Online Appendix to Reducing Unequal Representation: The Impact of Labor Unions on Legislative Responsiveness in the US Congress

Michael Becher
Daniel Stegmueller

Contents

A. Data

In this appendix we present additional details on our dataset including details on the creation of some control variables and descriptive statistics.

Matched roll calls Table A.2 displays Congressional roll calls matched to CCES items. We selected congressional roll calls based on content and, when several choices were available, based on their proximity to CCES fieldwork periods.

Income thresholds Table A.1 presents an overview of the income thresholds we use to classify CCES respondents into income groups. We use two thresholds separating the lowest and highest income terciles. We calculate them from yearly American Community Survey files excluding individuals living in group quarters. For each congress, Table A.1 shows the average of all district-specific thresholds as well as the smallest and largest ones.

Public unions Public unions captured (by name) in our data include the American Federation of State, County & Municipal Employees, National Education Association, American Federation of Teachers, American Federation of Government Employees, National Association of Government Employees, United Public Service Employees Union, National Treasury Employees Union, American Postal Workers Union, National Association of Letter Carriers, Rural Letter Carriers Association, National Postal Mail Handlers Union, National Alliance of Postal and Federal Employees, Patent Office Professional Association, National Labor Relations Board Union, International Association of Fire Fighters, Fraternal Order of Police, National Association of Police Organizations, various local police associations, and various local public school unions.

Descriptive statistics Table A.3 shows descriptive statistics for all variables used in our analysis. Note that these are for the untransformed variables. In our empirical models, we standardize all inputs to have mean zero and unit standard deviation.

Table A.1
Distribution of district income-group reference points. Average threshold over all districts, smallest and largest value.

	33th percentile			67th percentile		
Congress	Mean	Min	Max	Mean	Min	Max
109	38123	16800	73675	77964	39612	146870
110	40127	18000	77000	83047	43600	155113
111	39021	17500	78262	82440	46000	160050
112	37381	16500	81000	79868	38500	158654

Note: Calculated from American Community Survey 1-year files. Household sample excluding group quarters. Missing income information imputed using Chained Random Forests.

Table A.2
Matched CCES–House roll calls included in our analysis.

CCES Match	Bill	Date	Name	House Vote (Yea-Nay)	Bill Ideology†
(1)	HR 810	07/19/2006	Stem Cell Research Enhancement Act (Presidential Veto override)	235-193	L
(1)	HR 3	01/11/2007	Stem Cell Research Enhancement Act of 2007 (House)	253-174	L
(1)	S 5	06/07/2007	Stem Cell Research Enhancement Act of 2007	247-176	L
(2)	HR 2956	07/12/2007	Responsible Redeployment from Iraq Act	223-201	L
(3)	HR 2	01/10/2007	Fair Minimum Wage Act	315-116	L
(4)	HR 4297	12/08/2005	Tax Relief Extension Reconciliation Act (Passage)	234-197	C
(4)	HR 4297	05/10/2006	Tax Relief Extension Reconciliation Act (Agreeing to Conference Report)	244-185	C
(5)	HR 3045	07/28/2005	Dominican Republic-Central America-United States Free Trade Agreement Implementation Act	217-215	C
(6)	S 1927	08/04/2007	Protect America Act	227-183	C
(6)	HR 6304	06/20/2008	FISA Amendments Act of 2008	293-129	C
(7)	HR 3162	08/01/2007	Children's Health and Medicare Protection Act	225-204	L
(7)	HR 976	10/18/2007	Children's Health Insurance Program Reauthorization Act (Presidential Veto Override)	273-156	L
(7)	HR 3963	01/23/2008	Children's Health Insurance Program Reauthorization Act (Presidential Veto Override)	260-152	L
(7)	HR 2	02/04/2009	Children's Health Insurance Program Reauthorization Act	290-135	L
(8)	HR 3221	07/23/2008	Foreclosure Prevention Act of 2008	272-152	L
(9)	HR 3688	11/08/2007	United States-Peru Trade Promotion Agreement	285-132	C
(10)	HR 1424	10/03/2008	Emergency Economic Stabilization Act of 2008	263-171	L
(11)	HR 3080	10/12/2011	To implement the United States-Korea Trade Agreement	278-151	C
(12)	HR 3078	10/12/2011	To implement the United States-Colombia Trade Promotion Agreement	262-167	C
(13)	HR 2346	06/16/2009	Supplemental Appropriations, Fiscal Year 2009 (Agreeing to conference report)	226-202	L
(14)	HR 2831	07/31/2007	Lilly Ledbetter Fair Pay Act	225-199	L
(14)	HR 11	01/09/2009	Lilly Ledbetter Fair Pay Act of 2009 (House)	247-171	L
(14)	S 181	01/27/2009	Lilly Ledbetter Fair Pay Act of 2009	250-177	L
(15)	HR 1913	04/29/2009	Local Law Enforcement Hate Crimes Prevention Act	249-175	L
(16)	HR 1	02/13/2009	American Recovery and Reinvestment Act of 2009 (Agreeing to Conference Report)	246-183	L
(17)	HR 2454	06/26/2009	American Clean Energy and Security Act	219-212	L
(18)	HR 3590	03/21/2010	Patient Protection and Affordable Care Act	220-212	L
(19)	HR 3962	11/07/2009	Affordable Health Care for America Act	221-215	L
(20)	HR 4173	06/30/2010	Wall Street Reform and Consumer Protection Act of 2009	237-192	L
(21)	HR 2965	12/15/2010	Don't Ask, Don't Tell Repeal Act of 2010	250-175	L
(22)	S 365	08/01/2011	Budget Control Act of 2011	269-161	C
(23)	H CR 34	04/15/2011	House Budget Plan of 2011	235-193	C
(24)	H CR 112	03/28/2012	Simpson-Bowles/Copper Amendment to House Budget Plan	38-382	C
(25)	HR 8	08/01/2012	American Taxpayer Relief Act of 2012 (Levin Amendment)	170-257	L
(26)	HR 2	01/19/2011	Repealing the Job-Killing Health Care Law Act	245-189	C
(26)	HR 6079	07/11/2012	Repeal the Patient Protection and Affordable Care Act and [...]	244-185	C
(27)	HR 1938	07/26/2011	North American-Made Energy Security Act	279-147	C

Note: The matching of roll calls to CCES items can be many-to-one.
† Coding of a bill's ideological character as (L)iberal or (C)onservative based on predominant support of bill by Democratic or Republican representatives, respectively.

Table A.3
Descriptive statistics of analysis sample

	Mean	SD	Min	Max	N
Roll-call vote: yea	0.568	0.495	0.000	1.000	15780
Constituent preferences					
Low income	0.593	0.220	0.047	0.979	15934
High income	0.555	0.198	0.037	0.967	15934
Low-High Gap	0.172	0.121	0.000	0.588	15934
Union membership [log]	9.705	1.046	6.094	13.619	15934
Population	7.022	0.723	4.697	9.980	15934
Share African American	0.124	0.146	0.004	0.680	15934
Share Hispanic	0.156	0.174	0.005	0.812	15934
Share BA or higher	0.275	0.097	0.073	0.645	15934
Median income [$10,000]	5.177	1.356	2.282	10.439	15934
Share female	0.508	0.010	0.462	0.543	15934
Manufacturing share	0.110	0.047	0.025	0.281	15934
Urbanization	0.790	0.199	0.213	1.000	15934
Social capital [bowling establish./10]	0.900	1.259	0.024	5.800	15934
Certification elections [log]	3.347	0.861	0.000	5.100	15934
Congregations [per 1000 persons]	0.765	1.147	0.062	6.453	15934

Note: Calculated from American Community Survey, 2006-2013. Note that when entered in models, variables are scaled to mean zero and unit SD. Preference gap is absolute difference in preferences between low and high income constituents in sample. Urbanization is calculated as the share of the district population living in an urban area based on the Census' definition of urban Census blocks (matched to congressional districts using the MABLE database). Congregations per 1000 inhabitants calculated from RCMS 2000 (spatially interpolated).

