

# LOCAL UNION ORGANIZATION AND LAWMAKING IN THE U.S. CONGRESS\*

Michael Becher<sup>†</sup>  
Daniel Stegmueller<sup>‡</sup>  
Konstantin Käppner<sup>§</sup>

*Journal of Politics*, 80(2), 2018

## ABSTRACT

The political power of labor unions is a contentious issue in the social sciences. Departing from the dominant focus on membership size, we argue that unions' influence on national lawmaking is based to an important degree on their local organization. We delineate the novel hypothesis that the horizontal concentration of union members within electoral districts matters. To test it, we draw on administrative records and map the membership size and concentration of local unions to districts of the U.S. House of Representatives, 2003-2012. We find that, controlling for membership size, representatives from districts with less concentrated unions have more liberal voting records than their peers. This concentration effect survives numerous district controls and relaxing OLS assumptions. While surprising for several theoretical perspectives, it is consistent with theories based on social incentives. These results have implications for our broader understanding of political representation and the role of groups in democratic politics.

---

\*We are grateful to Duke University and the University of Konstanz (YSF) for financial support. MB also acknowledges financial support from the Agence Nationale de la Recherche (ANR)-Labex IAST. We thank our editor and five reviewers, as well as Nicole Rae Berg, Thomas Gschwend, Alexander Hertel-Fernandez, Paul Marx, Jonas Pontusson, Jake Rosenfeld, Roland Zullo, and seminar participants at APSA, Konstanz, and Mannheim for helpful comments and discussions, as well as Martina Buchinger, Spencer Dorsey, and Marco Morucci for excellent research assistance. None of them should be held responsible for the results.

<sup>†</sup>Institute for Advanced Study in Toulouse, University of Toulouse Capitole, michael.becher@iast.fr

<sup>‡</sup>Duke University, daniel.stegmueller@duke.edu

<sup>§</sup>University of Konstanz, konstantin.kaepner@uni-konstanz.de

## I. INTRODUCTION

The power of labor unions to influence elections and lawmaking is a central concern in the study of democratic representation. Research in political science, economics, and sociology has commonly conceived of union power in the political arena as a question of aggregate group size, assuming that more members bring more votes and, perhaps, money. For instance, scholars have examined the effect of unionization on legislative voting in the United States Congress (Box-Steffensmeier et al. 1997; Freeman and Medoff 1984; Kau and Rubin 1978; Seltzer 1995). Related research examines the macro-level relationship between unionization and outcomes such as turnout, economic policy, political equality, or poverty (Bartels 2008; Brady et al. 2013; Flavin 2016; Leighley and Nagler 2007; Radcliff and Davis 2000). This research views unions mainly through the lens of aggregate membership numbers, ignoring other organizational features. A recent review concludes that existing studies “commonly sum together membership over many different unions with the implicit assumption that organizational characteristics do not matter” (Southworth and Stepan-Norris 2009: 310).

In this paper, we argue that focusing solely on membership size provides an incomplete account of the organizational basis of union political power, and it significantly limits our broader understanding of political representation. Specifically, we develop and empirically assess the argument that the influence of unions on national lawmakers has significant roots in their local organization. We propose a novel hypothesis about the link between district-level union structure and lawmaking. It states that the horizontal distribution of union members across local units within a congressional district – a feature we call concentration – shapes legislative voting: representatives elected in districts where union members are relatively concentrated should be less supportive of union positions than representatives elected in districts where an equal number of union members is dispersed across several unions. Theoretically, the hypothesis that membership concentration reduces union political power is controversial. To motivate it, we draw on foundational theories of collective action that point to the importance of selective social incentives in groups for political action (Olson 1965) and seminal behavioral research on the political significance of social interactions in local unions (Berelson et al. 1954). Unions have a federal structure with numerous constituent units at the local (i.e., establishment) level. Local unions lie at the heart of organized labor in the U.S. (Freeman and Medoff 1984: 34; also see Olson 1965: 66-76). They form the organizational base of the union pyramid, and this is where workers interact with each other on a regular basis, in the workplace and after work. At the local level, social incentives shape to what degree political resources of union members and their networks – votes as well as time and money – can be effectively mobilized to influence national lawmaking. Because social incentives are more effective when groups are not too large, this

logic suggests that unions' political influence in a district is higher when union members are distributed across many local units than when they are concentrated in few large units.<sup>1</sup>

To empirically evaluate the concentration hypothesis, we draw on extensive and geographically fine-grained administrative records from the Department of Labor. This largely neglected data source allows us to precisely map union membership size and membership concentration to electoral districts of the House of Representatives between 2003 and 2012. Strikingly, the data show that the concentration of union membership is orthogonal to the number of union members in a district. In line with previous research, we find that district-level union membership is significantly linked to legislative voting. However, we also find that district-level union concentration is an important determinant of legislative votes. After accounting for membership levels, legislators from districts with a relatively low union concentration have a substantively more liberal legislative ideology and a higher propensity to support the union position on individual key votes (e.g., health care) than those from high concentration districts. This concentration effect holds after accounting for state effects, period effects, flexible time trends, and numerous district-level characteristics. Importantly, our empirical strategy accounts for district-level economic concentration, such as the number of firms and employment concentration, as well as campaign contributions from business. This rules out the possibility that the apparent effect of union concentration is explained by economic concentration. The results are also robust when relaxing functional form modeling assumptions. In addition, we provide some historical evidence to further clarify that economic structure alone does not explain concentration. Exploring channels of influence, we find that concentration is linked to campaign contributions and the election of Democratic representatives.

Taken together, our analysis demonstrates that the concentration dimension of local union organization substantively shapes lawmaking in Congress. These findings matter for our understanding of the relevance of unions for democratic representation as well as theories of groups in democratic politics more broadly. While union membership in the US and several other countries has receded far below its post-war apex (Rosenfeld 2014), union members still constitute one of the largest organized groups in the political arena. Our results suggest that their local organization affects the making of national laws on a broad range of policies, with implications for large segments of the population and political inequality. Theoretically, the concentration effect we find is difficult to reconcile with the view, expressed for instance in the seminal work of Key (1964), that organizational fragmentation necessarily undermines the

---

<sup>1</sup>Taking a different perspective, comparative political economists have studied the centralization of union organizations for wage bargaining as well as broader corporatist arrangements (e.g., Iversen 1999; Pontusson et al. 2002). Our approach is complementary to this important line of research, which does not examine the political effects of district-level union organization.

ability of groups to overcome collective action problems in politics. Instead, it is consistent with theories of political action in groups that account for social incentives. Our focus on local organization complements recent studies emphasizing that the political influence of unions is conditioned by the institutional environment, such as electoral rules and labor laws, or leadership (Ahlquist and Levy 2013; Anzia and Moe 2015; Anzia 2011; Flavin and Hartney 2015; Kim and Margalit 2016). It stands to reason that membership concentration may also matter for the political significance of other groups.

Another contribution of our paper lies in its empirical strategy. It departs from the heavy reliance on survey data in quantitative research on unions. This enables us to overcome two important problems (Southworth and Stepan-Norris 2009). First, a fundamental drawback of mass surveys is that they typically do not provide detailed information on the local union to which a member belongs. This precludes the possibility of examining the concentration of union membership. The second limitation concerns the counting of union members in a particular locality. State-level estimates of union density from the Current Population Survey (Hirsch et al. 2001) are used frequently across the social sciences. However, the number of survey respondents is too small to provide membership numbers for electoral districts. Hence measurement is bound to be quite noisy, and may be upward biased (Southworth and Stepan-Norris 2009). Many previous studies therefore rely on state-level measures only; others limit their coverage to metropolitan statistical areas (Box-Steffensmeier et al. 1997). In this paper, we use mandatory reports (so-called LM forms) filed by local unions to the Department of Labor. Their submission is a legal requirement for most unions, non-submission and incorrect submissions are penalized, and the Department of Labor conducts regular audits. We have retrieved all available raw data for around 30,000 individual unions (from 2000 to 2012) from the Department of Labor, and processed them such that we are able to construct annual measures of union membership and membership concentration by congressional district. The resulting measures provides a new empirical perspective on the structure of organized labor in the twenty-first century; they may also be used to address a variety of questions not considered in this paper.<sup>2</sup>

---

<sup>2</sup>Scholars have noticed the large potential of the Department of Labor's LM forms for social science research, though it appears that "the cost in synthesizing a large sample has thus far deterred systematic analysis" (Southworth and Stepan-Norris 2009: 312). There have been studies using smaller subsets of these records (Martin 2008; Zullo 2008). We provide the first comprehensive analysis for a multi-year period covering the whole country.

## II. LOCAL ORGANIZATION AND POLITICAL INFLUENCE

We proceed to lay out the theoretical perspectives that motivate our subsequent empirical sections. Going beyond the common conception of membership size as the essential power resource, but echoing several strands of research, we advance the argument that organized labor's effect on national legislators is strongly tied to the organization of local unions in an electoral district.

Following Olson (1965: 136) and many others, labor unions are viewed as groups that have been organized to serve the economic interests of their members relative to their employers, through collective bargaining over wages and benefits. Unions may turn to political action to pursue policy preferences and ideas of their members and/or leaders, though this does not happen by default. We contend that local organization shapes to what degree political resources of union members and their social networks can be effectively mobilized to shape policy.

### *II.A. The Concentration Hypothesis*

In a polarized two-party system, one party will be closer to the political leanings of most union members. Since the New Deal and especially early post-war years, that role has been played by the Democratic Party (Dark 1999; Lichtenstein 2013; Schlozman 2015). Democratic representatives in Congress are predictably more supportive of policies favored by unions than their Republican counterparts (Box-Steffensmeier et al. 1997; Seltzer 1995).<sup>3</sup> A majority of union members regularly reports favoring Democratic over Republican candidates (Rosenfeld 2014: 176). And while political economy models highlight that unions' narrow economic interests may diverge due to their specific occupational or sectoral structure, issue bundling inherent in two-party competition nonetheless produces a fairly stable alignment between Democrats and labor unions. Thus, the main problem of organized labor in the political arena is to get Democratic lawmakers elected to protect or expand liberal economic policies in Congress. This requires votes, money, and other resources to win office.<sup>4</sup>

We argue that local union organization matters for achieving this goal far more than has previously been recognized, and we focus on two important dimensions: union membership and the hitherto ignored degree of union *concentration*. By concentration we refer to the degree to which union members in a congressional district are concentrated in few local unions versus

---

<sup>3</sup>Excluding Southern Democrats.

<sup>4</sup>More generally, rational theories of electoral mobilization assume that turnout matters for policy mostly or exclusively by influencing which ideological type of politician wins office (e.g., Abrams et al. 2011). Research on legislative voting in Congress also suggests that voters affect lawmaking more by influencing who wins the election than by changing the position of elected politicians (Bartels 2008; Lee et al. 2004).

being distributed more evenly across several locals. It is usually argued that a higher number of union members entails more political influence, since it implies a larger pool of votes and other resources (Box-Steffensmeier et al. 1997; Masters and Delaney 2005: 369; Olson 1965: 68). We do not dispute this hypothesis. Everything else equal, increasing union membership brings more political clout. However, our point is that the usual focus on membership counts or union density leaves out a politically significant feature of union organization.<sup>5</sup>

The significance of local unions has already been discussed in a less widely cited part of Olson's (1965) seminal analysis of collective action. He points out that the federal structure of organized labor, with thousands of local unions as the basic organizational unit, may be conducive to collective action in a group that is large in the aggregate. Even in the absence of external enforcement, social interactions can sustain individual contributions toward the collective good — as long as local unions are not too large (Olson 1965: 66-97). This setting fosters social interactions between local union members and entails selective incentives to engage in costly activity on behalf of the group.<sup>6</sup>

Going beyond Olson (1965), we take this logic to apply to political behavior on behalf of the group (i.e., union) more broadly. This includes voting as well as contributing money or time to a political campaign. It may also include efforts to foster norms of solidarity or to shape policy preferences in line with group ideology (Ahlquist et al. 2014; Kim and Margalit 2016). The basic mechanism is that the prospect of approval, respect, or companionship being awarded or withdrawn alters the individual calculus of political action. For example, Abrams et al. (2011) present a formal model that elucidates how networks of family, friends, or co-workers shape participation in large elections (the same logic applies to political contributions). People vote despite a negligible probability of being pivotal if their social network attaches enough importance to voting for a particular party or candidate. Voting consistent with the group norm is rewarded while deviations are punished, and in equilibrium voting and punishment are self-enforcing. Social pressure remains a potent motive for political behavior in the internet era (Druckman and Green 2013). It does not require explicit action by group leaders, even though they may try to harness it.

