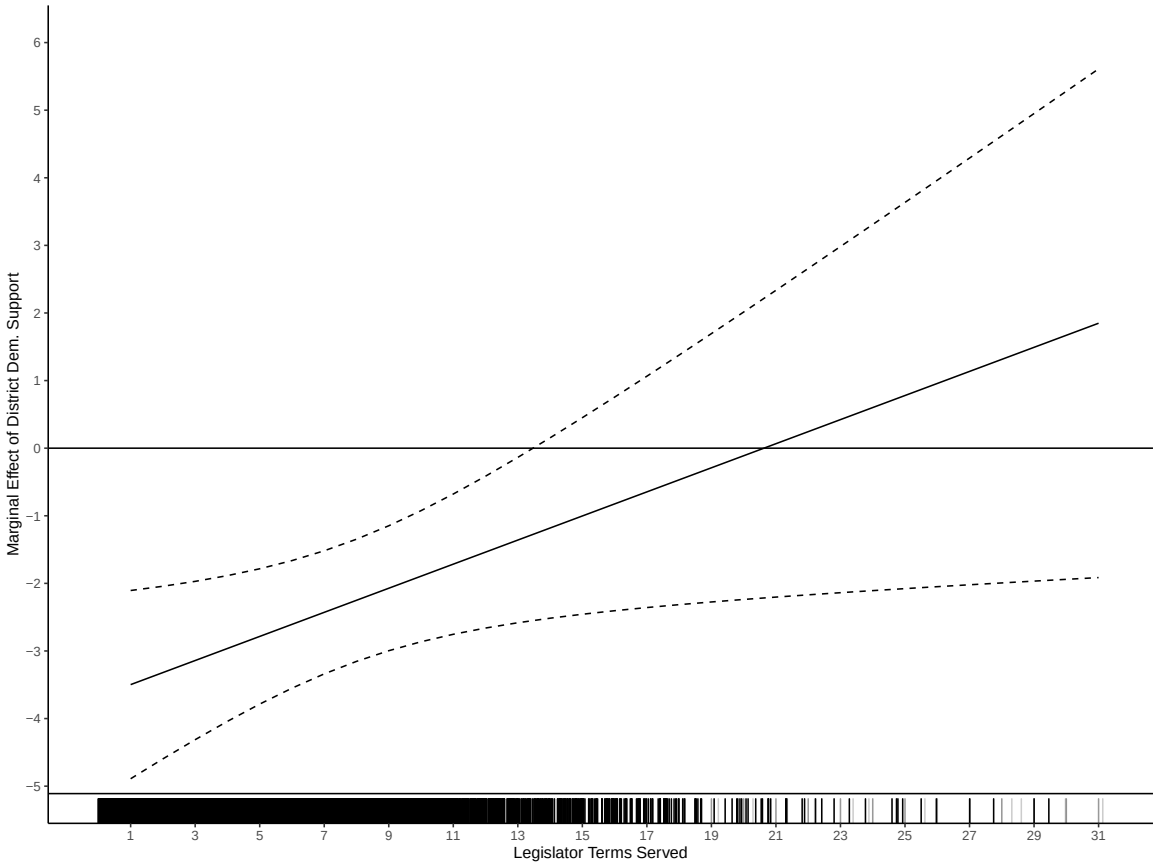
Figure A10: Fowler and Hall Conservative Vote Probabilities by Legislator Party, 104th-114th Congresses



Points are district-assignments by committee in 104th-114th Congresses. Some extreme district-assignments are omitted if their OC z-score is greater than two or less than -2.

Appendix B: Legislator Term Interaction Marginal Effects and Party Unity Results

Figure B1: Marginal Effect of District Partisanship Conditional on Legislator Terms Served



Marginal effects from model 3 in Table 3. The y-axis is the marginal effect of district Democratic support, and the x-axis is the number of terms served by a legislator. Rug plot at bottom shows distribution of number of legislator terms (values jittered).