B. Estimation of District Preferences

In this section we describe how we estimate district-level preferences using three different strategies: (i) small area estimation using a matching approach based on random forests (which we use in the main text of our paper), (ii) estimation using multilevel regression and post-stratification (MRP), and (iii) unadjusted cell means. Each approach invokes different statistical and substantive assumptions. In the spirit of consilience, our aim here is to show that our substantive results do not depend on any particular choice.

B.1. Small Area Estimation via Chained Random Forests

The core idea of our small area estimation strategy is based on the fact that we have access to two samples: one that is likely not representative of the population of all Congressional districts (the CCES), while the second one is representative of district populations by virtue of its sampling design (the Census or American Community Survey). By matching or imputing preferences from the former to the latter based on a common vector of observable individual characteristics, we can use the district-representative sample to estimate the preferences of individuals in a given district.[1]

Combining CCES and Census data using Random Forests Figure B.1 illustrates this approach in more detail. We have data from m individuals in the CCES and n individuals in the Census (with $n \gg m$). Both sets of individuals share K common characteristics Z_k, such as age, race, or education. The first task at hand is then to match P roll call preferences Y_p that are only observed in the CCES to the census sample. This is a purely predictive task and it is thus well suited for machine learning approaches. We use random forests (Breiman 2001) to lean about $Y_p = f(Z_1, \ldots, Z_K)$ for $p = 1, \ldots, P$ using the algorithm proposed by Stekhoven and Bühlmann (2011). This approach has two key advantages. First, as is typical for approaches based on regression trees, it deals with both categorical and continuous data, allows for arbitrary functional forms, and can include higher order interactions between covariates (such as age×race×education). Second, we can assess the quality of the predictions based on our model before we deploy it to predict preferences in the Census. With the trained model in hand we can use $\widehat{f}(Z_1, \ldots, Z_K)$ in combination with observed Z in the Census sample to fill in preferences (i.e., completing the square in the lower right of Figure B.1). Using the completed Census data, we can estimate constituent district preferences as simple averages by district and income group since the Census sample is representative for each Congressional district's population.

Data details Due to data confidentially constraints the Census Bureau does not provide district identifiers in its micro-data records. Instead, it identifies 630 Public Use Microdata areas. We

[1] See Honaker and Plutzer (2016) for a more explicit exposition of this idea, evidence for its empirical reliability, and a comparison to MRP estimates.

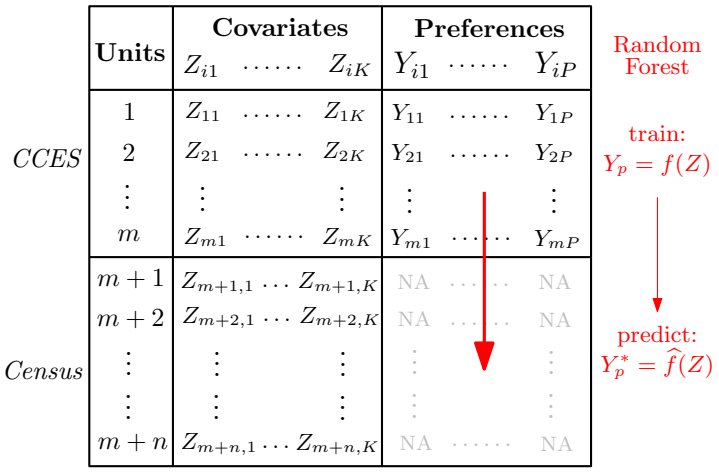

Figure B.1
Illustration of Small Area Estimation of District Preferences.

We use a sample of m individuals from the CCES that is not necessarily representative on the district-level, while a sample of n individuals from the Census is representative of district populations by design (Torrieri et al. 2014: Ch.4). We have access to bridging covariates Z_k that are common to both samples, while roll call preferences Y_p are only observed in the CCES. We train a flexible non-parametric model relating Y_p to Z and use it to predict preferences Y_p^* for Census individuals with characteristics Z. With preference values filled in, a district's income-group specific roll call preference can be estimated as the average of all units in that district.

create a synthetic Census sample for Congressional districts by sampling individuals from the full Census PUMA regions proportional to their relative share in a given districts. This information is based on a crosswalk from PUMA regions to Congressional districts created by recreating one from the other based on Census tract level population data in the MABLE Geocorr2K database. The 'donor pool' for this synthetic sample are the 1% extracts for the American Community Survey 2006-2011. We limit the sample to non-group quarter households and to individuals aged 17 and older providing us with data on 14 million (13,711,248) Americans. From this we create the synthetic district file which is comprised of 3,040,265 cases. This provides us with a Census sample including Congressional district identifiers. The sample for each district is representative of the district population (save for errors induced by the crosswalk). We thus use the distribution of important population characteristics (age, gender, education, race, income) to match data on policy preferences from the CCES.

We harmonize all covariates to be comparable between CCES and Census. For family income this entails an adjustment to the measure provided in the CCES. It asks respondents to place their family's total household income into 14 income bins.[2] We transform this discretized measure of income into a continuous one using a nonparametric midpoint Pareto estimator. It

[2]The exact question wording is: "Thinking back over the last year, what was your family's annual income?" The obvious issue here is that it is not clear which income concept this refers to (or, rather, which on the respondent employs). In line with the wording used in many other US surveys, we interpret it as referring to market income.

replaces each bin with its midpoint (e.g., the third category $20,000 to $29,999 gets assigned $25,000), while the value for the final, open-ended, bin is imputed from a Pareto distribution (e.g., Kopczuk, Saez, and Song 2010). Using midpoints has been recognized for some time as an appropriate way to create scores for income categories (without making explicit distributional modeling assumptions). They have been used extensively, for example, in the American politics literature analyzing General Social Survey (GSS) data (Hout 2004).

Algorithm details For easier exposition define a matrix D that contains both individual characteristics and roll call preferences. Let N be the number of rows of D. For any given variable v of D, D_v, with missing entries at locations $i_{mis}^{(v)} \subseteq \{1,\ldots,N\}$ we can separate out four parts:[3]

- Observed values of D_v: denoted as $y_{obs}^{(v)}$
- Missing values of D_v: $y_{mis}^{(v)}$
- Variables other than D_v with available observations $i_{obs}^{(v)} = \{1,\ldots,N\} \setminus i_{mis}^{(v)}$: $x_{obs}^{(v)}$
- Variables other than D_v with observations $i_{mis}^{(v)}$: $x_{mis}^{(v)}$

We now cycle through variables iteratively fitting random forest and filling in unobserved values until a stopping criterion c (indicating no further change in filled-in values) is met. Algorithmically, we proceed as follows:

Algorithm 1 Chained Random Forests

```
Start with initial guesses of missing values in D
w ← vector of column indices sorted by increasing fraction of NA
while not c do
    D_old^imp ← previously imputed D
    for v in w do
        Fit Random Forest: y_obs^(v) ~ x_obs^(v)
        Predict y_mis^(v) using x_mis^(v)
        D_new^imp ← updated imputed matrix using predicted y_mis^(v)
    Updated stopping criterion c
Return completed D^imp
```

To assess the quality of this scheme, we inspect the prediction error of the random forests using the out-of-bag (OOB) estimate (which can be obtaining during the bootstrap for each tree). We find it to be rather small in our application: most normalized root mean squared errors are around 0.11. This result is in line with simulations by Stekhoven and Bühlmann (2011) who compare it to other prediction schemes based on K nearest neighbors, EM-type

[3]Note that this setup deals transparently with missing values in individual characteristics (such as missing education).