Of course, empirical research started to document the relevance of social interactions in unions for politics well before the publication of Olson's (1965) theory. In their study of the 1948

---

<sup>5</sup>This is a *ceteris paribus* argument. We are not arguing that unions' political influence is exclusively due to local interactions. National unions and the competency and ideology of their leaders are important (e.g., see Ahlquist and Levy 2013; Dark 1999).

<sup>6</sup>While Olson (1965) recognizes the potential importance of social incentives for unions, he presumed that most local unions had become too large for this mechanism to be effective. However, our descriptive data shows otherwise for 2003-2012. While there are some very large local unions, the typical local union is small: it has slightly more than 100 members (see appendix A.2, Figure A.2.1; see also Freeman and Medoff 1984: 34-35).

presidential election, Berelson et al. conclude that social interactions within the same unionized plant were driving higher support for the Democratic party in a context where mobilization by local union leaders was low (Berelson et al. 1954: 37-53). Since then, a large body of research has confirmed the relevance of social incentives for political behavior more broadly, increasingly drawing on field experiments, though usually not considering unions (e.g., Druckman and Green 2013; Gerber et al. 2008).

Drawing on the preceding arguments, we propose that both the total number of union members and the horizontal concentration of members across different local unions within the same electoral district matter for democratic representation. It is uncontroversial to claim that a higher *number* of union members in a district should be linked to more consistent legislative support for pro-worker policies by elected lawmakers. Our novel hypothesis is that the *concentration* of members also matters: everything else equal, lawmakers from a district with highly concentrated union membership should be less supportive of pro-union legislation than lawmakers from districts where the same membership is distributed across more unions. The concentration hypothesis follows from the diminished efficacy of social incentives when many members are concentrated in one or few large unions.

Two theoretical qualifications are in order. First, concentration will not matter politically where union membership is so small that not even a fully mobilized membership is politically relevant. This threshold may vary depending on district characteristics (such as competitiveness). In the empirical section, we explore this issue in an interactive model. Second, the social interaction logic implies that high membership concentration undermines social incentives. It does not necessarily suggest that more membership fragmentation is always better. As unions become very small, approaching the (hypothetical) extreme of a one-member “group,” social pressure will not exist at all.<sup>7</sup> Our empirical concentration measure will reflect this consideration.

## *II.B. Countervailing Forces*

Other approaches in the literature point to countervailing mechanisms that push in the opposition direction of the concentration hypothesis. One view is that collective action problems between the leaders of different local unions undermine the electoral mobilization of union workers (e.g., see Key 1964: 66). With numerous unions in the same electoral district, union leaders deciding how much effort to put into political action are more likely to face a collective action problem in which they have incentives to free-ride on the political efforts of others. As a result, the dispersion of union members across multiple unions in a district should reduce, rather

---

<sup>7</sup>This point is highlighted in the formal model of Abrams et al. (2011: 244).

than increase, the political influence of organized labor. Thus, one may plausibly expect that, controlling for union membership, higher union concentration leads to more influence through the selection of more union-friendly legislators.<sup>8</sup> This is a clear-cut argument, consistent also with the standard model of Olson (1965). That said, there are theoretical reasons to suspect that the importance of collective action problems among local leaders is diminished by the same mechanisms that underlie the concentration hypothesis. Social incentives can encourage local leaders to do their part despite the temptation to free-ride. Less concentrated unions increase members' ability to hold their leaders to account, thus inducing them to political action via social pressure.

Another argument is that higher union concentration should strengthen the political power of unions if concentration goes hand in hand with more homogeneous policy preferences. Unions' narrow policy preferences may be based on occupation, industry, or firm interests. Higher union concentration may then imply that policy preferences are more homogeneous, making it easier perhaps to overcome collective action problems and increasing policymakers' incentives to be responsive (Busch and Reinhardt 2000). Prior research suggests that some heterogeneity in the narrow material interests of unions does not necessarily prevent joint political action. One point is that the dynamics of two-party competition leads to policy bundling along a single-dimension of politics despite a potentially large multi-dimensional policy space; it usually means that different unions side with the same political party on a broad range of issues (Rosenfeld 2014; Schlozman 2015). Unions have also constructed broader "communities of fate" that cut across narrow economic interests and sometimes take costly actions on behalf of other or broader groups (Ahlquist and Levy 2013; Lichtenstein 2013). Moreover, unions interested in their long-term survival have incentives to organize across industries (Kremer and Olken 2009), which implies that membership concentration need not reflect preference homogeneity.

This discussion highlights the fact that the effect of membership concentration on lawmakers is theoretically ambiguous. Only if the force of social mechanisms is sufficiently strong should we see that lower concentration leads to more legislative support of policies favored by unions and their members. In that sense, the concentration hypothesis is not "self-evident". It merits careful empirical testing.<sup>9</sup>

---

<sup>8</sup>A related argument is that organizational fragmentation may reduce labor's political clout through inefficient allocation of political resources (Lichtenstein 2013: 143-144).

<sup>9</sup>As discussed in an earlier note, a strong literature in comparative political economy studies the economy-wide centralization or coordination of wage bargaining (e.g., see Iversen 1999; Pontusson et al. 2002). Its focus on the vertical distribution of bargaining authority within union confederations is orthogonal to the focus of this paper on the horizontal concentration of union members at the level of electoral districts.



### III. MAPPING ORGANIZED LABOR

We map our two central features of local union organizations—their size and concentration—to electoral districts for the U.S. House of Representatives during the 109-112th Congress. We use administrative data covering almost 30,000 local unions for more than a decade, based on more than 300,000 individual reports. Our data set provides a new empirical perspective on the structure of organized labor in the US, and it allows for a re-examination of the political influence of labor unions on lawmaking. In this section, we describe our measurement strategy and data.<sup>10</sup>

#### *III.A. Using LM forms*

We analyze and aggregate mandatory reports filed annually by local unions with the U.S. Department of Labor. As this is not a widely used data source in political science, some exposition of its advantages and drawbacks will be helpful. The legal basis for these reports is the Labor-Management Reporting and Disclosure Act (LMRDA) of 1959, which started as a movement for greater union democracy (initiated by the ACLU) but was transformed by legislators into a push to curtail the economic and organizational power of unions (Aaron 1960).<sup>11</sup> The act introduced a comprehensive system of reporting: unions have to file an initial report with the Office of Labor-Management Standards (OLMS) followed by a yearly report using a so-called LM form. While the level of reporting detail varies by union income, all yearly forms include information on the number of union members and the address of the union office. For the public sector, the Civil Service Reform Act (CSRA) of 1978 affirmed union rights in the public sector created by previous presidential executive orders (Coleman 1980: 202). It also created a system of reporting similar to the one in place for the private sector (also overseen by OLMS).

Their legal basis makes LM forms a reliable source of information on unions and their members. Filing LM forms is mandated by law and failure to report, or reporting falsified information, is made a criminal offense in the LMRDA punishable by a fine of up to \$100,000 and/or imprisonment of up to 1 year.<sup>12</sup> As a further incentive against misreporting, the validity of union reports is continuously verified. Under its Compliance Audit Program, OLMS selects a

---

<sup>10</sup>Our focus is on the House of Representatives rather than the Senate because this allows us to exploit within-state variation in union organization. This is especially useful because the organizational features move slowly during our period of observation.

<sup>11</sup>While the ACLU has pushed for greater internal union democracy since 1947, LMRDA legislation occurred in the wake of the ‘Select Committee on Improper Activities in the Labor Management Field’ hearings chaired by Senator McClellan from 1957–59. For a summary of LMRDA’s legislative history, see Aaron (1960).

<sup>12</sup>The corresponding figure for unions covered by CSRA is \$250,000 and 5 years imprisonment.

number of unions for detailed audit each year.<sup>13</sup>

Using LM forms has important advantages over using measures derived from surveys. Multi-purpose surveys usually only ask about union membership without any reference to the particular union in question. This precludes studying the organizational basis of unions and probably explains the focus on membership size in the literature on the political power of unions. Even with respect to union membership there are well-known measurement problems. For example, the Current Population Survey (the most widely-used source for union membership) uses a rather broad question wording, querying respondents if the person in question is a “member of a labor union or of an employee association similar to a union”. This invites over-reporting leading to an overestimate of union membership (Southworth and Stepan-Norris 2009: 311). In addition, some systematic misclassification by respondents has been documented (Card 1996). A related issue is unit non-response, which plagues every survey.<sup>14</sup>

There are, of course, disadvantages to using administrative data. Perhaps the main potential drawback when using LM forms is that some unions are exempt from filing requirements. While each and every private sector union has an associated LM form, some public sector unions are not covered by the relevant laws and are not required to file regular reports. All unions representing postal or federal employees are covered. But unions that exclusively represent state, county, or municipal government employees are exempt. However, note the strong definition of exclusivity here. Unions almost exclusively representing, say, municipal government employees nonetheless do have to file if only one of their members is employed in the private sector or by the federal government. For instance, a police union that also includes private security employees will have to file a report. The same holds for a union of municipality workers that includes postal staff. Since the latter part of the twentieth century unions are no longer purely organized along craft and industry lines, but enroll workers of different sectors and occupations (Lichtenstein 2013: 249):

[...] the UAW [United Automobile Workers] has recruited health-insurance clerks and prison guards, and its largest local west of the Mississippi represents teaching assistants, tutors, and readers at the University of California’s ten campuses. Meanwhile, the United Steelworkers have organized Pittsburgh grocery workers, and the Communication Workers of America negotiate for state and municipal employees.

---

<sup>13</sup>Unions were chosen based on random selection, OLMS’s internal discretionary criteria, or based on complaints of union members.

<sup>14</sup>While the CPS has very low non-response rates, they (in line with the general trend) have risen from about 7% in 1995 to 11% in 2014 (Robison and Grievies 2014: 1339), increasing the need to rely on weighting procedures to produce national or state estimates of union membership.

Similarly, the American Federation of State, County, and Municipality Employees also organizes private sector workers in education, health and home care. This behavior is consistent with evolutionary models of unions (Kremer and Olken 2009). Altogether, this suggests that the degree of non-coverage in our data is probably limited. Validating our data against governmental aggregated sources, we indeed find a very high degree of coverage (see below and Appendix A.1).

### *III.B. Measuring Membership and Concentration*

From the Department of Labor we obtain a database of digitized LM forms. The process of turning these raw data into measures of union organization for each House district consists of three steps: cleaning the data from entry errors, spelling mistakes etc.; geo-coding; and district-level aggregation using the relevant quantities of interest.<sup>15</sup>

Using cleaned addresses, we produce a geo-location for each union address by processing each address via natural language processing and then matching it geographically.<sup>16</sup> Since not all addresses are necessarily complete, we use exact matching with attribute relaxation. This means that while exact matches are preferred, they can be produced at decreasing levels of precision: (1) The most precise geo-location is produced by an exact match to a segment of a street based on street and house or building numbers; (2) at the second level, matches are based on the ZIP portion of an address; (3) if this fails, matches are based on the city or on county-regions; (4) if a whole address cannot be resolved, matches are based on the state portion of an address. However, the latter three strategies are rarely needed. More than 99% of the 358,051 union forms filed between 2000 and 2013 are matched based on either an exact address or a ZIP code. Given a union's unique geo-location, it is placed into the map of districts for the House of Representatives for each Congress using the cartography shapefiles for congressional districts from the Census Bureau.

With the geo-located union data in hand, we calculate two main measures of local union organization. First, the total number of union members in the district is the sum of members in all local unions. Given census-based congressional apportionment, population size is roughly constant across districts (and often identical within states), except in the seven at-large districts.

---

<sup>15</sup>We conducted checks of the DoL's digitization of LM forms and find a high level of accuracy. For a randomly selected sample of 2,291 LM forms, we verified the coding of their content against the original documents. In 96.8% of all cases, digitized information on address and number of union members exactly match the original form submitted by the union officer. Most issues with the remaining 3.2% of deviating forms were minor and concerned mostly address details, such as small typographical errors. A number of those deviations result from mistaken entries by union officers, for example, by entering street addresses into the field intended for PO boxes (and vice versa). We address these issues first through data-cleaning of all 358,051 LM form address fields, and second through a flexible natural language processing step preceding our geocoding procedure.

<sup>16</sup>We use Texas A&M University's geocoding services; see Goldberg et al. (2007) for details.

Hence, it makes little difference whether we simply look at membership counts or normalize them by population. As discussed above, membership size is the dominant measure of union political power in the social science literature. Compared to surveys, our LM-based measure does not suffer from small sample problems and is geographically more accurate than those used in previous research (see, for instance Box-Steffensmeier et al. 1997).