The first three models in Table B1 leverage within district changes over time to predict party unity, and also include committee and congress fixed effects (model 1) and committee-by-congress fixed effects (model 2). Model 3 estimates first-difference changes within district. The independent variables are the absolute district partisanship as measured by the district's difference from the Democratic presidential vote share in the previous presidential election, where higher values indicate a more partisan district, and legislator party.

In each of the first three models, as a district becomes more partisan, the party unity score of the member from that district increases by about by .071% for each additional percentage more

Table B1: Two-Way Fixed Effects Estimates for Party Unity, 104th-114th Congresses

	District F. E.			Member F. E.		
	(1)	(2)	(3)	(4)	(5)	(6)
Absolute District Partisanship	7.11*	7.62*	8.03*	1.46	1.46	9.20*
	(2.20)	(2.32)	(3.48)	(2.43)	(2.63)	(4.26)
Legislator Party (GOP=1)	1.98*	2.05*	1.91*	89.40*	90.88*	
	(0.627)	(0.632)	(0.966)	(0.444)	(1.50)	
Unit Fixed Effects	Yes	Yes		Yes	Yes	
Committee Fixed Effects	Yes		Yes	Yes		
Congress Fixed Effects	Yes		Yes	Yes		
Committee-by-Congress Fixed Effects		Yes			Yes	
First Differences			Yes			Yes
No. of Unit Fixed Effects	983	983		1,001	1,001	
R-Squared	0.017	0.025	0.005	0.001	0.001	0.032
N	5,162	5,162	3,332	5,162	5,162	3,002

*p<.05. Models are panel linear regression where the dependent variable is a district's (models 1-3) or legislator's (models 4-6) party unity score on committee *j* in congress *t*. Cases indicate number of district or legislator fixed effects. Standard errors clustered by unit (number of clusters equals number of unit fixed effects, in model 3 number of clusters is 790 and in model 6 number of clusters is 794). In model 6 legislator party cannot be estimated because it does not change within legislator across time.

partisan the district is compared to the country. An increase from the minimum level of district partisanship to the mean produces an increase in party unity voting of about 3.33% (95% CI: 1.31% to 5.36%) (model 1). While not a huge substantive increase in party unity voting, it is equal to about .45 of a standard deviation. Party unity voting on committees is extremely high especially in recent congresses, and only 45 district-assignments out of 463 in the 114th Congress, had a party unity score 3.33 percentage points less than 100. The substantive effects are very similar in models 2 and 3, and combined, these result demonstrate that more partisan districts produce greater party unity within committee voting.

Legislator party is also positive and statistically significant. Consistent with research on asymmetric polarization (Thomsen 2009), even when accounting for district partisanship, Republicans have a higher party unity score than Democrats. Republican party unity scores are about 1.98% higher than Democrats, or slightly less than half of the effect size as absolute

district partisanship.

The results in models 4 and 5 of Table B1 show two-way fixed effects estimates leveraging variation within members rather than within districts. There are no significant results for either district partisanship or party, indicating that party unity voting in committees is responsive to district characteristics, but due to replacement of members rather than changes in behavior. The estimate for the differenced value is significant at the .05 level, and in the expected direction. This is the only significant estimate from the legislator-assignment identified models, however. These results are consistent with those for OC z-scores.

Appendix C: Coefficients for Committee Heterogeneity Interaction Models

Table C1 shows the coefficients for the substantive results plotted in Figure 1. The marginal effects in the left panel use the results from model 1, and are calculated by varying percentage of the committee controlled by the majority party, holding committee staff (logged) at its maximum. The marginal effects of district Democratic support in the right panel also use model 1 and are calculated by varying committee staff, holding percentage of the committee controlled by the majority party at its maximum. Figure 2 plots the substantive effects of absolute district partisanship from model 2 by varying percentage of the committee controlled by the majority party, holding committee staff (logged) at its maximum.

Table C2 shows the coefficients for the substantive results plotted in Figure 3. The left panel shows the marginal effect of district Democratic presidential support varying incumbent vote share in the previous election while holding all other variables at their mean. The right panel does the same, but shows absolute district Democratic presidential vote.