LASSO algorithms, or multivariate normal schemes and find it to perform comparatively well with both continuous and categorical variables.[4]

B.2. Multilevel Regression and Poststratification

The approach described in the last section is closely related to MRP (Gelman and Little 1997; Park, Gelman, and Bafumi 2006; Lax and Phillips 2013), which has become quite popular in political science. Both strategies involve fitting a model that is predictive of preferences given observed characteristics followed by a weighting step that re-balances observed characteristics to their distribution in the Census. What differentiates MRP from the previous approach is that it imposes more structure in the modeling step both in terms of functional form and distributional assumptions. By utilizing the advantages of hierarchical models with normally distributed random coefficients it produces preference estimates that are shrunken towards group means (Gelman et al. 2013: 116f.).[5] No such structural assumptions are made when using Random Forests. It will thus be instructive to compare how much our results depend on such modeling choices.

MRP implementation For each roll call item in the CCES we estimate a separate model expressing the probability of supporting a proposal as a function of demographic characteristics. The demographic attributes included in our model broadly follow Lax and Phillips (2009, 2013) and are race, gender, education, age, and income.[6] Race is captured in three categories (white, black, other), education in five (high school or less, some college, 2-year college degree, 4-year college degree, graduate degree). Age is comprised of 6 categories (18-29, 30-39, 40-49, 50-59, 60-69, 70+) while income is comprised of 13 categories (with thresholds 10, 15, 20, 25, 30, 40, 50, 60, 70, 80, 100, 120, 150 [in $1,000]). Our model also includes district-specific intercepts. For each roll-call, we estimate the following hierarchical model using penalized maximum likelihood (Chung et al. 2013):

$$Pr(Y_i = 1) = \text{logit}^{-1}\left(\beta^0 + \alpha^{race}_{j[i]} + \alpha^{gender}_{k[i]} + \alpha^{age}_{l[i]} + \alpha^{educ}_{m[i]} + \alpha^{income}_{n[i]} + \alpha^{district}_{d[i]}\right) \tag{B.1}$$

We employ the notation of Gelman and Hill (2007) and denote by $j[i]$ the category j to which individual i belongs. Here, β_0 is an intercept and the αs are hierarchically modeled effects for

[4] See Tang and Ishwaran (2017) for further empirical validation of this strategy. See also Honaker and Plutzer (2016), who compare a similar matching strategy (but based on a multivariate normal model) with MRP estimated preferences using the CCES.

[5] This might be especially appropriate when some groups are small. The median number of respondents per district in the CCES is 506 and no district has fewer than 192 sampled respondents. But since we slice preferences further by income sub-groups, one may be worried that the sample size in some districts is small. MRP deals with this potential issue at the cost of making distributional assumptions.

[6] We also estimated a version of the model including a macro-level predictor, which has been found to improve the quality of the model. We use the demographically purged state predictor of Lax and Phillips (2013: 15), that is, the average liberal–conservative variation in state-level public opinion that is not due to variation demographic predictors. In our case this produces rather similar MRP estimates.

the various demographic groups. Each is drawn from a common normal distribution with mean zero and estimated variance σ^2:

$$\alpha_j^{race} \sim N\left(0, \sigma^2_{race}\right), \quad j = 1, \ldots, 3 \tag{B.2}$$

$$\alpha_k^{gender} \sim N\left(0, \sigma^2_{gender}\right), \quad k = 1, \ldots, 2 \tag{B.3}$$

$$\alpha_l^{age} \sim N\left(0, \sigma^2_{age}\right), \quad l = 1, \ldots, 6 \tag{B.4}$$

$$\alpha_m^{educ} \sim N\left(0, \sigma^2_{educ}\right), \quad m = 1, \ldots, 5 \tag{B.5}$$

$$\alpha_n^{income} \sim N\left(0, \sigma^2_{income}\right), \quad n = 1, \ldots, 13 \tag{B.6}$$

This setup induces shrinkage estimates for the same demographic categories in different districts. Note that using fixed effects for characteristics with few categories (Specifically, gender) does not impact our results. The district intercepts are drawn from a normal distribution with state-specific means $\alpha_{s[d]}$ and freely estimated variance:

$$\alpha_d \sim N\left(\alpha^{state}_{s[d]}, \sigma^2_{state}\right). \tag{B.7}$$

Our final preferences estimates for each income group on each roll call are obtained by using cell-specific predictions from the above hierarchical model, weighted by the population frequencies (obtained from our Census file) for each cell in each congressional district.

B.3. Model results under various preference estimation strategies

The estimates of district-level preferences obtained via our SAE approach and MRP are in broad agreement: The median difference in district preferences between SAE and MRP is 2.5 percentage points for low income and -0.1 percentage points for high income constituents. A large part of this difference is due to the heavier tails of the distribution of district preferences for each roll call estimated by our approach—perhaps not surprising given the shrinkage characteristics of MRP. To what extent do these differences in the distribution of preferences affect our estimated union effects?

Panel (A) of Table B.1 shows estimates for our six main specifications using MRP-based preferences. The results are unequivocal: using MRP estimated preferences leads to more pronounced estimates in all specifications. Using specification (6), which includes state policies, measures of district social capital and district covariates interacted with preferences, as well as district fixed effects, we find that a unit increase in union membership increased responsiveness of legislators towards the preferences of low income constituents by about 8 (± 2) percentage points (compared to only 5 points using our measurement strategy). Responsiveness estimated for high income preferences are similarly larger. Note that while larger, all estimates also carry increased confidence intervals.

As a further point of comparison, panel (B) shows preferences estimated via raw cell means in the CCES. Due to the issues discussed above, the raw data cannot be taken as gold

Table B.1
Model results using different strategies to estimate district-level preferences.

	(1)	(2)	(3)	(4)	(5)	(6)
A: Multilevel Regression & Poststratification						
Low income preferences	0.182	0.158	0.181	0.185	0.115	0.081
	(0.021)	(0.024)	(0.026)	(0.022)	(0.022)	(0.022)
High income preferences	−0.136	−0.119	−0.139	−0.137	−0.091	−0.064
	(0.017)	(0.019)	(0.021)	(0.017)	(0.018)	(0.018)
B: Raw CCES means						
Low income preferences	0.080	0.061	0.063	0.080	0.043	0.028
	(0.010)	(0.011)	(0.012)	(0.010)	(0.011)	(0.011)
High income preferences	−0.027	−0.013	−0.010	−0.025	−0.018	−0.011
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)	(0.009)

Note: Specifications (1) to (6) are as in Table I in the main text but using different strategies to estimate district-level preferences of three income groups. Entries are effects of standard deviation increase in union membership on marginal effect of income group preferences on legislator vote.

standard, but it is nonetheless informative to see how much the results vary. Our core results even obtain when we simply use raw cell means without any statistical modeling to counter non-representative distributions of individual characteristics and small cell sizes. We find that in our strictest specification, a unit increase in union membership still increases responsiveness towards low income constituents by about 3 (±1) percentage points.