Second, following our theoretical discussion we calculate the degree of concentration vs. dispersion of union members in an electoral district. Our main measure of union concentration is the 4-union concentration ratio (CR4). It captures the share of all union members in a district that belong to the largest four unions. It is thus bounded between 0 and 1. In the case of perfect concentration, all members are concentrated in the top four unions. As membership becomes more dispersed, the ratio declines. Empirically, at the minimum the largest four unions capture only 14 percent of all members. The concentration ratio is intuitive and straightforward to interpret. Another desirable property is that it is not mechanically related to the overall union membership in a district. Group size and concentration may still be correlated empirically, but this will not be by construction. While the four-largest-unions threshold is somewhat arbitrary, it captures a large amount of variation in the data; it also follows a long tradition in the empirical analysis of firm concentration (Curry and George 1983), and it seems natural to apply the same criteria to unions. Alternatively, we can calculate the effective number of unions (analogous to the well-known effective number of parties measure), which turns out to be highly correlated with the concentration ratio ( $r > 0.9$ ). All results reported below also obtain when using it.<sup>17</sup> Both measures capture the theoretical intuition that social pressure works best when membership is not too concentrated. This logic does not suggest that more fragmentation in the tail is always beneficial. In Appendix A.4, we provide evidence that union concentration is rooted in historical organizational preferences.

Before describing the structure of local union organization, we report how we validated our data. While there is no “gold standard” of accurate union membership numbers, we compare our data to the widely used CPS-based measure of state-level union density. The two measures agree to a large extent (their correlation, averaged over all years, is 0.86). On average unionization levels based on estimates from the CPS are 1.9 percentage points higher than counts of members from LM forms. This difference is consistent with some degree of over-reporting, induced by CPS’s broad question wording (Southworth and Stepan-Norris 2009: 311). It can also be interpreted as an upper bound for the non-coverage of some public sector unions in our data, confirming that LM forms provide a rather comprehensive accounting of unions. For a more

---

<sup>17</sup> Appendix A.7 reports results for the effective number of unions; it also constructs a one-dimensional measure of union organization based on the ratio of union concentration to union membership suggested by a reviewer.

detailed discussion, see appendix A.1.

### *III.C. Two Dimensions of Union Organization*

A cross-sectional snapshot of union membership numbers (as population share) and membership concentration (CR4) across the map of congressional districts for the 109-112th Congress is shown in Figure I. Panel (a) highlights that there is considerable variation in union membership even within states. For instance, states with relatively high average union membership, such as California or New York, show substantial geographic variation of union membership; even states with low average union membership, such as Florida or Texas, show pockets of high union density. Panel (b) shows that there is also considerable variation in union concentration among districts.

We have argued theoretically that membership numbers and concentration are separate dimensions of union organization. A district with a large number of union members might see these concentrated in a few unions, or dispersed among many. Empirically, these two characteristics turn out to vary independently as well. In Panel (a) of Figure II, we plot membership concentration against the logged number of union members for each of the 435 districts. The absence of any clear correlation ( $r = -0.05$ ) is apparent from the widely scattered observations. This striking pattern suggests that we can, for the sake of exposition, distinguish between four combinations of membership size and concentration (high-high, high-low, low-high, low-low). Panel (b) plots these combinations in the map of House districts. While this simplified classification understates variability, one can see that there is variation within states that is not captured by membership counts. For instance, while California's 4th and 12th congressional districts both contain an almost identical number of union members (about 40,000 during 110th Congress), they vary dramatically in their level of concentration. In the 4th district, the four largest unions capture 92% of all members while capturing only about 50% in the 12th. In a similar vein, Texas's 7th and 11th congressional districts feature an equally low number of union members (about 2,500), but members are dispersed in the former (with a concentration ratio of 55%), and concentrated in the latter (97%).

### *III.D. What Explains Concentration?*

Figures I and II raise the question of what explains variation in union organization. While there is a well-developed literature on the determinants of union membership (for a review, see Wallerstein and Western 2000), we are aware of no prior study that explains union concentration. In Appendix A.4, we explore this question in more detail. To summarize, we suggest that district-

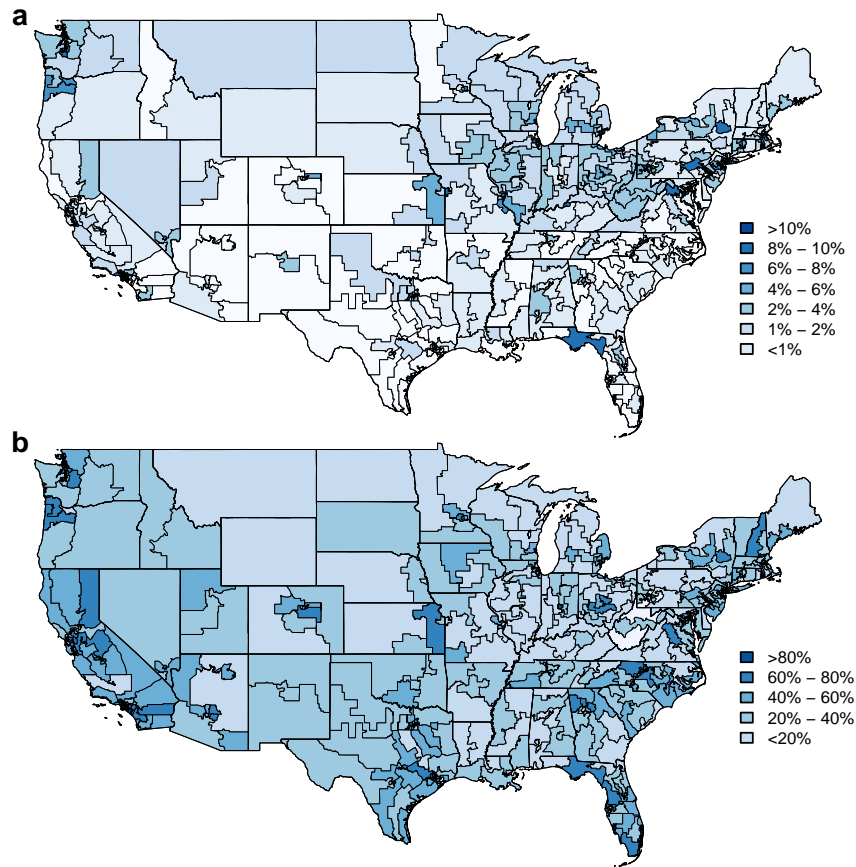


FIGURE I  
Union membership and concentration in the United States.

The map shows average union membership and union concentration for each district of the 109th to 112th Congress. Entries are averages over time. Panel (A) shows union members as percentage of the total population, panel (B) shows union concentration (measured as 4-union concentration ratio).

level union concentration is partly rooted in history, shaped by the interaction of economic structure and historical preferences for how unions should be organized. While the distribution of economic establishments puts a constraint on union concentration, their relatively large number (see Appendix A.2) leaves a considerable amount of slack to be shaped by organizational preferences that, once established, are sticky. Using survey data from the New Deal era, a turning point in the history of organized labor (Lichtenstein 2013), we find that public preferences on *how* unions should be organized – narrowly along craft-lines or more broadly – are clearly distinct from support for union rights, that is, whether unions should exist at all (Appendix A.4., Fig. A.4.1). Furthermore, in a regression analysis at the level of congressional districts we find that the contemporaneous link between economic concentration and union concentration is

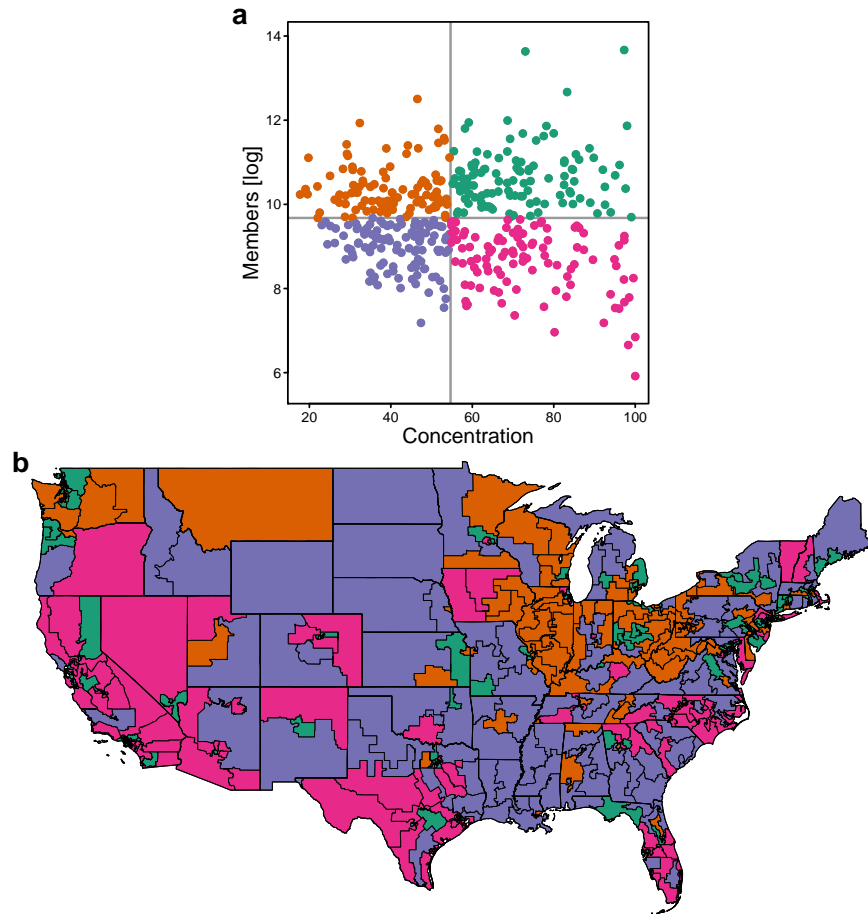


FIGURE II  
Relationship between union membership and concentration.

This figure illustrates the independent variation of membership and concentration during the 109th to 112th Congress. Panel (a) plots CR4 concentration against (logged) membership in each district. Different colors mark the four quadrants obtained by splitting both measures at the median. Panel (b) maps the distribution of the four membership–concentration combinations. All entries are averages for 2005–2012.

significantly moderated by preferences concerning union structure from the 1930s (Table A.4.1). These findings are broadly consistent with an organizational perspective of union behavior (Ahlquist and Levy 2013), which suggests that members of successful unions embrace the larger organizational goals or strategies set by (founding) leaders. They clarify that significant variation in union concentration can be traced back to preferences that are not hardwired into the current economic structure.

## IV. LOCAL UNIONS AND VOTING IN CONGRESS

Does the organizational structure of unions in a congressional district influence how the district’s representative votes in Congress? Our dataset allows us to directly assess our hypothesis that the political clout of organized labor is not only shaped by the number of union members (as found by previous studies), but also by the distribution of members across local unions.

### IV.A. Empirical Strategy

We estimate a series of models for legislators’ voting behavior as a function of local union characteristics and suitable controls. Our main model is:

$$y_{dst} = \mathbf{x}'_{dt}\boldsymbol{\beta} + \mathbf{u}'_{dt}\boldsymbol{\gamma} + \theta_s + \tau_t + \boldsymbol{\xi}_s\psi(t) + \varepsilon_{dst}.$$

Our dependent variable is  $y_{dst}$ , a summary index of the voting behavior of the House representative from district  $d$  in state  $s$  during congressional term  $t$  based on a large number of votes. We rely on the most popular measure, the first dimension of DW-NOMINATE scores calculated by Poole and Rosenthal (1997).<sup>18</sup> DW-NOMINATE scores measure legislators’ revealed ideology on the dominant left-right dimension, where  $-0.75$  is the most liberal and  $+1.36$  the most conservative position in our data, and are comparable over time. The unit of analysis is the individual legislator in a given term.<sup>19</sup> In addition, we report results for individual key votes where unions took an explicit position.

The vector  $\mathbf{u}_{dt}$  captures our two main variables characterizing local union organization as calculated from LM forms: the logged number of union members and the 4-union concentration ratio.<sup>20</sup> As we have seen, different theoretical perspectives suggest opposite predictions about the direction of the concentration effect; our argument implies that higher concentration leads to less responsive policymakers for a given level of unionization. To capture common time shocks, such as mid-term versus presidential electoral cycles or shifts in national mood, our model allows for Congress-specific effects  $\tau_t$ . All our specifications include a set of time-varying district-level controls,  $\mathbf{x}_{dt}$ , that follow the literature on roll call voting (e.g., McCarty et al. 2006):

<sup>18</sup>Note that an alternative measure, pro-union voting scores assembled by the AFL-CIO, produces comparable results (in terms of estimates and significance); see appendix A.6 for details.

<sup>19</sup>Including special elections to fill vacancies, the total number of possible DW-NOMINATE observations is 1782. The analysis covers 1767 because the dataset retrieved from Poole’s website ([www.voteview.com](http://www.voteview.com)) does not include scores for 15 legislators with too few individual votes to produce reliable estimates.