Table C1: District-Assignment Two-Way Fixed Effects Estimates of the Conditional Effect of Partisan- and Information-Induced Preferences on Ideology, 104th-114th Congresses

	DV=Ideology (1)	**DV=Absolute Ideology** (2)
District Dem. Presidential Support	34.07	
	(18.18)	
Absolute District Presidential Support		1.09
		(10.76)
Legislator Party (GOP=1)		-0.065*
		(0.024)
Percent of Committee Controlled by Majority	3.45	0.576
	(4.86)	(3.45)
Committee Staff (Logged)	0.395	0.245
	(0.741)	(0.471)
District Dem. Support x Percent of Comm. x Staff	17.16*	-0.946
	(7.94)	(4.58)
District Dem. Support x Percent of Comm.	-69.98*	0.453
	(32.10)	(18.73)
District Dem. Support x Staff	-9.21*	0.355
	(4.50)	(2.63)
Percent of Comm. x Staff	-0.885	-0.341
	(1.24)	(0.816)
District-Assignment Fixed Effects	Yes	Yes
Committee Fixed Effects	Yes	Yes
Congress Fixed Effects	Yes	Yes
No. of Unit Fixed Effects	980	980
R-Squared	0.315	0.046
N	5,154	5,154

*p<.05. Models are panel linear regression where the dependent variable is a district's ideology z-score on committee *j* in congress *t* (model 1), or a district's absolute ideological z-score (model 2). Standard errors clustered by district (number of clusters equals number of unit fixed effects).

Table C2: District-Assignment Two-Way Fixed Effects Estimates of the Conditional Effect of Vote Share on Ideology, 104th-114th Congresses

	DV=Ideology	**DV=Absolute Ideology**
	(1)	(2)
District Dem. Presidential Support	-5.90*	
	(0.724)	
Vote Share	0.014	-0.170*
	(0.111)	
Absolute District Presidential Support		0.155
		(0.463)
Legislator Party (GOP=1)		-0.062*
		(0.024)
Percent of Committee Controlled by Majority	0.042	0.576
	(0.988)	(0.559)
Committee Staff (Logged)	-0.058	0.025
	(0.076)	(0.036)
District Dem. Support x Vote Share	3.21*	0.622
	(0.838)	(0.612)
District-Assignment Fixed Effects	Yes	Yes
Committee Fixed Effects	Yes	Yes
Congress Fixed Effects	Yes	Yes
No. of Unit Fixed Effects	980	980
R-Squared	0.328	0.045
N	5,128	5,128

*p<.05. Models are panel linear regression where the dependent variable is a district's ideology z-score on committee *j* in congress *t* (model 1), or a district's absolute ideological z-score (model 2). Number of unit fixed effects indicate number of observed districts. Standard errors clustered by district (number of clusters equals number of unit fixed effects).

Appendix D: Main Results Separated by Constituent and Non-Constituent Oriented Committee

The dependent variables capture overall district ideology. It is possible that within-issue ideology for some districts is orthogonal to overall district ideology. This may be especially true for issues which are distributive in nature. As a robustness check, I replicate the main results separating committee jurisdiction using the Frisch and Kelly (2004) and Deering and Smith (1997) committee classification scheme which characterizes some committees as "constituent oriented". These committees are: Agriculture, Armed Services, Natural Resources, Transportation and Infrastructure, Science, Space, and Technology, and Small Business. To these, I add Appropriations, which Deering & Smith/Frisch & Kelly classify as a "prestige" committee. There are 3,305 observations (64.03%) not on these committees and 1,857 observations (35.97%) on these committees. As the results in Table D1 show, the results for each set of committees are nearly identical.