In sum, all three approaches lead to the same qualitative conclusions about the moderating effect of unions on unequal representation in Congress. The two alternative approaches to deal with the problem that CCS surveys are not representative of congressional districts by design suggest that a larger effect of unions than the naive approach using the unadjusted survey data. Quantitatively, our preferred estimates are based on small area estimation via random forests as they are less reliant on normality assumptions and are systematically more conservative than those based on MRP.

C. Alternative Income Thresholds

This section discusses the impact of different income thresholds on our results. In Table I in the main text, preferences of income groups are based on a *district-specific* income thresholds splitting the population into three groups (at the 33rd and 66th percentile). Thus, voters are classified as 'low income' relative to other voters in their congressional district. For example, during the 111th Congress a voter with an income of $40,000 would be part of the low income group in most of Massachusetts' districts (where low income thresholds vary from about $40,000

to $50,000), but not in the 8th (where the threshold is about $30,000). If income threshold were *state-specific* instead, he or she would be considered low income everywhere in the state (as the state-specific low income threshold is now ≈$47,000).

Table C.1
Model results using different definitions of income groups.

	(1)	(2)	(3)	(4)	(5)	(6)
A: State-specific income thresholds						
Low income preferences	0.105	0.082	0.097	0.107	0.067	0.044
	(0.013)	(0.015)	(0.016)	(0.013)	(0.014)	(0.014)
High income preferences	−0.062	−0.036	−0.052	−0.065	−0.049	−0.027
	(0.012)	(0.013)	(0.014)	(0.013)	(0.013)	(0.013)
B: Shifted income thresholds: p20 - p80						
Low income preferences	0.098	0.077	0.09	0.100	0.063	0.042
	(0.012)	(0.013)	(0.014)	(0.012)	(0.013)	(0.012)
High income preferences	−0.054	−0.031	−0.046	−0.057	−0.044	−0.025
	(0.011)	(0.012)	(0.012)	(0.011)	(0.012)	(0.012)

Note: Specifications (1) to (6) are as in Table I in the main text but with income groups defined via different income thresholds. Entries are estimates for η^l and η^h with cluster-robust standard errors in parentheses.

Not all states display as much variation in income-group thresholds. Thus, using state- instead of district-specific thresholds does not alter our core results in an appreciable way. As Panel (A) shows, the resulting marginal effects estimates for all six model specifications are remarkably similar when using preferences of income groups defined by state-specific thresholds. In panel (B) we no longer divide the population into three equally sized income groups. Instead, we restrict the low-income group to only those below the 20th percentile of the (district-specific) income distribution. Similarly, we classified as high income only those above the 80th percentile. Our resulting estimates for the union-responsiveness marginal effects are slightly smaller, but still of a substantively relevant magnitude and statistically different from zero.

D. Measures of District Organizational Capacity

In the empirical analysis reported in this Appendix, we use the number of religious congregations as another proxy for associational life, complementing the social capital measure used in the main text. In a previous version of the paper, we also use certification elections as a proxy for unions' mobilization capacity. Here we provide some background and explain in more detail how we calculate both variables.

Congregations As one proxy for district level social capital we use the number of congregations per inhabitant. The number of congregations in a given district is not readily available for the years covered in our study. Therefore, we spatially aggregate county-level measures from the 2010 Religious Congregations and Membership Study to the congressional district level using areal interpolation techniques that take into account the population distribution between counties and districts. We use a geographic country-to-district equivalence file calculated from Census shapefiles. This is combined with population weights for each country-district intersection derived using the Master Area Block Level Equivalency database v1.3.3 (available from the Missouri Census Data Center), which calculates them based on about 5.3 million Census blocks. With these weights in hand we can interpolate county-level to district-level congregation counts using weighted means (for states with at-large districts, this reduces to a simple summation, as counties are perfectly nested within districts).

NLRB certification elections In a previous version of the paper, we also used union certification elections as a proxy for workers' capability to organize for collective action. As has been pointed out by readers and discussants, one concern with this variable is that it may be driven to a significant extent by the existing stock of local unions, as unionization requires people and resources. While it may be useful to distinguish realized union membership from unionization effort and our results are robust to accounting for NLRB elections, in line with the suggestions we dropped this robustness test. For completeness and consistency, we document the construction of the measure.

The formation of unions is regulated by the National Labor Relations Act (NLRB) enacted in 1935 (see Budd 2018: ch. 6). A successful union organization process usually requires an absolute majority of employees voting for the proposed union in a certification election held under the guidelines of the NLRB. Getting the NLRB to conduct an election requires that there is sufficient interest among employees in an appropriate bargaining unit to be represented by a union. For proof of sufficient interest, the NLRB requires that at least 30% of employees sign an authorization card stating they authorize a particular union to represent them for the purpose of collective bargaining. Building support and collecting the required signatures takes organizational effort. For workers, unionization has features of a public good. Everybody may gain through better conditions from collective bargaining, but contributing to the organizational drive is costly for each individual. Beyond mere opportunity costs, there also is a non-zero risk of being (illegally) fired by the employer for those especially active. If more than 50% of employees sign authorization cards, then the union can request voluntary recognition without a certification election. However, the employer has the right to deny this, in which case a certification election is held. In his labor relations textbook, Budd (2018: 199) notes that voluntary card check recognition is "the exception rather than the norm because employers typically refuse to recognize unions voluntarily."

We use the NLRB's database on election reports to extract all attempts to certify (or de-certify) a local union. They are available from `www.nlrb.gov`. Each database entry is a vote concerning a bargaining unit; the average unit size is 25 employees. There are about 2200

elections each year. Each individual case file usually provides address information on the employer and the site where the election was held. Using this information, we geocode each individual case report and locate it in a congressional district.

E. Additional Robustness Tests

In this section we describe several additional robustness tests.

Redistricting First, we address the fact that several cases of court-ordered redistricting in Georgia and Texas lead to inter-Census changes in district boundaries. We exclude both states in specification (1) of Table E.1 and find our results unchanged.

Alternative measures of social capital Next we consider two alternative measures of social capital. First, the number of bowling alleys in an area (Putnam 2000). As the social capital index used in the main text, this variables comes from Rupasingha and Goetz (2008) and was spatially reweighted to districts. Second, the number of congregations per inhabitant. The construction of this measure is explained above (Appendix D). These two measures are less likely than the social capital index to be the result of unionization. We find that these changes in measurement do not qualitatively alter our findings. Unsurprisingly, the estimated union effects are somewhat larger than in the specification adjusting for the social capital index.

1:1 mapping of CCES preferences to roll calls We begin by limiting our sample by creating a unique mapping between preferences and roll call votes. Some of our CCEs preferences estimates are linked to more than one Congressional roll call. To investigate if this affects our results, specification (3) uses a 1:1 map dropping additionally available roll calls after the first match. This reduces the sample size to 11,104 respondents. We find that our results are not influenced by this change.

Extreme preferences excluded In specification (4) we investigate if extreme district preferences on some roll calls drive our results. To do so, we trim the distribution of preferences at the bottom and the top. For each roll call we exclude districts with preference estimates below the 5th and above the 95th percentile. Using only trimmed preferences has no appreciable impact on our estimates.

New York excluded Another test estimates our model with the state of New York excluded from the sample. While Becher, Stegmueller, and Kaeppner (2018) found that LM form estimates of union strength correlate highly with aggregated state-level estimates derived from the Current Population survey, they note that this correlation is lower for New York. In specification (5) we thus show that our results are not affected by its exclusion.

Union Concentration Our data on local unions are from Becher, Stegmueller, and Kaeppner (2018), who also find that the local concentration of unions is an important dimension. While Becher, Stegmueller, and Kaeppner (2018) show that both dimensions (membership

Table E.1
Additional robustness tests. LPM coefficients with robust standard errors in parentheses.