<sup>20</sup>We take the log as the distribution of members is right-skewed. In appendix A.6 we show that an alternative measure based on the share of union members yields the same substantive (in terms of sign and significance) results.



median family income, racial composition (percentage white), and level of education (percentage with BA degree or higher). We also control for the share of a district’s workforce employed in the service sector, as this sector has been more resistant to unionization (Freeman and Medoff 1984: ch. 13), the share of agricultural employment, and a district’s degree of urbanization. To further assess the concern that union concentration is driven by economic concentration, we will estimate specifications including additional controls for districts’ economic structure. Appendix A.2 provides descriptive statistics for all variables used in our analysis.

The impact of time-invariant state-level characteristics is captured by state fixed effects,  $\theta_s$ . Most notably, “right-to-work” laws in some states allow workers in unionized companies to opt out of union membership and paying associated dues, and thus shape the mobilization capacity of organized labor.<sup>21</sup> Collective bargaining rights for public sector workers also vary between states (Flavin and Hartney 2015). However, the inclusion of state fixed effects alone leaves open the possibility that omitted *time-varying* confounders affect our results. While an observational study is never able to rule out this possibility completely, we can account for linear and nonlinear time trends in state-level unobservables by including state-specific functions of time in  $\xi_s \psi(t)$ .<sup>22</sup> Finally,  $\varepsilon_{dst}$  are residuals. We employ standard errors robust to within-state heteroscedasticity and correlation (Cameron and Trivedi 2005: 834).<sup>23</sup>

Our analysis focuses on the 109-112th Congress (2005-2012). This reflects research design considerations and data constraints. Since we pool observations from different districts, we focus on a single apportionment period during which district borders remain constant (i.e., 108-112th Congress based on the 2000 census), with the exception of several cases of court-ordered redistricting in Georgia and Texas (we ensure that our results hold when these are removed).<sup>24</sup>

#### IV.B. Results

Table I presents results from six specifications. All include both state and congress fixed effects, as well as basic district characteristics. Column (1) includes our measures of union membership (the logged number of members) and concentration (the district 4-union concentration ratio).

---

<sup>21</sup>Their influence is captured by state-specific constants, since right-to-work laws are constant in our estimation sample (it ends before the only two reforms, which occurred 2012 in Indiana and Michigan, had time to materialize).

<sup>22</sup>We specify  $\psi(t)$  as restricted cubic splines with three knots.

<sup>23</sup>The short times series (T=4) and the slow-moving nature of union organization imply that a within-district analysis is not informative, as district fixed effects cannot be distinguished from the variables of interest.

<sup>24</sup>We exclude the 108th Congress because our district-level controls from the American Community Survey are not available. But note that including the 108th Congress when backwards-interpolating missing control variables via state-specific linear and quadratic time trends does not change our results (see appendix A.6).

These results confirm previous findings linking union membership to more liberal congressional voting outcomes (e.g., Box-Steffensmeier et al. 1997; Freeman and Medoff 1984; Kau and Rubin 1978; Seltzer 1995). Turning to the concentration hypothesis, our results show that concentration matters for legislative outcomes in the direction suggested by arguments about the importance of social incentives for group political action: Holding overall membership constant, a unit increase in local union concentration is related to a 0.37 unit increase in conservative legislative ideology of that district’s representative. In specification (2) we add the interaction of both union variables (each centered for easier interpretation) and do not find model-based evidence for a conditional relationship between membership and concentration. Given the empirical pattern evident in panel (a) of Figure II, this is not surprising.

TABLE I  
Union membership, concentration, and legislative voting.

	(1)	(2)	(3)	(4)	(5) <sup>d</sup>
Union members	−0.075 (0.022)	−0.084 (0.019)	−0.073 (0.022)	−0.072 (0.023)	−0.073 (0.020)
Concentration	0.368 (0.116)	0.374 (0.122)	0.369 (0.113)	0.376 (0.118)	0.311 (0.102)
Members×concentration		0.044 (0.098)			
State fixed effects	✓	✓	✓	✓	✓
Congress fixed effects	✓	✓	✓	✓	✓
District characteristics <sup>a</sup>	✓	✓	✓	✓	✓
District economic structure <sup>b</sup>			✓	✓	✓
State time trends <sup>c</sup>				✓	✓
N	1767	1767	1767	1767	1767
R <sup>2</sup>	0.512	0.496	0.514	0.533	0.543

*Note:* 109th to 112th Congress. Standard errors robust to arbitrary within-state correlation and heteroscedasticity. Analytical standard errors are possibly sensitive to small or unevenly sized clusters. We show in appendix A.6 that our results hold when employing a cluster bootstrap or a wild cluster bootstrap (Cameron et al. 2008).

<sup>a</sup> Controls for income (median household income), race (percent white), education (percent with college or post-college degree), employment (percent in service industries, and percent in agriculture), degree of urbanization.

<sup>b</sup> Controls for number of firms and dispersion of employment across sectors. Firm data are spatially matched to congressional districts from Economic Census data, see appendix A.3 for details.

<sup>c</sup> State-specific time trends via restricted cubic splines.

<sup>d</sup> Partially linear model with flexible structure of confounders estimated via Post-double-selection LASSO to minimize omitted variable bias (Belloni et al. 2013, 2017). See appendix A.5 for details.

An alternative explanation not fully ruled out by the first two specifications is that union concentration might simply be a mechanical consequence of industrial structure. This is a relevant endogeneity problem, as firm or occupational structure (or the unobservables predicting them)

may well affect both union concentration and legislative outcomes. Our empirical exploration of the determinants of union concentration in Appendix A.4 indicates that economic structure alone is not sufficient to explain concentration. However, addressing this endogeneity concern also requires controlling for suitable proxies of economic structure in the analysis of legislative voting. We include two additional variables. The first is the number of firms, which we calculate from the Census Bureau’s Economic Census.<sup>25</sup> The second is the dispersion of employment over occupational sectors calculated from American Community Survey data on sectoral employment shares. Together, both measures capture key features of industrial structure. As can be seen from column (3), their inclusion does not appreciably change the estimated union effects. This strengthens our interpretation that union concentration is a distinct and relevant phenomenon and not just a mirror image of economic structure.

Our previous specifications focus on within-state variation between districts, and rule out as confounders both time-invariant state characteristics (such as culture) as well as idiosyncrasies of a given Congress. In specification (4) we aim to move closer to ruling out some time-varying state confounders as well, by including flexible state-specific time trends. Thus, any remaining omitted state-level variable would have to exert a highly nonlinear influence on congressional voting to bias our results. In specification (5) we move to subject our analysis to an even stricter test by estimating a partially linear model using the post-double selection estimator (Belloni et al. 2013, 2017). The key idea of this model is to account for confounders in a very flexible way by dropping the restriction of a linear effect of covariates and by allowing for arbitrary higher order interactions among covariates. This leads to hundreds of covariate terms, of which a subset is selected using standard variable selection tool (the LASSO). Importantly, this approach goes beyond a causally naive model selection tool by jointly estimating three equations: the outcome equation, where legislative ideology is the dependent variable, and selection (“treatment”) equations, where the dependent variables are union concentration and membership. This explicitly accounts for the logic of omitted variable bias: confounders (and their respective transformations) that matter in the selection stage are kept in the model even if they have only moderate weight in the outcome equation. For a more detailed exposition, see appendix A.5.

In both extended specifications, we find that the coefficient of union membership changes very little. Accounting for flexible state-specific time trends slightly increases our estimate for union concentration, while it decreases somewhat in the final specification. However, both specifications confirm our hypothesis that there is a statistically and substantively significant link between the concentration of union membership and legislative outcomes. Substantively, our

---

<sup>25</sup>See appendix A.3 for details.

results imply that a standard deviation increase in (log) union membership is related to a 0.16 ( $\pm 0.05$ ) standard deviation increase in liberal legislative ideology, while a standard deviation decrease in concentration is related to a 0.14 ( $\pm 0.04$ ) standard deviation increase in liberal ideology. These are politically relevant magnitudes and we demonstrate below that they are indeed sufficient to affect the outcomes of major votes.

#### *IV.B.1. Assessing Proximate District Confounders*

Table II assesses how robust our results are to the inclusion of what we call proximate district-level confounders. These are political confounders that may in part also be the result of union organization. One may think that pre-existing political preferences explain union size and concentration as well as why representatives are more left-leaning in some districts than in others. Several scholars argue that union strength is plausibly exogenous to political preferences, conditional on the macro context (i.e., union laws) and economic conditions, because the motivation to join or even form a union is mainly economic. For example, Olson (1965: 153-155) argues that union political action is a by-product of economic activism (see also Ahlquist and Levy 2013: 16). Nonetheless, it is prudent to consider the possibility that prior political preferences are an omitted variable. Otherwise one may remain worried that our results simply reflect district partisanship. To address this issue, we employ two proxies. Our first measure is the district-level Democratic vote share in the 2000 presidential election. Since election outcomes are partly a consequence of union political mobilization and perhaps even “preference activation” (Ahlquist and Levy 2013; Kim and Margalit 2016), we use the last presidential election prior to our period of analysis. Our second proxy is the average district-level political ideology of campaign donors, as inferred from their donation patterns by Bonica (2014). Including these variables means that our analysis partials out unions’ prior impact on district preferences and election outcomes. Specifications (1) and (2) in Table II show that adding both proxies does not affect the estimated effect of union concentration. The coefficient on union members is somewhat dampened, but it remains substantively important and significant at conventional levels.

We also account for (logged) corporate contributions to each legislator. This allows us to assess the concern that the concentration effect merely mirrors the structure of firms or business interests. Again, our findings on union organization remain robust.<sup>26</sup> Another specification includes the (logged) number of public union members in a district. Recent research has stressed

---

<sup>26</sup>We calculate corporate contributions similar to Carnes (2013) using the raw campaign finance contribution data from the Center for Responsive Politics. We sum FEC reported contributions to candidates from all “business” sectors (i.e., excluding labor and single-issue donations). Our count includes both individuals and PACs (but using either alone does not change our results).

TABLE II  
Proximate district level confounders.

	(1)	(2)	(3)	(4)	(5)
Union members	−0.059 (0.024)	−0.045 (0.022)	−0.065 (0.023)	−0.076 (0.021)	−0.032 (0.021)
Concentration	0.399 (0.105)	0.296 (0.102)	0.387 (0.124)	0.373 (0.117)	0.318 (0.111)
Presidential vote share <sup>a</sup>	✓				✓
Donor Ideology <sup>a</sup>		✓			✓
Public union members			✓		✓
Corporate contributions <sup>b</sup>				✓	✓
N	1767	1767	1767	1767	1767
R <sup>2</sup>	0.549	0.632	0.515	0.513	0.637

*Note:* 109th to 112th Congress. All models include state and congress fixed effects and district-level controls for income, race, education, employment in services and agriculture, number of firms, sectoral employment dispersion, and urbanization. Standard errors robust to arbitrary within-state correlation and heteroscedasticity.

<sup>a</sup> Presidential vote share in 2000 election. Donor ideology estimated by Bonica (2014) in the 108th congress. Missing values due to re-districting or insufficient information filled in using stochastic mean imputation (comprising 5.7% and 0.2% respectively).

<sup>b</sup> Total corporate contributions (cube-root transformed) to candidate calculated from data of the Center of Responsive Politics. Missing values (0.7%) filled in using stochastic mean imputation.

the particular characteristics of public unions and their political influence (e.g., Anzia and Moe 2015; Flavin and Hartney 2015; Moe 2006). Hence, one may ask whether our results are mostly driven by public unions or if the sign of the effect changes when excluding public unions. Since the LM data does not allow us to exactly identify all public unions and their members (recall the discussion above), we calculate a district’s number of public union members from an indicator created by selecting likely public unions based on their name.<sup>27</sup> Even after accounting for the degree of public membership in a district, our results show rather little change in our two union variables.

In sum, these additional results rule out pre-existing political preferences or partisan leanings, as well as corporate strength, as an alternative explanation. A number of further specification and robustness tests, reported in Appendix A.6, show that our findings hold under various alternative measurements and modeling assumptions.

<sup>27</sup>In other words, we select unions that are likely to contain public employees by using regular expressions containing terms such as “firefighters”, “police”, “county” or “public” employees. While this does not, of course, yield a precise classification of public unions, it captures the degree to which (likely) public employees are present in a given district.