Table D1: District-Assignment Estimates for Ideology, 104th-114th Congresses

	DV=Ideology			DV=Absolute Ideology		
	Constituent Committees					
	(1)	(2)	(3)	(4)	(5)	(6)
District Dem. Presidential Support	-3.23*	-3.27*	-2.66*			
	(0.440)	(0.434)	(0.958)			
Absolute Dem. Pres. Support				0.583*	0.544*	0.189
				(0.188)	(0.186)	(0.332)
Legislator Party (GOP=1)				-0.074	-0.085*	-0.052*
				(0.040)	(0.039)	(0.017)
District-Assignment Fixed Effects	Yes	Yes		Yes	Yes	
Committee Fixed Effects	Yes		Yes	Yes		Yes
Congress Fixed Effects	Yes		Yes	Yes		Yes
Committee x Congress Fixed Effects		Yes			Yes	
First Differences			Yes			Yes
No. of Unit Fixed Effects	579	579		579	579	
R-Squared	0.249	0.244	0.032	0.058	0.065	0.021
N	1,857	1,857	1,083	1,857	1,857	1,083
	Non-Constituent Committees					
	(1)	(2)	(3)	(4)	(5)	(6)
District Dem. Presidential Support	-3.13*	-3.22*	-2.25*			
	(0.438)	(0.421)	(0.527)			
Absolute Dem. Pres. Support				0.658*	0.629*	0.481
				(0.195)	(0.190)	(0.265)
Legislator Party (GOP=1)				-0.101*	-0.095*	-0.040*
				(0.038)	(0.039)	(0.012)
District-Assignment Fixed Effects	Yes	Yes		Yes	Yes	
Committee Fixed Effects	Yes		Yes	Yes		Yes
Congress Fixed Effects	Yes		Yes	Yes		Yes
Committee x Congress Fixed Effects		Yes			Yes	
First Differences			Yes			Yes
No. of Unit Fixed Effects	780	780		780	780	
R-Squared	0.337	0.330	0.020	0.039	0.068	0.007
N	3,305	3,305	2,209	3,305	3,305	2,209

*p<.05. Models are panel linear regression where the dependent variable is a district's ideology z-score on committee *j* in congress *t* (models 1-3), or a district's absolute ideological z-score (models 4-6). Number of unit fixed effects indicate number of observed districts. Standard errors clustered by district (number of clusters equals number of unit fixed effects, in models 3 and 6 number of clusters is 364).

Appendix E: Main Results Using Groseclose-Levitt-Snyder Adjusted Conservative Vote Probabilities and OC Scores.

This robustness check replicates the main results using conservative vote probabilities (Fowler and Hall 2012). Conservative vote probabilities indicate the probability a given legislator votes conservatively relative to the median legislator. The measure uses bill and legislator fixed effects to develop these probabilities. Note however, that the measure is not constrained between 0 and 1 because it captures the difference relative to the median legislator. For example, a value of .1 indicates a legislator is 10% more likely to vote conservatively than the median. The measure also requires an anchor (similar to W-NOMINATE scores); here, the anchor is the first Republican ICPSR identifier for any given set of votes.

To scale these across committees and congresses, I apply the transformation articulated by Groseclose, Levitt and Snyder Jr. (1999) in the context of ADA scores. These adjusted (sometimes called "turbo") ADA scores account for differences across chambers and congresses by allowing the ideological space to shift and stretch. (This stretching and shifting of the ideological space is also why OC scores should not be compared across time or across committees). The Groseclose-Levitt-Snyder adjustment corrects for these problems by deriving an index, which they compare to an inflation index. Adjusted conservative vote probabilities correlate with OC z-scores at .74, while adjusted OC scores correlate with OC z-scores at .88.

I apply the adjustment to both the conservative vote probabilities measure and the committee-congress level OC scores I create in the paper. As the tables below show, both show robust effects for the main results using the two-way fixed effects models. For conservative vote probabilities, a change in district Democratic presidential support from the minimum to the maximum results in a 168% (95% CI: 147% to 188%) decrease in the chances of voting conservative compared to the median voter. The results for the other specifications are substantively similar. Table E2 shows the same change in district Democratic presidential support decreases the

adjusted OC z-score by 2.31, similar to the main effect of 3.41 in the main results (Table 1). The effect of absolute district Democratic presidential support on adjusted OC z-scores is not statistically significant, but the coefficients are in the correct direction. Because OC Z-scores are more easily interpretable, I present those in the main results rather than the adjusted values. Further, adjusted OC z-scores may be over-determined or too manipulated to be substantively meaningful.