	Low income preferences		High income preferences		N
(1) Redistricting	0.067	(0.014)	−0.051	(0.013)	12,784
(2) Social capital: churches	0.072	(0.015)	−-0.051	(0.014)	14,282
(3) Injective preference roll call map	0.063	(0.013)	−0.041	(0.013)	11,104
(4) Extreme preferences excl.	0.074	(0.016)	−0.048	(0.015)	13,308
(5) New York excluded	0.070	(0.015)	−0.048	(0.014)	14,730
(6) Local Union Concentration	0.065	(0.014)	−0.047	(0.014)	15,780
(7) Trimmed LPM estimator	0.074	(0.015)	−0.055	(0.014)	15,426
(8) Errors-in-variables	0.062	(0.004)	−0.054	(0.004)	15,345
(9a) No fixed effects	0.068	(0.014)	−0.041	(0.013)	14,282
(9b) Two-way fixed effects (roll calls)	0.060	(0.014)	−0.040	(0.013)	14,282
(10a) CCES 2006-based roll calls excl.	0.065	(0.014)	−0.043	(0.015)	11,180
(10b) Influential roll calls excluded	0.073	(0.015)	−0.057	(0.014)	12,367

Note: Based on specification (5) of Table I. Entries are estimates for η^l and η^h with cluster-robust standard errors in parentheses, except for (8) which is estimated in the Bayesian framework (entries are posterior means and standard deviations). See text for all specification details.

and concentration) vary independently, it is prudent to check if our results on the impact of union membership on representation still obtain when accounting for the structure of union organization. In specification (6) we show this to be the case.

Trimmed LPM estimator A seventh, more technical, specification implements the trimmed estimator suggested by Horrace and Oaxaca (2006). It accounts for the fact that we estimate a linear probability model to a binary dependent variable, which entails the possibility that the model-implied linear predictor lies outside the unit interval. Our results in Table E.1 indicate that this change does not materially affect our core results (if anything, they become slightly larger).

Errors-in-variables Our penultimate test accounts for the errors-in-variables problem caused by the fact that our district preference measures are based on estimates. While, in general, standard errors for our district-level estimates are quite small relative to the quantity being measured and one expects a downward bias in parameter estimates in a linear model with errors-in-variables, we estimate this specification to get a sense of the quantitative magnitude of the change in parameter estimates.[7] We find that adjusting for measurement error produces very

[7]We implement this model in a Bayesian framework, where we incorporate the measurement error model directly into the posterior distribution. To specify the variance of the measurement error for low and high income group preferences, we average the standard errors of the district-group means from the raw CCES data (pre-Census matching). Measurement error variance is slightly larger for low income preferences (0.029) than for high

little quantitative change; both estimates are within the confidence bounds of our non-corrected estimates.

Different fixed effects specifications In our main models we include district fixed effects in order to capture the possibility that there are (district-specific) systematic differences between legislators and survey respondents on the same issues (Hill and Huber 2019). However our main analysis does not depend on the presence of fixed effects (partly due to the fact that there is ideological variation in the content of the bills studied, partly due to our use of demanding interactive controls). In an empty model of roll call votes and preferences the correlation of the fixed effects with the linear predictor is 0.05, which drops to 0.01 in a specification with all controls. This is confirmed in specification (9a) which excludes district fixed effects and produces results very similar to those reported in the main text.

Alternatively, one can turn to a more demanding specification where fixed effects capture a larger fraction of district×roll call-specific unobservables. We do so in specification (9b) where we estimate a two-way fixed effects model, which adds roll-call fixed effects. The correlation between roll-call fixed effects and the linear predictor is −0.37 (after including a full set of preference-control interactions), which suggests a higher relevance of this second set of fixed effects. However, our estimates in Table E.1 show that this more demanding set-up does not substantively alter our conclusions (this specification brings our estimate close to the post-double LASSO selection estimate which uses more flexible functional forms of covariates to reduce omitted variable bias).

Influential roll calls Our main model includes preference estimates using CCES waves 2006 to 2012 in order to cover a broad range of policy issues. Even though the quality of the CCES is generally high and the assumptions needed to construct model-based estimates are comparable to those needed to properly model non-response in classical phone (RDD) surveys, one of our reviewers pointed us to "teething" problems with the first wave (cf. the discussion in Hill et al. 2007; Vavreck and Rivers 2008). We inspected if roll calls matched to survey responses including the CCES 2006 wave show systematically different responsiveness estimates by extending our main model with $\eta_l \times \xi_l CCES_{06}$ and $\eta_h \times \xi_h CCES_{06}$ terms ($CCES_{06}$ is an indicator variable marking roll calls matched to preference estimates involving the 2006 wave). A joint F-test of ξ_l, ξ_h yields a value of $F = 2.64$ with a corresponding p-value of 0.073 providing limited evidence for a systematic deviation. More straightforwardly, we re-estimated our main model excluding any roll call for which citizen preference estimates involve the 2006 wave of the

income preferences (0.025). We use the setup proposed in Richardson and Gilks (1993), implemented in Stan (v.2.17.0) and estimated (due to the size of our data set) using mean field variational inference. We use normal priors with mean zero and standard deviation (SD) of 100 for all regression coefficients, and inverse Gamma priors with shape and scale 0.01 for residuals. In the measurement error equation, we use normal priors with mean zero and SD of 10 for the mean of the measurement error and a student-t prior with 3 degrees of freedom and mean 1, SD 10 for the standard deviation of the measurement. The reported entries are posterior means and standard deviations.

CCES. The resulting estimates, in specification (10a) of Table E.1, show that our substantive conclusions do not differ from the ones reported in Table I.

More generally, we examine if specific roll calls are overly influential for our responsiveness estimates. Beyond the impact of specific CCES waves, this might be the result of differential measurement bias on some items, for example, when citizens are uninformed on certain issues or assign them low priority and/or their representatives face strategic voting incentives (Hill and Huber 2019: 614). Instead of creating a classification of 'importance' or 'difficulty' of roll call votes for citizens (which is possibly heterogenous over districts), we estimate influence statistics for each roll call. This allows us to identify influential roll calls and exclude them from our analysis as a robustness check. We calculate roll call-specific leverage statistics for low and high income preferences. We use DFBETA as a measure of the standardized absolute difference between the estimate with a roll call included and the estimate without it (Belsley, Kuh, and Welsch 1980). We do not find that any roll call is particularly influential for our estimates of responsiveness to low and high income groups. The median influential roll call shifts our estimate of low income responsiveness by +0.012 standard errors and our estimate of high income responsiveness by −0.027 standard errors. Nevertheless, we selected all roll calls whose influence statistic exceeded 0.25 (i.e., shifting our estimate by more than a quarter of a standard error) and excluded them from the analysis. The resulting estimates in specification (10b) show a slightly increased level of responsiveness towards the preferences of low income citizens (which, however, still lies within the confidence bound of our preferred specification in the main text).