#### *IV.B.2. Contributions and Partisan Selection*

In the following, we try to shed some light on potential channels of influence. Based on our argument, the two dimensions of union organization are linked to the *selection* of partisan politicians, with unions rallying around Democratic candidates. We expect lower levels of concentration to be associated with a higher probability of a Democratic victory. Similarly, lower concentration should be associated with more *financial contributions* by unions and union members to sitting members of Congress. Again, alternative arguments based on coordination problems or preference heterogeneity suggest, if anything, the opposite. Table III indicates the empirical relevance of union organization for both contributions and partisan selection, supporting the logic behind our main hypothesis. The first column presents results from a linear probability model where the dependent variable indicates if a legislator is affiliated with the Democratic Party. The model accounts for observable district characteristics, as well as state and Congress fixed effects. Clearly, districts with more union members and lower levels of concentration show a higher propensity of selecting a Democrat, all else equal. In the second column we study the influence of local union characteristics on financial contributions. The dependent variable is the amount of candidate contributions from labor.<sup>28</sup> We find that local union structure relates to candidate contributions: a unit increase in (logged) union membership in a district raises labor contributions by \$11,500. Even more notably, a unit decrease in concentration of members leads to a \$66,800 increase in labor contributions, holding everything else constant.

#### *IV.C. Union influence on key votes in the 111th Congress*

We have shown that union characteristics matter for legislators' revealed ideology. So far, we have measured the former as the one-dimensional (unobserved) variable underlying representatives' (observed) voting record. We now turn to individual votes on key legislation where unions took explicit positions. In particular, we study all 35 major pieces of legislation in the 111th Congress that are part of the political scorecard of the AFL-CIO and code whether a representative votes in line with the union recommendation. These votes include, among others, the final passage of the Affordable Care Act (H.R. 3590) on March 21, 2010, and the Jobs for Main Street Act (H.R. 2847) on December 16, 2009. Health care reform is widely considered one of the major reforms of the new century, something that unions had supported for decades but failed to achieve. Both votes were razor-tight and highly partisan (219 vs. 211 and 217 vs. 212 in favor).

---

<sup>28</sup>It is calculated in the same way as corporate contributions (see above), using contributions from the labor sector.

TABLE III  
Candidate selection and labor contributions.

	Democrat <sup>a</sup>	Labor contributions <sup>b</sup>
Union members	0.068 (0.019)	0.115 (0.038)
Concentration	-0.319 (0.114)	-0.668 (0.245)
N	1765	1767
R <sup>2</sup>	0.430	0.309

*Note:* Models include state and congress fixed effects and district-level controls (see Table II). Standard errors robust to arbitrary within-state correlation and heteroscedasticity.

*a* Linear probability model.

*b* Labor contributions are in 100,000s of Dollars. Model estimated on cube-root transformed LHS, with estimated coefficients transformed back to Dollar scale.

We estimate logit models for each key vote accounting for district characteristics (the same as in Table I) and state random effects with standard errors robust to within-state correlations.<sup>29</sup> Figure III plots first differences in the probability of a representative casting a vote in line with the AFL-CIO supported position together with 95% confidence intervals.<sup>30</sup> It underscores the substantive effect of union organization on legislative behavior. In many of the votes that matter to the AFL-CIO, a standard deviation increase in membership increases the probability of a representative voting the pro-union position by almost 10 percentage points. The respective effect size for the influence of union concentration is around 5 points (all else equal). In the case of the final passage of the Affordable Care Act, a standard deviation increase in union membership increases the probability of a supportive vote by 11.7 ( $\pm 3.7$ ) percentage points, while an equally sized decrease in union concentration increases it by 6.5 ( $\pm 2.0$ ) points. In contrast, voting in the far less divisive “cash for clunkers” bill (H.R. 3435, passed 316 vs. 109 in favor) is comparatively less related to union influence: an increase in union membership increases the vote probability of a supportive vote by 4.8 ( $\pm 2.5$ ) points, while a decrease in union concentration increases it by 4.4 ( $\pm 1.9$ ) percentage points.

<sup>29</sup>The small number of districts per state (and the existence of at-large districts) would make models including state dummies (instead of random effects) subject to incidental parameters bias.

<sup>30</sup>We calculate this difference in probability by increasing the membership variable and respectively decreasing the concentration variable by one standard deviation from the mean while holding all other covariates at their observed value.

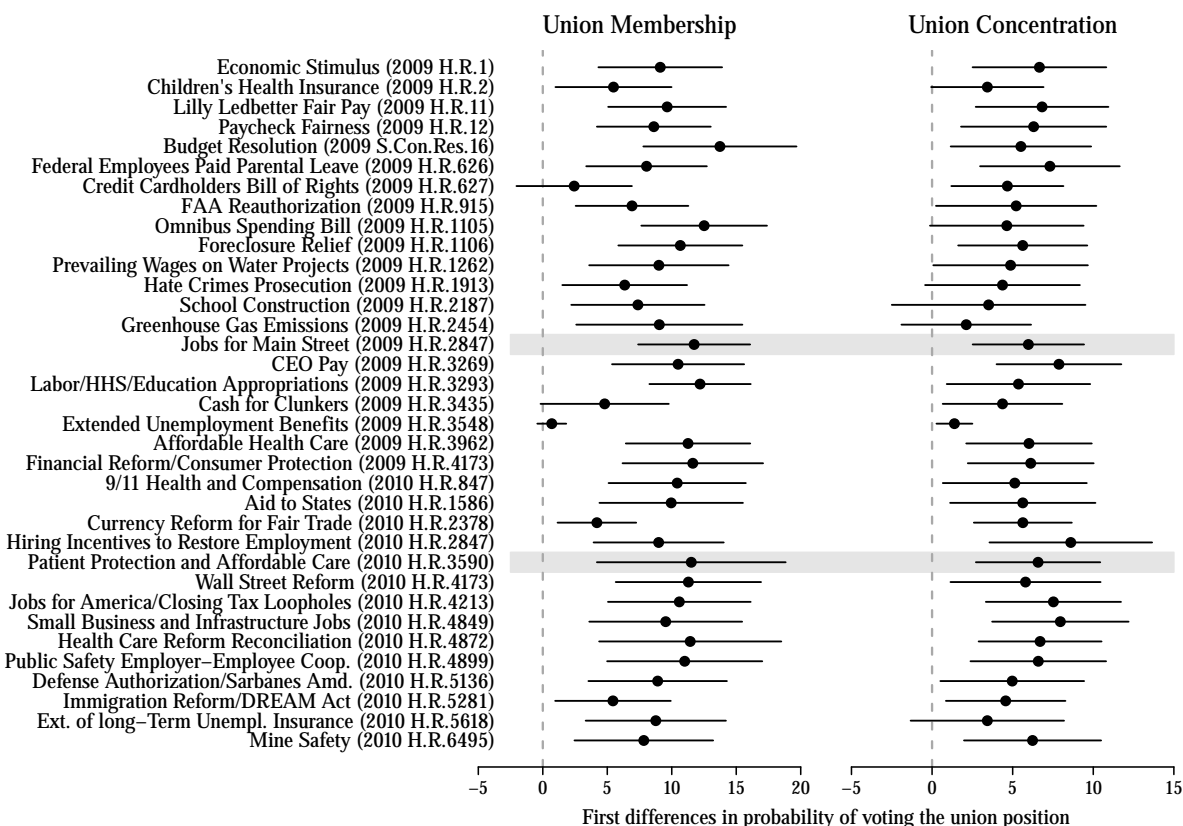


FIGURE III  
Unions and key votes in the 111th Congress

This figure shows the influence of union characteristics on key votes (as defined by AFL-CIO) in 2009 and 2010. Each row entry represents first differences in the probability of a representative casting a vote in line with the AFL-CIO supported position as function of a SD increase in membership and a SD decrease in concentration. All models include district controls and state random effects and allow for arbitrary within-state correlation. Line segments are 95% confidence intervals.

## V. CONCLUSION

How does the local organization of membership groups affect their political power to shape lawmaking in the national legislature? Going beyond the literature's often exclusive focus on membership size, we have provided a theoretical rationale for taking into account other organizational features and conducted an empirical investigation of their political effects. Our focus has been on one of the most important membership groups active in democratic politics: labor unions. Drawing on extensive and geographically fine-grained administrative data, we were able to measure two distinct dimensions of local union organization at the level of congressional districts: membership size and the horizontal concentration vs. dispersion of union members in



different organizational units. Analyzing legislative voting in the U.S. House of Representatives, this is the first study to uncover what we call the concentration effect: For a given level of union membership, representatives from districts where membership is relatively dispersed across multiple unions have more liberal voting records than representatives from districts where members are concentrated. Our analysis demonstrates that the effect of union concentration on legislative voting is not merely an artifact of economic structure, the political power of business, district partisanship, or model assumptions.

Our results have important implications concerning the relevance of labor unions for democratic politics. Previous research has ignored the concentration dimension of local unions, and has underestimated the overall political significance of organized labor for national lawmaking. To be clear, we confirm the standard view that group size matters. In addition, however, the district-level concentration of union membership is of nearly equal importance. Our analysis of both DW-NOMINATE scores and individual key votes demonstrates that the two dimensions of local union organization, size and concentration, affect the making of national laws on a broad range of policies. These include policies with large ramifications for the economy and where mass policy preferences are significantly polarized by income. Over the last decade, research has shown that elected representatives are more responsive to high income groups (e.g., Bartels 2008). Our findings suggest that not just higher overall membership (in line with the state-level analysis of Flavin 2016) but also lower concentration implies a move towards more political equality. The role of concentration has largely been missing from the debate about union membership decline. While concentration is more or less constant in the period we study, our results suggest that a decrease in concentration may partially offset, up to a point, the political significance of declining membership; in contrast, increasing concentration would augment it.

Furthermore, our findings shed light on different theoretical perspectives on the effect of group organization on political mobilization. A widely held view is that organizational dispersion of group members is a political disadvantage because it increases coordination or collective action problems between group leaders. In contrast, our findings are consistent with the theory of political mobilization based on social interactions in small groups. They suggest that social incentives for political mobilization may be stronger than the presumed perils of a more dispersed membership. This is broadly consistent with a large literature on the significance of social incentives for political behavior. Social incentives are a powerful engine for political action, and they are more pronounced where interactions are less anonymous. One may also conjecture that there is an additional indirect effect. Local group leaders might be more accountable to their members when membership is less concentrated, and so free-riding by leaders is reduced.

As a result, we think that, beyond unions, the concentration dimension of group organization

makes for a relevant ingredient in more realistic theories of democratic politics. The mechanism underlying our concentration hypothesis is general. Social pressure as a motive for political action is not restricted to unions. Thus, one may expect a similar pattern for, say, churches of the same denomination. Testing this possibility is a clear avenue for further research.

## REFERENCES

- Aaron, B. (1960). The labor-management reporting and disclosure act of 1959. *Harvard Law Review* 73(5), 851–907.
- Abrams, S., T. Iversen, and D. Soskice (2011). Informal social networks and rational voting. *British Journal of Political Science* 41(2), 229–257.
- Ahlquist, J. S., A. B. Clayton, and M. Levi (2014). Provoking preferences: Unionization, trade policy, and the ilwu puzzle. *International Organization* 68, 33–75.
- Ahlquist, J. S. and M. Levi (2013). *In the Interests of Others*. Princeton and Oxford: Princeton University Press.
- Anzia, S. F. (2011). Election timing and the electoral influence of interest groups. *Journal of Politics* 73(2), 412–427.
- Anzia, S. F. and T. M. Moe (2015). Public sector unions and the costs of government. *Journal of Politics* 77(1), 114–27.
- Bartels, L. (2008). *Unequal Democracy*. Princeton and Oxford: Princeton University Press.
- 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, A., V. Chernozhukov, and C. B. Hansen (2013). Inference for high-dimensional sparse econometric models. In D. Acemoglu, M. Arellano, and E. Dekel (Eds.), *Advances in Economics and Econometrics: Tenth World Congress*, Volume 3, pp. 245–295. Cambridge: Cambridge University Press.
- Berelson, B. R., P. F. Lazarsfeld, and W. McPhee (1954). *Voting: A Study of Opinion Formation in a Presidential Campaign*. Chicago and London: University of Chicago Press.
- Bonica, A. (2014). Ideology and interests in the political marketplace. *American Journal of Political Science* 57(2), 294–311.
- Box-Steffensmeier, J. M., L. W. Arnold, and C. J. W. Zorn (1997). The strategic timing of position taking in congress: A study of the north american free trade agreement. *American Political Science Review* 91(2), 324–338.
- Brady, D., R. S. Baker, and R. Finnigan (2013). When unionization disappears: State-level

- unionization and working poverty in the united states. *American Sociological Review* 78(5), 872–896.
- Busch, M. L. and E. Reinhardt (2000). Geography, international trade, and political mobilization in u.s. industries. *American Journal of Political Science* 44(4), 703–719.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics* 90(3), 414–427.
- Cameron, C. A. and P. K. Trivedi (2005). *Microeconometrics. Methods and Applications*. Cambridge: Cambridge University Press.
- Card, D. (1996). The effect of unions on the structure of wages: A longitudinal analysis. *Econometrica* 64(4), 957–979.
- Carnes, N. (2013). *White-Collar Government: The Hidden Role of Class in Economic Policy Making*. Chicago, IL: University of Chicago Press.
- Coleman, C. J. (1980). The civil service reform act of 1978: Its meaning and its roots. *Labor Law Journal* 31(4), 200–207.
- Curry, B. and K. D. George (1983). Industrial concentration: A survey. *The Journal of Industrial Economics* 31(3), 203–255.
- Dark, T. E. (1999). *The Unions and the Democrats*. Cornell: Cornell University Press.
- Druckman, J. N. and D. P. Green (2013). Mobilizing group membership: The impact of personalization and social pressure e-mails. *SAGE Open* 3(2), 1–6.
- Flavin, P. (2016). Labor union strength and the equality of political representation. *British Journal of Political Science*, (forthcoming).
- Flavin, P. and M. 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.
- Freeman, R. B. and J. Medoff (1984). *What Do Unions Do?* New York: Basic Books.
- Gerber, A. S., D. P. Green, and C. W. Larimer (2008). Social pressure and voter turnout: Evidence from a large-scale field experiment. *American Political Science Review* 102(1), 33–48.
- Goldberg, D. W., J. P. Wilson, and C. A. Knoblock (2007). From text to geographic coordinates: The current state of geocoding. *Urban and Regional Information Systems Association Journal* 19(1), 33–46.
- Hirsch, B., D. Macpherson, and W. Vroman (2001). Estimates of union density by state. *Monthly Labor Review* 124(7), 51–55.
- Iversen, T. (1999). *Contested Economic Institutions: The Politics of Macroeconomics and Wage Bargaining in Advanced Democracies*. Cambridge and New York: Cambridge University Press.