Finally, Tables E1 and E2 are identified using two-way effects with district as the panel. Consistent with the main results, there is no significant effect on district Democratic presidential support when the models are identified with legislators as the panels (results not shown, but code provided in replication file). Again, this suggests that legislators do not change their behavior over time to adapt to their district.

Table E1: District-Assignment Estimates for Adjusted Conservative Vote Probabilities, 104th-114th Congresses

	DV=Ideology			**DV=Absolute Ideology**		
	(1)	(2)	(3)	(4)	(5)	(6)
District Dem. Presidential Support	-2.12*	-2.15*	-1.77*			
	(0.132)	(0.132)	(0.195)			
Absolute Dem. Pres. Support				0.206*	0.209*	0.642*
				(0.045)	(0.046)	(0.139)
Legislator Party (GOP=1)				-0.653*	-0.655*	-0.08*
				(0.009)	(0.009)	(0.007)
District-Assignment Fixed Effects	Yes	Yes		Yes	Yes	
Committee Fixed Effects	Yes		Yes	Yes		Yes
Congress Fixed Effects	Yes		Yes	Yes		Yes
Committee x Congress Fixed Effects		Yes			Yes	
First Differences			Yes			Yes
No. of Unit Fixed Effects	963	963		963	963	
R-Squared	0.560	0.196	0.116	0.899	0.902	0.086
N	4,906	4,906	3,129	4,906	4,906	3,129

*p<.05. Models are panel linear regression where the dependent variable is a district's adjusted conservative vote probability *j* in congress *t* (models 1-3), or a district's absolute adjusted conservative vote probability (models 4-6). Number of unit fixed effects indicate number of observed districts. Standard errors clustered by district (number of clusters equals number of unit fixed effects, in models 3 and 6 number of clusters is 719).

Table E2: District-Assignment Estimates for Adjusted OC Scores, 104th-114th Congresses

	DV=Ideology			**DV=Absolute Ideology**		
	(1)	(2)	(3)	(4)	(5)	(6)
District Dem. Presidential Support	-2.31*	-2.34*	-1.32			
	(1.06)	(1.18)	(2.82)			
Absolute Dem. Pres. Support				1.83	0.877	3.72
				(1.16)	(0.623)	(2.51)
Legislator Party (GOP=1)				-0.50*	0.108	-0.12
				(0.189)	(0.009)	(0.069)
District-Assignment Fixed Effects	Yes	Yes		Yes	Yes	
Committee Fixed Effects	Yes		Yes	Yes		Yes
Congress Fixed Effects	Yes		Yes	Yes		Yes
Committee x Congress Fixed Effects		Yes			Yes	
First Differences			Yes			Yes
No. of Unit Fixed Effects	983	983		983	963	
R-Squared	0.002	0.02	0.06	0.194	0.737	0.173
N	5,162	5,162	3,332	5,162	4,906	3,332

*p<.05. Models are panel linear regression where the dependent variable is a district's adjusted OC scores *j* in congress *t* (models 1-3), or a district's absolute adjusted OC score (models 4-6). Number of unit fixed effects indicate number of observed districts. Standard errors clustered by district (number of clusters equals number of unit fixed effects, in models 3 and 6 number of clusters is 740).

Appendix F: Additional Details on the Committee Votes Data Collection Process

The data were scraped from committee report text and include member names and recorded vote positions, along with the vote type (vote to report or not), the committee report number, and the bill number to which the committee report pertains. Votes to report bills were identified using the vote descriptions contained in the committee reports. Names and committees were matched with Stewart's Committee Data (Stewart and Woon 2016) and with DW-NOMINATE data to identify individual members (Poole and Rosenthal 2007).

The text of committee reports was collected through a variety of sources, including Congress.gov and ProQuest's electronic collection of the U.S. Government's Serial Set. A number of points about the data bear mentioning. First, votes within House committees are classified according to the name of the committee as of the 114th Congress. Though the substantive focus and jurisdiction of committees largely stays the same, the names frequently change. For example, in the last few years, the Education and Labor Committee has been called the Economic and Educational Opportunities Committee and the Education and the Workplace Committee. I treat all votes taken in these committees as belonging to the same committees and I also consider the Select Committee on Homeland Security and the Homeland Security Committees as the same committee, as the Select Committee was created before the committee became permanent in 2005.