F. Post Double Selection Estimator

The post-double-selection model provides a relaxation of the linearity and exogeneity assumptions made in the baseline specification. To do so we use the double-post-selection estimator proposed by Belloni et al. (Belloni, Chernozhukov, and Hansen 2013; Belloni et al. 2017). Specifically, this model setup aims to reduce the possible impact of omitted variable bias by accounting for a large number of confounders in the most flexible way possible. This can be achieved by moving beyond restricting confounders to be linear and additive, and instead considering a flexible, unrestricted (non-parametric) function. This leads to the formulation of the following partially linear model (Robinson 1988) equation (for ease of exposition we omit district fixed effects in the notation and ignore i subscripts):

$$y_{jd} = \mu^l \theta^l_{jd} + \mu^h \theta^h_{jd} + \eta^l U_d \theta^l_{jd} + \eta^h U_d \theta^h_{jd} + g(Z_d) + \epsilon_{jd} \tag{F.1}$$

with $E(\epsilon_{jd}|Z_s, U_d, \theta_{jd}) = 0$. Here, y is the vote of a representative in a given district, U_d is the level of union density. The function $g(Z_d)$ captures the possibly high-dimensional and nonlinear influence of confounders (interacted with income group preferences). The utility of this specification as a robustness tests stems from the fact that it imposes no a priori restriction

on the functional form of confounding variables. A second key ingredient in a model capturing biases due to omitted variables is the relationship between the treatment (union density) and confounders. Therefore, we consider the following auxiliary treatment equation

$$U_d = m(Z_d) + v_i, \quad E(v_i|Z_d = 0), \tag{F.2}$$

which relates treatment to covariates Z_d. The function $m(Z_d)$ summarizes the confounding effect that potentially create omitted variable bias if $m \neq 0$, which is to be expected in an observational study such as ours.

The next step is to create approximations to both $g(\cdot)$ and $m(\cdot)$ by including a large number (p) of control terms $w_d = P(Z_d) \in \mathbb{R}^p$. These control terms can be spline transforms of covariates, higher order interaction terms, etc. Even with an initially limited set of variables, the number of control terms can grow large, say $p > 200$. To limit the number of estimated coefficients, we assume that g and m are approximately sparse (Belloni, Chernozhukov, and Hansen 2013) and can be modeled using s non-zero coefficients (with $s \ll p$) selected using regularization techniques, such as the LASSO (see Tibshirani 1996; see Ratkovic and Tingley 2017 for a recent exposition in a political science context):

$$y_{jd} = \mu^l \theta^l_{jd} + \mu^h \theta^h_{jd} + \eta^l U_d \theta^l_{jd} + \eta^h U_d \theta^h_{jd} + w'_d \beta_{g0} + r_{gd} + \zeta_{jd} \tag{F.3}$$

$$U_d = w'_d \beta_{m0} + r_{mi} + v_d \tag{F.4}$$

Here, r_{gi} and r_{mi} are approximation errors.

However, before proceeding we need to consider the problem that variable selection techniques, such as the LASSO, are intended for prediction, not inference. In fact, a "naive" application of variable selection, where one keeps only the significant w variables in equation (F.3) fails. It relies on perfect model selection and can lead to biased inferences and misleading confidence intervals (see Leeb and Pötscher 2008). Thus, one can re-express the problem as one of prediction by substituting the auxiliary treatment equation (F.4) for D_d in (F.3) yielding a reduced form equation with a composite approximation error (cf. Belloni, Chernozhukov, and Hansen 2013). Now both equations in the system represent predictive relationships and are thus amenable to high-dimensional selection techniques.

Note that using this dual equation setup is also necessary to guard against variable selection errors. To see this, consider the consequence of applying variable selection techniques to the outcome equation only. In trying to predict y with w, an algorithm (such as LASSO) will favor variables with large coefficients in $\bar{\beta}_0$ but will ignore those of intermediate impact. However, omitted variables that are strongly related to the treatment, i.e., with large coefficients in β_{m0}, can lead to large omitted variable bias in the estimate of η even when the size of their coefficient in $\bar{\beta}_0$ is moderate. The Post-double selection estimator suggested by Belloni, Chernozhukov, and Hansen (2013) addresses this problem, by basing selection on *both* reduced form equations. Let $\hat{I}_1$ be the control set selected by LASSO of y_{jd} on w_d in the first predictive equation, and let $\hat{I}_2$ be the control set selected by LASSO of U_d on w_d in the second equation. Then, parameter

estimates for the effects of union density and the regularized control set are obtained by OLS estimation of equation (F.1) with the set $\hat{I} = \hat{I}_1 \cup \hat{I}_2$ included as controls (replacing $g(\cdot)$). In our implementation we employ the root-LASSO (Belloni, Chernozhukov, and Wang 2011) in each selection step.

This estimator has low bias and yields accurate confidence intervals even under moderate selection mistakes (Belloni and Chernozhukov 2009; Belloni, Chernozhukov, and Hansen 2014).[8] Responsible for this robustness is the indirect LASSO step selecting the U_d-control set. It finds controls whose omission leads to "large" omitted variable bias and includes them in the model. Any variables that are not included ("omitted") are therefore at most mildly associated to U_d and y_{jd}, which decidedly limits the scope of omitted variable bias (Chernozhukov, Hansen, and Spindler 2015).

G. Nonparametric Evidence for Union Preferences Interaction

As discussed in the main text, we want to estimate a specification that makes as little *a priori* assumptions about functional form relationships between variables (including their interactions). Thus, we non-parametrically model $y_{ijd} = f(z)$ with $z = [\theta^l_{jd}, \theta^h_{jd}, U_d, X_d]$ by approximating it via Kernel Regularized Least Squares (Hainmueller and Hazlett 2014), $y = Kc$. Here, K is an $N \times N$ Gaussian Kernel matrix

$$K = \exp\left(\frac{-\|Z_d - z_j\|^2}{\sigma^2}\right) \tag{G.1}$$

with an associated vector of weights c. Intuitively, one can think of KRLS as a local regression method, which predicts the outcome at each covariate point by calculating an optimally weighted sum of locally fitted functions. The KRLS algorithm uses Gaussian kernels centered around an observation. The weights c are chosen to produce the best fit to the data. Since a possibly large number of c values provide (approximately) optimal weights it makes sense to prefer values of c that produce "smoother" function surfaces. This is achieved via regularization by adding a squared L2 penalty to the least squares criterion:

$$c^* = \underset{c \in \mathscr{R}^D}{\operatorname{argmin}} \left[(y - Kc)'(y - Kc) + \lambda c' K c \right], \tag{G.2}$$

which yields an estimator for c as $c^* = (K + \lambda I)^{-1} y$ (see Hainmueller and Hazlett 2014, appendix). This leaves two parameters to be set, σ^2 and λ. Following Hainmueller and Hazlett (2014), we set $\sigma^2 = D$ the number of columns in z and let λ be chosen by minimizing leave-one-out loss.

[8] For a very general discussion see Belloni et al. (2017).

The benefit of this approach is twofold. First, it allows for an approximation of highly nonlinear and non-additive functional forms (without having to construct non-linear terms as we do in the post-double selection LASSO). Second, it allows us to check if the marginal effects of group preferences changes with levels of union density *without* explicitly specifying this interaction term (and instead learning it from the data). To do the latter one can calculate pointwise partial derivatives of y with respect to a chosen covariate $z^{(d)}$ (Hainmueller and Hazlett 2014: 156). For any given observation j we calculate

$$\frac{\widehat{\partial y}}{\partial z_j^{U_d}} = \frac{-2}{\sigma^2} \sum_i c_i \exp\left(\frac{-\|Z_d - z_j\|^2}{\sigma^2} \right) \left(Z_d^{U_d} - z_j^{U_d} \right). \tag{G.3}$$

These yields as many partial derivatives as there are cases. We apply a thin plate smoother (with parameters chosen via cross-validation) to plot these against district-level union membership in Figure G.1. Perhaps unsurprisingly, we find that the assumption of an exactly linear interaction specification is too restrictive, especially in the case of the preferences of high income constituents.