- Kau, J. B. and P. H. Rubin (1978). Voting on minimum wages: A time-series analysis. *Journal of Political Economy* 86(2), 337–342.
- Key, V. O. (1964). *Politics, Parties, and Pressure Groups* (5 ed.). New York: Thomas Y. Crowell Company.
- Kim, S. E. and Y. Margalit (2016). Informed preferences? the impact of unions on workers’ policy views. *American Journal of Political Science*, (forthcoming).
- Kremer, M. and B. A. Olken (2009). A biological model of unions. *American Economic Journal: Applied Economics* 1(2), 150–75.
- Lee, D. S., E. Moretti, and M. J. Butler (2004). Do voters affect or elect policies? evidence from the u. s. house. *Quarterly Journal of Economics* 119(3), 807–859.
- Leighley, J. E. and J. Nagler (2007). Unions, voter turnout, and class bias in the u.s. electorate, 1964–2004. *Journal of Politics* 69(2), 430–441.
- Lichtenstein, N. (2013). *State of the Union: A Century of American Labor* (2 ed.). Princeton and Oxford: Princeton University Press.
- Martin, A. W. (2008). The institutional logic of union organizing and the effectiveness of social movement repertoires. *American Journal of Sociology* 113(4), 1067–1103.
- Masters, M. F. and J. T. Delaney (2005). Organized labor’s political scorecard. *Journal of Labor Research* 26(3), 365–392.
- McCarthy, N., K. T. Poole, and H. Rosenthal (2006). *Polarized America*. Cambridge, MA; London, England: MIT Press.
- Moe, T. M. (2006). Political control and the power of the agent. *Journal of Law, Economics, & Organization* 22(1), 1–29.
- Olson, M. (1965). *The Logic of Collective Action*. Cambridge: Harvard University Press.
- Pontusson, J., D. Rueda, and C. R. Way (2002). Comparative political economy of wage distribution: The role of partisanship and labour market institutions. *British Journal of Political Science* 32(2), 281–308.
- Poole, K. T. and H. Rosenthal (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Radcliff, B. and P. Davis (2000). Labor organization and electoral participation in industrial democracies. *American Journal of Political Science* 44(1), 132–141.
- Robison, E. and C. Grieses (2014). Panel analysis of household nonresponse and person coverage in the current population survey. Paper presented in the Survey Research Methods Section at the Joint Statistical Meetings of the American Statistical Association, Boston, 2014.
- Rosenfeld, J. (2014). *What Unions No Longer Do*. Cambridge, MA and London, England: Harvard University Press.

- Schlozman, D. (2015). *When Movements Anchor Parties*. Princeton and Oxford: Princeton University Press.
- Seltzer, A. J. (1995). The political economy of the fair labor standards act of 1938. *Journal of Political Economy* 103(6), 1302–1342.
- Southworth, C. and J. Stepan-Norris (2009). American trade unions and data limitations: A new agenda for labor studies. *Annual Review of Sociology* 35, 297–320.
- Wallerstein, M. and B. Western (2000). Unions in decline? what has changed and why. *Annual Review of Political Science* 3, 355–377.
- Zullo, R. (2008). Union membership and political inclusion. *Industrial and Labor Relations Review* 62(1), 22–38.

# ONLINE APPENDIX TO LOCAL UNION ORGANIZATION AND LAWMAKING IN THE U.S. CONGRESS

## CONTENTS

A.1	Comparison of LM form and CPS measures of union density . . . . .	1
A.2	Descriptive information . . . . .	3
A.3	Construction of district-level number of firms and urbanization . . . . .	4
A.4	What explains union concentration? . . . . .	7
A.5	Double-post-selection LASSO estimation . . . . .	13
A.6	Additional robustness tests . . . . .	15
A.7	The structure of union membership . . . . .	16
A.8	Interactive specifications . . . . .	18

### *A.1 Comparison of LM form and CPS measures of union density*

How does our union data drawn from public records compare to existing measures of union density drawn from survey data? In Figure A.1.1 we aggregate our measure of unionization based on LM forms filed to the Department of Labor to yearly state averages (on the x-axis) and plot it against unionization estimates based on the Current Population Survey calculated by Hirsch and Macpherson (2003). Their data are available at [unionstats.com](http://unionstats.com) and widely used in political science, economics, and sociology. Note that they also offer sub-state data, but these are confined to metropolitan statistical areas. See Hirsch et al. (2001) for more details.

First, note the high correspondence between the two measures. On average unionization levels based on estimates from the CPS are about 1.9 percentage points higher than counts of members from LM forms, and the correlation is 0.86. As has been noted before, the wording of the CPS question is likely to induce over-reporting of union membership as it asks about the membership in a union or “similar” employee associations (Southworth and

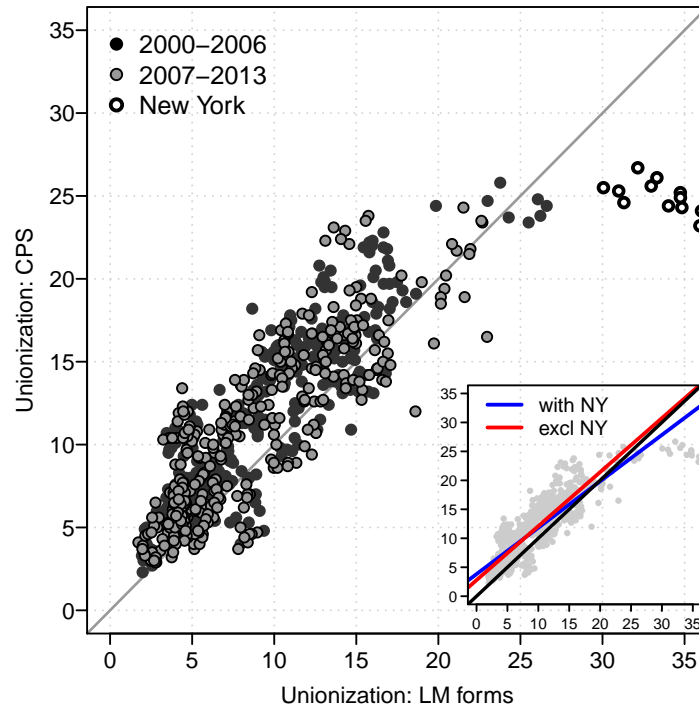


FIGURE A.1.1

#### Comparison of LM-form and CPS estimates of state unionization levels

This figure shows the close correspondence between state-level unionization calculated based on LM forms and CPS-based estimates. We split our data into two periods (pre- and post-2006) to illustrate the lack of time patterns. Open circles mark the state of New York. Red and blue lines in the bottom right plot are least-square fits excluding and including New York.

Stepan-Norris 2009: 311). Thus the difference is an upper bound of the non-coverage of some public sector unions in our data. The high similarity between our aggregated LM forms data and the CPS measure is confirmed by the linear fit between both measures being close to 1, plotted by the blue line in the bottom right corner of Figure A.1.1. Second, note that a comparison of LM-CPS pairs in Figure A.1.1 for the years prior and post 2006 yields no apparent pattern, indicating that the gap between both measures does not increase over time. This can be quantified by an F-test of different trends between both periods. The null hypothesis of no difference in trends cannot be rejected at  $F = 1.60$ .

In some cases greater differences are apparent. The most obvious case is that of New York (marked by circles), where in all years the share of unionized workers based on counts of LM form reports is higher than that based on CPS survey responses. Note that if we exclude New York, the correspondence between our measure and the CPS-based one is even higher, as shown by the second linear fit line (in red). A plausible explanation is that differences in this case are due to differences between the location of the union office and workplace (New York city) and workers' residences (in nearby New Jersey).

## A.2 Descriptive information

Table A.2.1 shows descriptive statistics and sources for the variables included in our models.

TABLE A.2.1  
Descriptive statistics

	Mean	SD	Min	Max	N
W-Nominate Dim 1 scores <sup>a</sup>	0.14	0.53	-0.75	1.36	1767
Union members [log]	9.66	1.11	2.67	13.73	1767
Union concentration [CR4]	0.57	0.20	0.14	1.00	1767
Median HH income [10,000\$] <sup>b</sup>	5.19	1.38	2.11	10.74	1767
Share white [0-1] <sup>b</sup>	0.65	0.23	0.02	0.97	1767
Share BA or higher [0-1] <sup>b</sup>	0.28	0.10	0.07	0.66	1767
Share service sector empl. [0-1] <sup>b</sup>	0.18	0.03	0.09	0.41	1767
Share agriculture empl. [0-1] <sup>b</sup>	0.01	0.01	0.00	0.22	1767
Number of Firms [10,000s] <sup>c</sup>	1.71	0.36	0.78	4.27	1767
Employment dispersion [0-1] <sup>d</sup>	0.62	0.10	0.24	0.96	1767
Share of urban households [0-1] <sup>c</sup>	0.80	0.19	0.23	1.00	1767
Presidential vote share 2000 [0-1] <sup>e</sup>	0.52	0.15	0.20	0.94	1711
Donor scores <sup>e</sup>	0.09	0.40	-1.10	0.92	1763
Representative's party affiliation [0,1] <sup>e</sup>	0.51	0.50	0.00	1.00	1765
Corporate contributions [10 <sup>6</sup> \$] <sup>f</sup>	0.89	1.17	-0.01	15.98	1755
Labor contributions [10 <sup>6</sup> \$] <sup>f</sup>	0.10	0.11	0.00	1.16	1755

<sup>a</sup> Data from Poole and Rosenthal (1997); <http://www.voteview.com/>

<sup>b</sup> Data from American Community Survey.

<sup>c</sup> Calculated from County Business Patterns Data from the Economic Census Economic Census, and from the general population Census. See Section A.3 for details.

<sup>d</sup> Calculated for each district as  $([\sum_{i=1}^K s_i]^{-1} - 1)/(K - 1)$ , where  $s \in [0, 1]$  are employment shares in  $K = 6$  occupational categories in the American Community Survey (table S2401).

<sup>e</sup> Data from Bonica (2014).

<sup>f</sup> Calculated from raw data collected by the Center for Responsive Politics ([www.opensecrets.org](http://www.opensecrets.org)). Negative values are possible since donation amounts include refunds.

Figure A.2.1 shows the distribution of the median size of unions in house districts in the 109th to 112th Congress.

Figure A.2.2 plots district-level union concentration over time.