All roll call votes to report to the House floor taken in committee and listed in the report issued by the committee which accompanies the bill are recorded. Only committee reports which contain the term "vote" were examined for a committee vote. Far more committee reports mention the word "vote" than actually contain a recorded vote, but this strategy was used to screen out hundreds of committee reports which do not contain a roll call vote. A significant number of votes are reported in tables which are not machine-readable. These votes were hand-coded.

A number of votes are embedded in the committee report of a different committee. This is the case for a large number of votes taken by other committees and reported in the Budget Committee's report for reconciliation bills. Where identified, these votes are considered to be votes in the committee of record for the members, not for the Budget Committee. Finally, the level of detail in the extracted data varies significantly in how member names are reported with some committees reporting full names, some reporting last names, and some reporting states. In some committees where two members have the same last name and no state identification is given, members cannot be matched. These discrepancies were ignored if members voted the same or if they were of the same party as these two factors do not affect inferences on party votes as described below. Still, there are approximately 1,200 member-votes out of the roughly 325,000 total votes that cannot be matched due to insufficient information reported in the committee report. Given that a non-matched member is the result of a random process (i.e.g, having the same last name as another member on the same committee in the same Congress), these non-matched members should not affect the causal claims drawn.

Voice votes and unanimous consent votes are not included in the data for a number of reasons. First, some committees report voice votes and some do not, and even among the committees which do report voice votes, record-keeping appears to be inconsistent over time. Voice votes, when reported, are often mentioned in the text of the committee report rather than in the reporting votes section of the report. Unanimous consent votes do not appear to be recorded by any committee. Division votes are occasionally taken by committees, but not reported in the committee reports. Both division votes and unanimous consent votes are more likely on bills which are less substantive or important (e.g., naming federal buildings, commemorating people or events).

Additional Details on the Creation of OC Scores

OC scores differ from W-NOMINATE scores in that they maximize the correct classification of legislators' choices; it is a "non-parametric procedure that requires no assumptions about the parametric form of the legislators' preference functions, other than assuming that they are symmetric and single-peaked," (Lo 2020, 1, also see Poole 2000). OC scores are a single-dimension ideological measure ranging from most liberal (-1) to most conservative (1).

To be included in the sample, a legislator must have taken ten votes within a committee-congress. Further, while legislator i may cast more than ten votes, for some committee-congresses there are not enough other legislators to scale legislator i's voting record. See Table A1 in the Appendix for the number of legislators within each committee-congress who have an OC score.

References

Deering, Christopher J. and Steven S. Smith. 1997. *Committees in Congress*. Washington D.C.: Congressional Quarterly Press.

Fowler, Anthony and Andrew B. Hall. 2012. "Conservative Vote Probabilities: An Easier Method for the Analysis of Roll Call Data." Accessed at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2120720.

Frisch, Scott A. and Sean Q Kelly. 2004. "Self-Selection Reconsidered: House Committee Assignment Requests and Constituency Characteristics." *Political Research Quarterly* 57(2).

Groseclose, Timothy, Steven D. Levitt and James M. Snyder Jr. 1999. "Comparing Interest Group Scores across Time and Chambers: Adjusted ADA Scores for the US Congress." *The American Political Science Review* 93(1):33–50.

Poole, Keith T. and Howard Rosenthal. 2007. *Ideology and Congress*. 2nd ed. New Brunswick, New Jersey: Transaction Publishers.

Stewart, III, Charles and Jonathan Woon. 2016. "Congressional Committee Assignments, 103rd to 114th Congresses, 1993–2017: House of Representatives.". http://web.mit.edu/17.251/www/data_page.html.

Thomsen, Danielle M. 2009. "Ideological Moderates Won't Run: How Party Fit Matters for Partisan Polarization in Congress." *Journal of Politics* .