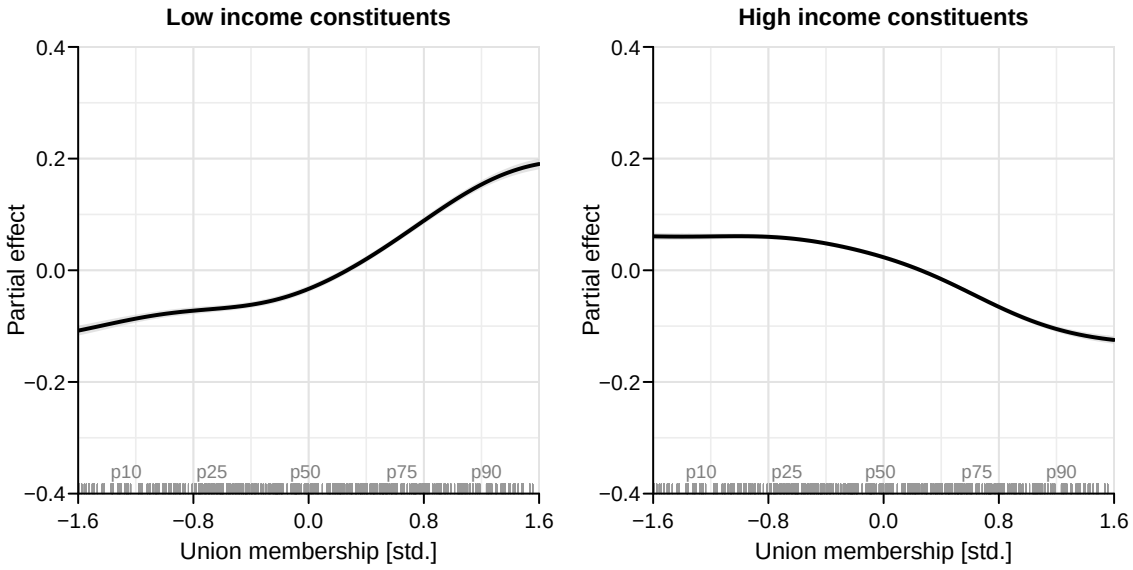

Figure G.1
Nonparametric estimate of interaction between union membership and preferences

Note: This figure plots partial effects (summarized using thin-plate spline smoothing) of preferences of low and high income constituents on legislative votes at levels of district union membership. Estimates obtained via KRLS.

However, the most noteworthy result clearly is the fact that, using a non-parametric model not including an *a priori* interaction between union membership and preferences, we find clear evidence that union membership moderates the relationship between preferences and legislative voting. For low income constituents, increasing district-level union membership steadily increases the marginal effect of their preferences on legislators' vote choice. Moving from

low levels of union membership (at the 25th percentile) to median levels of union membership increase low-income preference responsiveness by about 5 percentage points. An equally sized increase from the median to the 75th percentile increases responsiveness by almost 8 percentage points. We also find similar (albeit weaker) evidence for an interaction between high income group preferences and union membership.

H. Heterogeneity

Union type Is our finding driven by a particular type of union? A recent strand of research stresses the special characteristics of public unions and their political influence (e.g., Anzia and Moe 2016; Flavin and Hartney 2015). Hence, one may ask whether our findings mainly reflect the influence of private-sector unions since public sector unions are too narrow in their interests to mitigate unequal responsiveness. Panel (A) of Table H.1 provides some evidence on this question. The administrative forms used to measure union membership do not distinguish between private and public unions, and local unions may contain workers from both the private and the public sector. To calculate an approximate measure of district public union membership, we identify unions with public sector members (based on their name) and create separate union membership counts for "public" and the remaining "non-public" unions (see appendix A for details).

Our findings suggests that the coefficient for the impact of a districts' public union membership on the responsiveness of legislators to the preferences of the poor is sizable (at about 7 percentage points) and clearly statistically different from zero. At the same time, the coefficient for the remaining "non-public" unions is slightly reduced. The difference between the two estimates is not statistically distinguishable from zero. This finding does not support the hypothesis of a null-effect of public sector unions. It also suggests that the changing private-public union composition will not necessarily lead to less collective voice in Congress.

Bill ideology Panel (B) explores whether the effect of unions varies with the ideological direction of the bill that is voted on. Based on the partisan vote margin of the roll call vote, we define an indicator variable for conservative roll calls and estimate separate coefficients for each bill type. We find that union effects are relevant (and significant) for both bill types, they are larger for conservative votes. A standard deviation increase in union membership increases responsiveness to the preferences of low-income constituents by about 9 ($\pm$2) percentage points for conservative bills compared to about 5 ($\pm$1) points for liberal bills. The difference is larger for the preferences of high income constituents. In both cases the difference in marginal effects between liberal and conservative bills is statistically significant. Our findings suggest that union influence is more relevant for bills that have (potentially) adverse consequences for low income constituents. We trace this issue further in the next specification.

Union voting recommendations In panel (C) we consider bills with economic content and that have (or have not) been endorsed explicitly by the largest union confederation, the AFL-CIO. Our

Table H.1
Effect heterogeneity. Marginal effects of unionization on legislative responsiveness to low and high income groups.

	Low income	High income
(A) Private vs. Public unions		
Public unions	0.074 (0.016)	−0.058 (0.015)
Non-public unions	0.054 (0.016)	−0.027 (0.016)
(B) Bill ideology		
Conservative bill	0.086 (0.017)	−0.086 (0.018)
Liberal bill	0.052 (0.014)	−0.028 (0.013)
(C) AFL-CIO endorsement		
No position	0.054 (0.014)	−0.054 (0.013)
Endorsement	0.077 (0.015)	−0.040 (0.014)

Note: Estimates for η^l and η^l with cluster-robust standard errors in parentheses. N=15,780. Panel (A) shows separate effects for district counts of union members for unions classified as public or non-public (see text). Statistical tests for the difference in union type yield $p = 0.172$ for low income preferences and $p = 0.027$ for high income ones. Panel (B) estimates separate effects for bills classified as conservative or liberal based on their predominant party vote. Tests for significance of difference: $p = 0.009$ for low and $p = 0.000$ for high income preferences. Panel (C) classifies bills with economic content where the AFLCIO has taken a public stand for or against it (depending on bill content). Tests for significance of difference: $p = 0.003$ for low income, $p = 0.049$ for high income preferences.

definition of endorsement is based on voting recommendations made publicly by the AFL-CIO.[9] AFL-CIO recommendations signal the salience of the issue to unions, and they were made for more than half of the votes in the analysis. The mainly cover redistributive and economic issues. From the roll-call votes in our sample, the AFL-CIO took no position on stem cell research, Iraq redeployment, foreign intelligence surveillance, the repeal of Don't Ask Don't Tell, energy security, and several fiscal appropriations. Intuitively, we find that that the union effect is larger for issues where the AFL-CIO has made a clear endorsement. Panel (C) shows that the impact of union membership on legislators' responsiveness for bills especially relevant to low-income citizens is about 2 percentage points larger for votes on which the AFL-CIO had taken a prior position. This difference is statistically different from zero ($p = 0.003$).[10] The fact that districts with higher union membership see better representation of the less affluent more so when issues are salient to unions bolsters the interpretation that our main result is actually driven by unions' capacity for political action. This finding is also consistent with micro-level studies of the effects of union position-taking (Ahlquist, Clayton, and Levi 2014; Kim and Margalit 2017). There remains a smaller but significant union effect for the other issues as well. This

[9]Taken from the AFL-CIO "legislative scorecard", `https://aflcio.org/what-unions-do/social-economic-justice/advocacy/scorecard`.