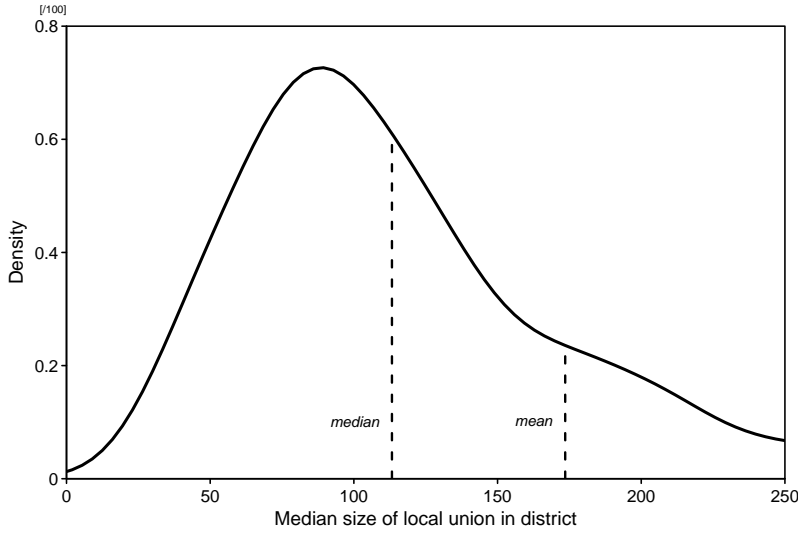


FIGURE A.2.1  
Distribution of Median Size of Local Unions in House Districts.

This figure shows a kernel density plot of the median size of local unions in each house district in the 109th to 112th Congress. Mean and median of this distribution are indicated by dashed lines.

### *A.3 Construction of district-level number of firms and urbanization*

In order to rule out potential confounders of the relation between our union concentration measure and the legislative ideology scores of a district's representative, we control for the structure of employment and urbanization.

Measures of the number of firms in a given district, as well as a districts level of urbanization are not readily available for the years covered in our study.<sup>1</sup> Thus, we spatially aggregate county-level measures from the Economic and the General Census to the congressional district level using asymmetric weighting (Mennis and Hultgren 2006). See Goplerud (2015) for a recent introduction to political science and an illustration of its superiority to simply weighting by area. To employ this technique, we require three spatial layers of which the first (counties) is to be mapped to the second (congressional districts). The third layer provides auxiliary information used to predicting the spatial distribution of the quantity of interest in the first spatial unit (counties) based on a third variable. We use voting districts and their population in the 2010 census as the auxiliary layer. They are ideal for our application for three reasons. First, while voting district boundaries can change over time, they are generally nested within both counties and congressional districts. Second, they are geographically fine-grained (typically covering a low 4-digit number of voters). Third, the population distribution across them is likely strongly related to the distribution of firms

<sup>1</sup>For instance, data from the economic census aggregated to the congressional district level are only available from the 113th congress onwards.

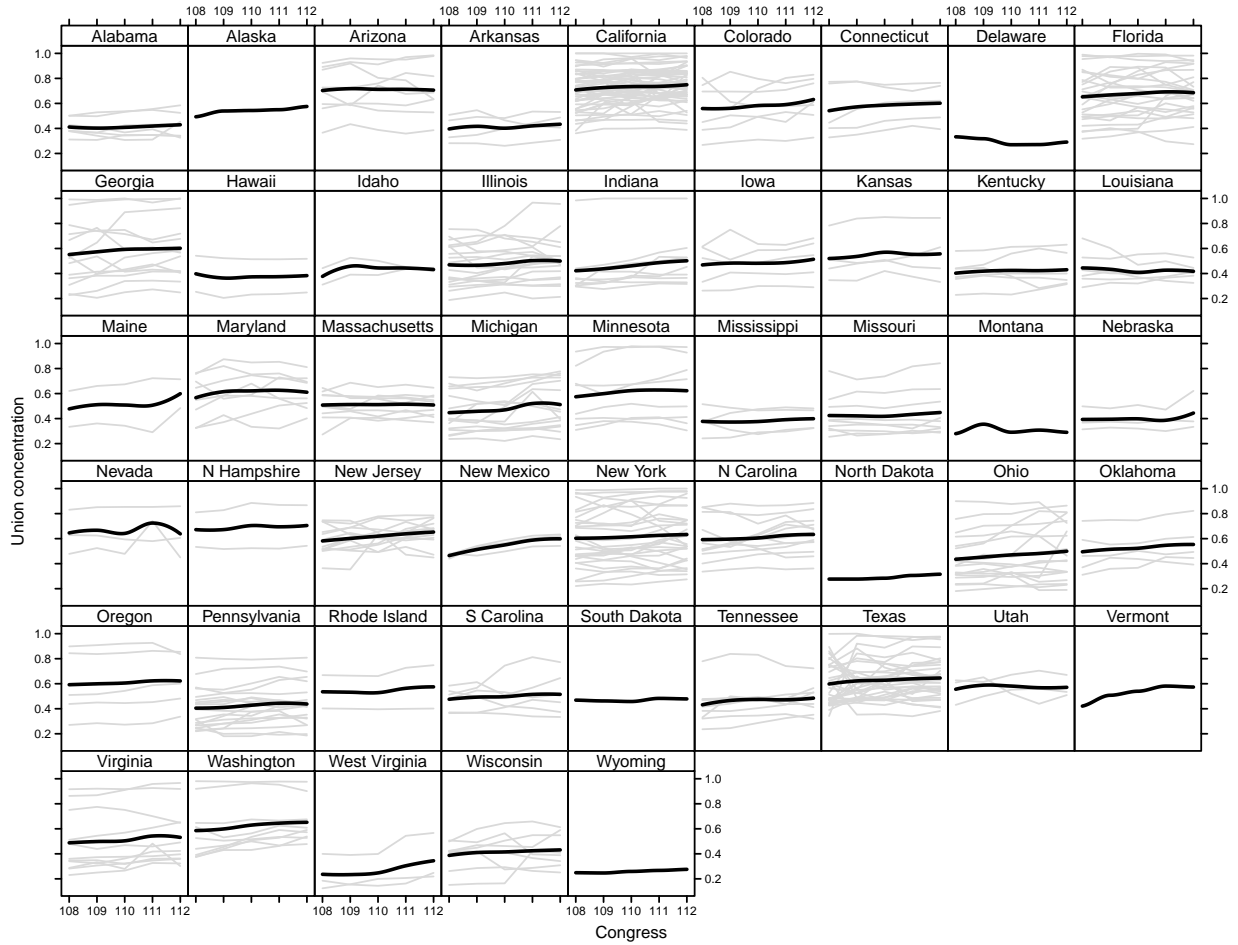


FIGURE A.2.2  
Union concentration over time.

This figure shows union concentration from the 108th to 112th Congress. Shown are levels for each district in each state (thin gray lines), as well as lowess-smoothed averages by state (black lines).

and the rate of urbanization. We use shapefiles provided by the Topologically Integrated Geographic Encoding and Referencing program of the US Census Bureau for counties in 2013, and voting districts in 2010 to map county-level data to each congressional districts in a given year.<sup>2</sup>

For every congress, we aggregate (1) the number of establishments (firms) from the County Business Patterns Series from the respective Economic Census; (2) the share of

<sup>2</sup>For states with at-large districts, the aggregation procedure reduces to a simple summation, as counties by definition are perfectly nested within districts/states. For three states (Kentucky, Oregon, Rhode Island), no or incomplete voting district data is available from the Census Bureau, and we use 2010 county-subdivision as the auxiliary layer instead. For the remaining states, around 3.5% of the precinct-level population estimates were missing and ignored in the procedure.

households in urbanized areas out of total number of households from the Census's American Fact Finder.

To examine the accuracy of our aggregation strategy, we benchmark it against available district-level data from official sources in single years. The first panel of Figure A.3.1 below shows the relation between the number of firms in the districts of the 113th congress (available from the County Business Patterns database) to our county-based aggregation. The second panel compares our aggregated measure of urban households to data available from the American Fact Finder for the 111th congress. The high correlations in Figure A.3.1 underline the high quality of the aggregation procedure.

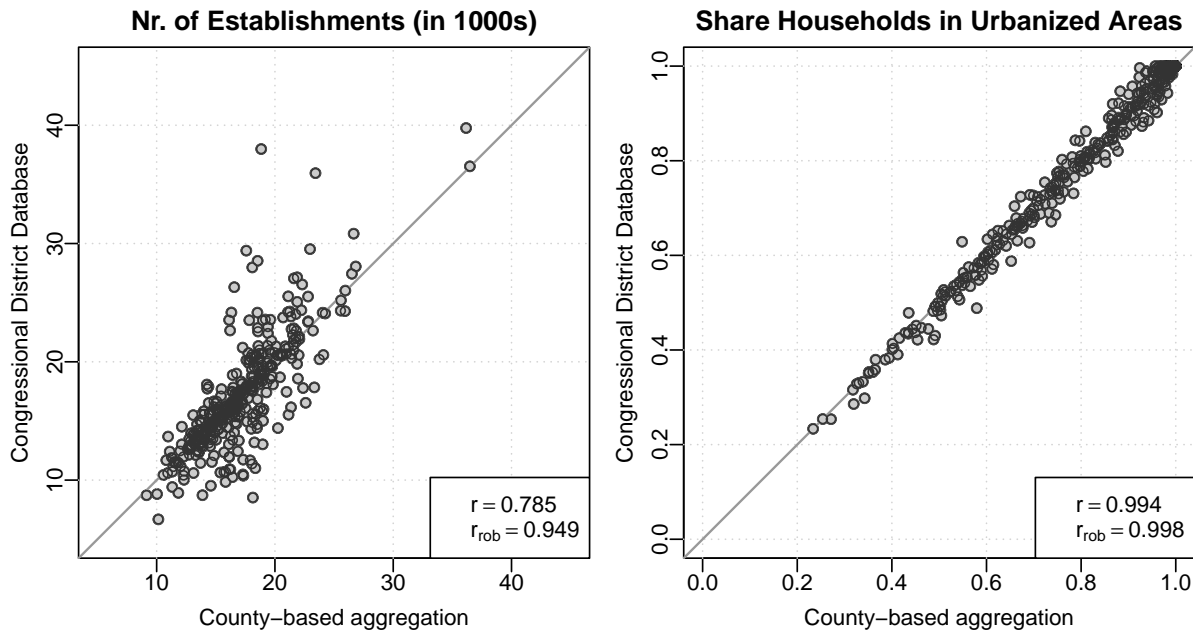


FIGURE A.3.1  
Benchmark of County-Based Spatial Aggregation

This figure shows the relationship between district-level measures of business structure and urbanization provided directly by the Census for the 113th and 111th congress (y-axis) with our spatially interpolated and aggregated measures (x-axis). Displayed in the bottom right corner are correlation and robust-correlation coefficients.

#### *A.4 What explains union concentration?*

Why are some congressional districts characterized by more union concentration than others despite similar levels of overall union membership (recall figures I and II)? While there is a well-developed multi-disciplinary literature on the determinants of union membership and union density (for a review, see Wallerstein and Western 2000), we are aware of no study that explains union concentration. One may conjecture that economic structure – the distribution of firms, occupations, or sectors – more or less completely determines union concentration. Our empirical results on the effects of union organization on legislative voting already indicate that this is not the case. Recall that there is a clear political effect of union concentration even after accounting for aspects of a district's economic structure, such as the number of firms, employment concentration, the relevance of agriculture, services, and the degree of urbanization. This section aims to shed more light on the question of varying union concentration.

Theoretically, we argue that union concentration is shaped by the interaction between economic structure and organizational preferences of workers and their union leaders that are rooted in history. Empirically, we draw on public opinion data from the 1930s – a turning point in the history of organized labor in the US where membership increased dramatically – to provide an initial empirical assessment of the argument. In line with relevant implications of the argument, we find that already in the New Deal era public opinion concerning labor unions clearly distinguishes between (1) the support of unions' rights to organize and strike, and (2) *how* unions should be organized, in particular, whether they should be encompassing or concentrated. Furthermore, a regression analysis finds that historical preferences for union concentration moderate the relationship between economic structure and union concentration in the twenty-first century. In contrast (conducting an informal placebo test) we do not find a significant moderating effect of historical preferences for union rights.

The years between the two world wars were characterized by competing views within the labor movement on whether to organize workers along craft lines, as advocated by the pre-dominant American Federation of Labor, or more broadly along industry lines, as proposed by the break-away Congress of Industrial Organization (e.g., Lichtenstein 2013). The question of whether unions' organizational structure should be relatively encompassing or narrow was separate from the question of unions' rights to organize workers. Different organizational goals may ultimately lead to varying union concentration in a similar economic environment. While it is clear that the distribution of economic establishments puts a constraint on union structure, it leaves a considerable amount of slack to be shaped by organizational preferences that, once established, are quite persistent. Congressional districts include a multitude of establishments that may or may not be organized and, if organized, may be organized by the same or different unions. As is shown in Table A.2.1, the average number of firms per district is above 17,100 and the minimum is 7,800. Competing

unions may organize workers in different establishments in the same industry, or workers in the same occupation in different establishments. By the same token, the same union may succeed to organize workers in different industries.