[10]The high-income preferences estimate is smaller for endorsed bills but still significantly different from zero.

makes sense because union endorsements are not exhaustive, they may reflect some strategic considerations and policy issues are somewhat bundled.

References

Ahlquist, John S., Amanda B. Clayton, and Margaret Levi. 2014. "Provoking Preferences: Unionization, Trade Policy, and the ILWU Puzzle." *International Organization* 68(1): 33–75.

Anzia, Sarah F., and Terry M. Moe. 2016. "Do Politicians Use Policy to Make Politics? The Case of Public-Sector Labor Laws." *American Political Science Review* 110(4): 763–777.

Becher, Michael, Daniel Stegmueller, and Konstantin Kaeppner. 2018. "Local Union Organization and Law Making in the US Congress." *Journal of Politics* 80(2): 39–554.

Belloni, A, V Chernozhukov, and C Hansen. 2014. "Inference On Treatment Effects After Selection Amongst High-Dimensional Controls." *Review of Economic Studies* 81: 608–650.

Belloni, A, V Chernozhukov, I Fernández-Val, and C Hansen. 2017. "Program Evaluation and Causal Inference With High-Dimensional Data." *Econometrica* 85(1): 233–298.

Belloni, Alexandre, and Victor Chernozhukov. 2009. "Least squares after model selection in high-dimensional sparse models." *Bernoulli* 19(2): 521–547.

Belloni, Alexandre, Victor Chernozhukov, and Christian B. Hansen. 2013. "Inference for High-Dimensional Sparse Econometric Models." In *Advances in Economics and Econometrics: Tenth World Congress*, eds. Daron Acemoglu, Manuel Arellano, and Eddie Dekel. Vol. 3 Cambridge: Cambridge University Press , 245–295.

Belloni, Alexandre, Victor Chernozhukov, and Lie Wang. 2011. "Square-root lasso: pivotal recovery of sparse signals via conic programming." *Biometrika* 98(4): 791–806.

Belsley, David A, Edwin Kuh, and Roy E Welsch. 1980. *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: John Wiley & Sons.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45(1): 5–32.

Budd, John W. 2018. *Labor Relations: Striking a Balance*. 5 ed. New York, NY: McGraw-Hill Education.

Chernozhukov, Victor, Christian Hansen, and Martin Spindler. 2015. "Valid post-selection and post-regularization inference: An elementary, general approach." *Annual Review of Economics* 7(1): 649–688.

Chung, Yeojin, Sophia Rabe-Hesketh, Vincent Dorie, Andrew Gelman, and Jingchen Liu. 2013. "A nondegenerate penalized likelihood estimator for variance parameters in multilevel models." *Psychometrika* 78(4): 685–709.

Flavin, Patrick, and Michael T. Hartney. 2015. "When Government Subsidizes Its Own: Collective Bargaining Laws as Agents of Political Mobilization." *American Journal of Political Science* 59(4): 896–911.

Gelman, Andrew, and Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel / Hierarchical Models*. Cambridge University Press.

Gelman, Andrew, and Thomas C Little. 1997. "Poststratification into many categories using hierarchical logistic regression." *Survey Methodologist* 23: 127–135.

Gelman, Andrew, Hal S Stern, John B Carlin, David B Dunson, Aki Vehtari, and Donald B Rubin. 2013. *Bayesian data analysis*. Third ed. Boca Raton: CRC Press.

Hainmueller, Jens, and Chad Hazlett. 2014. "Kernel Regularized Least Squares: Reducing Misspecification Bias with a Flexible and Interpretable Machine Learning Approach." *Political Analysis* 22(2): 143–168.

Hill, Seth J., and Gregory A. Huber. 2019. "On the Meaning of Survey Reports of Roll-Call "Votes"." *American Journal of Political Science* 63(3): 611–625.

Hill, Seth J., James Lo, Lynn Vavreck, and John Zaller. 2007. "The Opt-in Internet Panel: Survey Mode, Sampling Methodology and the Implications for Political Research." Unpublished paper, Working paper.

Honaker, James, and Eric Plutzer. 2016. "Small Area Estimation with Multiple Overimputation." Unpublished paper, Manuscript. [http://hona.kr/papers/files/smallAreaEstimation.pdf].

Horrace, William C, and Ronald L Oaxaca. 2006. "Results on the bias and inconsistency of ordinary least squares for the linear probability model." *Economics Letters* 90: 321–327.

Hout, M. 2004. "Getting the most out of the GSS income measures. GSS Methodological Report 101." : Forthcoming.

Kim, Sung Eun, and Yotam Margalit. 2017. "Informed Preferences? The Impact of Unions on Workers' Policy Views." *American Journal of Political Science* 61: 728–743.

Kopczuk, Wojciech, Emmanuel Saez, and Jae Song. 2010. "Earnings Inequality and Mobility in the United States: Evidence from Social Security Data since 1937." *Quarterly Journal of Economics* 125(1): 91–128.

Lax, Jeffrey R, and Justin H Phillips. 2009. "How should we estimate public opinion in the states?" *American Journal of Political Science* 53(1): 107–121.

Lax, Jeffrey R, and Justin H Phillips. 2013. "How should we estimate sub-national opinion using MRP? Preliminary findings and recommendations." Unpublished paper, Paper presented at the Annual Meeting of the Midwest Political Science Association, Chicago.

Leeb, Hannes, and Benedikt M Pötscher. 2008. "Can one estimate the unconditional distribution of post-model-selection estimators?" *Econometric Theory* 24(2): 338–376.

Park, David K, Andrew Gelman, and Joseph Bafumi. 2006. "State-level opinions from national surveys: Poststratification using multilevel logistic regression." In *Public opinion in state politics*, ed. Jeffrey E. Cohen. Stanford: Stanford University Press , 209–28.

Putnam, Robert. 2000. *Bowling Alone. The collapse and revival of american community*. New York: Simon and Schuster.

Ratkovic, Marc, and Dustin Tingley. 2017. "Sparse Estimation and Uncertainty with Application to Subgroup Analysis." *Political Analysis* 25(1): 1–40.

Richardson, Sylvia, and Walter R Gilks. 1993. "A Bayesian approach to measurement error problems in epidemiology using conditional independence models." *American Journal of Epidemiology* 138(6): 430–442.

Robinson, Peter M. 1988. "Root-N-consistent semiparametric regression." *Econometrica* 56(4): 931–954.

Rupasingha, Anil, and Stephan J Goetz. 2008. "US county-level social capital data, 1990-2005." Unpublished paper, Penn State University, University Park, PA.

Stekhoven, Daniel J, and Peter Bühlmann. 2011. "MissForest. Non-parametric missing value imputation for mixed-type data." *Bioinformatics* 28(1): 112–118.

Tang, Fei, and Hemant Ishwaran. 2017. "Random forest missing data algorithms." *Statistical Analysis and Data Mining: The ASA Data Science Journal* 10: 363–377.

Tibshirani, Robert. 1996. "Regression shrinkage and selection via the lasso." *Journal of the Royal Statistical Society B* 58(1): 267–288.

Torrieri, Nancy, ACSO, DSSD, and SEHSD Program Staff. 2014. "American Community Survey Design and Methodology." Unpublished paper, United States Census Bureau. [www.census.gov/programs-surveys/acs/methodology/design-and-methodology.html].

Vavreck, Lynn, and Douglas Rivers. 2008. "The 2006 cooperative congressional election study." *Journal of Elections, Public Opinion and Parties* 18(4): 355–366.