Drawing on the theoretical framework of Ahlquist and Levy (2013) helps to clarify possible mechanisms that link initially divergent views about organizational structure to variation in union concentration. In the framework, the development of an individual union is shaped to a significant degree by the preferences and skills of the founding leadership, which is subject due to idiosyncratic shocks. As in political parties or firms, asymmetric information is a common problem in principal agent relationships. If a union leader proves successful in delivering tangible economic benefits to members and the union thrives, members are more likely to embrace the broader organizational goals set by the leaders as their own. This leads to some degree of path-dependency. Applied to the question of organizational structure, the framework suggests that unions steeped in a tradition of narrow organizational structure will be less able and willing to organize beyond their initial membership base, and the result will be less union concentration. For our purposes, it is not important whether elite preferences shape mass preferences or vice versa, but the organizational theory of Ahlquist and Levy (2013) highlights one plausible direction of influence, consistent with recent evidence (Ahlquist et al. 2014; Kim and Margalit 2016).

Taken together, these considerations lead to several observable implications that we can assess with available data. Going back to the critical period of union growth in the New Deal years, we should observe that preferences for union organization are distinguishable from preferences for union rights more broadly. Furthermore, there should be a link between historical organizational preferences and current levels of union concentration. In particular, the hypothesis is that historical organizational preferences condition the link between contemporaneous economic structure and union concentration. Where support for more encompassing union organization was high about 80 years ago, the relationship between economic structure and union structure should be weaker.

To empirically assess these implications, we draw on national public opinion surveys that were conducted by Gallup in 1936-7 and are available from the Roper Center for Public Opinion Research ([ropercenter.cornell.edu](http://ropercenter.cornell.edu)). Reflecting the large salience of the union issue at the time, the surveys (dataset USAIPO1936-3637) include several items on the respondents' view about organized labor. Directly tapping into the organizational dimension, one item concretely asks whether respondents prefer a narrow or encompassing union structure (N = 5,790). It concerns the elite divisions about organizational strategy mentioned above. More specifically, the question wording is as follows:

Which do you favor:

1. One single union for all workers in an industry (Industrial Union), OR
2. Separate unions for each craft in an industry (like carpenters, masons, machinists, etc)?

Respondents choosing the first option are effectively favoring union concentration. Respondents choosing the second option are against union concentration. Another item deals with union rights rather than with how unions should be organized. In particular, the question ask respondents whether the militia should be “called out whenever strike trouble threatens” (N=2,904). Respondents answering yes take a clear anti-union stand. We aggregate responses to the state level (there are no sub-state identifiers). As the surveys were conducted using quota-control methods rather than random sampling, which can lead to systematic deviations between the survey samples and the population, we follow Berinsky (2006) and employ a post-stratification weighting adjustment. Using 1940s state-level census estimates of population proportions, we we achieve a weighted distribution of sampled groups that correspond to the population.

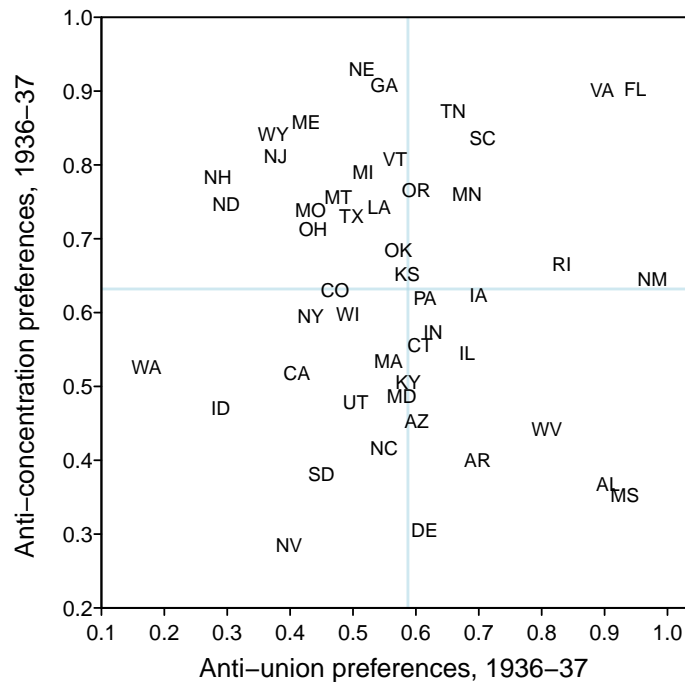


FIGURE A.4.1  
Historical preferences against unions and union concentration.

Based on Gallup surveys from 1936-7, this figure shows the state-level relationship between the fraction of respondents who are against unions’ right to strike (x-axis) and the fraction of respondents who prefer unions to organize along narrow (i.e., craft) lines (y-axis). We use post stratification weights to balance the distribution from the Gallup quota sample against population cells obtained from the 1940 Census.

Figure A.4.1 plots the resulting state-level measures of anti-union preferences (fraction who support using the militia against strikes) and preferences concerning union concentration (fraction who prefer separate unions) the late 1930s. As is evident from the scatterplot, the correlation between two dimensions is low ( $r = -0.05$ ). Thus, aggregate preferences

for how unions should be organized do not reduce to aggregate opposition to union rights. The descriptive evidence bolsters the case that concentration is an important if neglected dimension of organized labor. Note that there is a resemblance between Figure A.4.1 and Figure II in the main text, which shows that (district-level) union membership and union concentration are two distinct dimensions of realized union organization. Of course, one may ask what explains public support for concentration in the 1930s; settling this question is beyond the scope of this paper.

TABLE A.4.1  
Historical origins of union concentration

	(1)	(2)	(3)	(4)	(5)
Employment dispersion					
×anti-concentration preferences 1930s	−0.025 (0.010)	−0.025 (0.010)	−0.025 (0.010)	−0.027 (0.010)	−0.028 (0.009)
×anti-union preferences 1930s	−0.009 (0.010)	−0.007 (0.010)	−0.007 (0.010)	−0.007 (0.010)	−0.007 (0.009)
Employment dispersion	−0.074 (0.009)	−0.074 (0.009)	−0.074 (0.010)	−0.083 (0.011)	−0.106 (0.021)
Anti-Concentration preferences	0.007 (0.015)	0.003 (0.015)	0.003 (0.014)	0.004 (0.014)	0.004 (0.013)
Anti-Union preferences	−0.020 (0.017)	−0.021 (0.017)	−0.023 (0.016)	−0.027 (0.017)	−0.029 (0.017)
Slavery, 1860 (state-level) <sup>a</sup>		✓	✓	✓	✓
Union legislation (state-level) <sup>b</sup>			✓	✓	✓
Number of firms (district-level)				✓	✓
Socio-economic profile (district-level) <sup>c</sup>					✓
N	432	429	429	429	429

*Note:* Dependent variable is average district-level union concentration (CR4), 2003-2012. Historical union preference variables are standardized to have mean zero and unit variance. All models contain state-specific constants. RE-GLS with robust standard errors allowing for heteroscedasticity and within-state correlation.

<sup>a</sup> State-share of slaves in 1860 based on Mitchener and McLean (2003).

<sup>b</sup> (Anti-)union legislation measured using two variables. These count the share of years between 1950 and 2004 where a state had right-to work or collective bargaining legislation. Based on Flavin and Hartney (2015) and Holmes (1998).

<sup>c</sup> Summary measure of 5 district variables: income, education, share of black, service sector employment, and urbanity. First principal factor explaining 63% of total variance (first two eigenvalues are 2.45, 1.05).

Next, we turn to examining the hypothesis that historical concentration preferences condition the contemporaneous link between economic structure and union concentration. The analysis is somewhat indirect in that we have to rely on state-level measures of concentration preferences. Table A.4.1 reports a series of cross-sectional regressions that regress average district-level union concentration (CR4) 2003-2012 on district-level employment dispersion, state-level anti-concentration preferences from 1936-7 and their multiplicative interaction (both variables are standardized as deviations from their mean

for easier interpretation). The model also includes anti-union preferences from 1936-7 and their interaction with employment dispersion. This allows us to check if the effect of concentration preferences is indeed distinct from that of anti-union sentiment more broadly. As most of the variation in union concentration is between districts, we focus on period averages (increasing the number of observations by using congressional terms does not change the results).

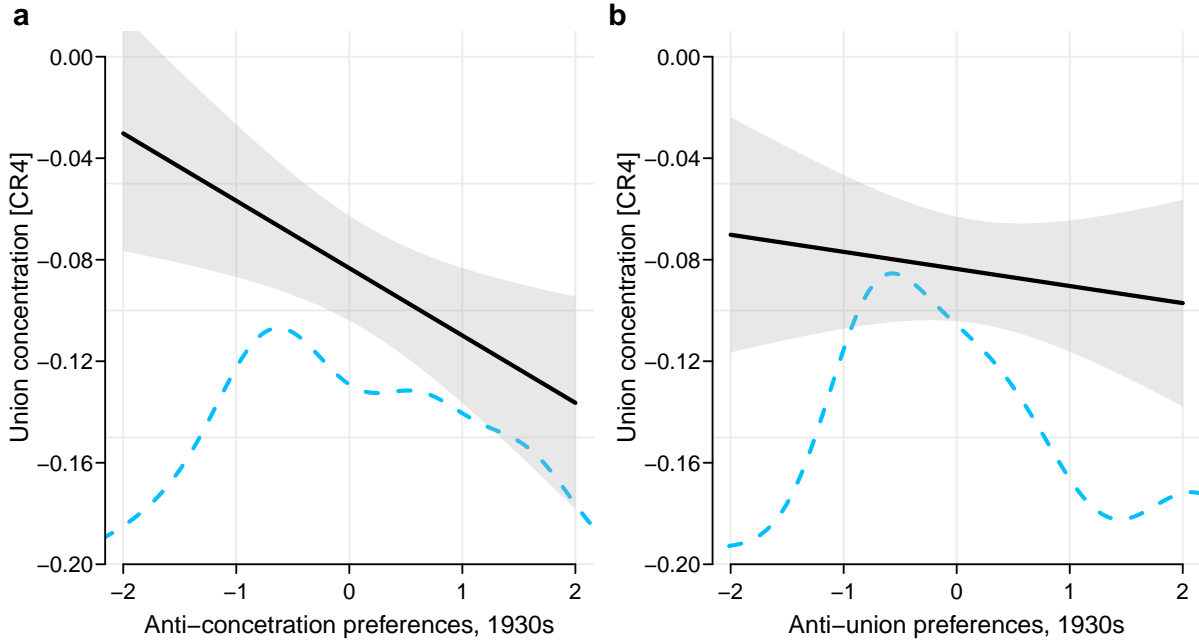


FIGURE A.4.2

#### Historical preferences and contemporary union concentration.

This figure plots the marginal effect of contemporary employment dispersion on union concentration evaluated at levels of historical anti-concentration preferences (panel a) and anti-union preferences (panel b). Both historical preferences are standardized to mean zero with unit standard deviation; the dashed blue lines show their density.

Consistent with the hypothesis, the estimates reported in Table A.4.1 show that the link between employment dispersion and union concentration is moderated by historical anti-concentration preferences. The negative relationship between employment dispersion and union concentration is more pronounced in states with a higher share of anti-concentration sentiment in the 1930s. In contrast, there is no evidence that anti-union preferences have the same effect. The slope on the interaction term less than half the size and not statistically significant. Figure A.4.2 clarifies the respective marginal effects implied by the model. The result is robust across various specifications, which sequentially add state-level and district-level controls. At the state level, we control for other historical factors such as slavery (data is missing for three cases), which have been linked to union development (Schlozman 2015; Lichtenstein 2013), and post-war union policies (right to work laws and collective bargaining



rules). At the district-level, we control for a summary measure of income, education, share of black, service sector employment, and urbanity.

In sum, this section provides a first step in explaining variation in union concentration based on the interaction of economic structure and historical preferences for union organization. It also helps to interpret the main empirical results in the paper, by clarifying that variation union concentration can in part be traced to preferences that are not hardwired into the economy.

## A.5 Double-post-selection LASSO estimation

To relax OLS modeling assumptions, we report model 5 in Table I using a double-post-selection estimator proposed by Belloni et al. (Belloni et al. 2013, 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 (we omit the notation of time and state fixed effects for an easier exposition):

$$y_i = u_i\alpha + g(z_i) + \zeta_i, \quad E(\zeta_i|z_i, u_i) = 0. \quad (\text{A.5.1})$$

Here,  $y_i$  is the DW-Nominate score of a representative in a given district,  $u_i$  is the level of union concentration, and  $\alpha$  captures its effect on legislative ideology. The function  $g(z_i)$  captures the possibly high-dimensional and nonlinear influence of confounders. 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 concentration) and confounders. Therefore, we consider the following auxiliary treatment equation

$$u_i = m(z_i) + v_i, \quad E(v_i|z_i) = 0, \quad (\text{A.5.2})$$

which relates treatment to a covariates  $z_i$ . The function  $m(z_i)$  summarizes the confounding effect and creates omitted variable bias.

The next step is to create approximations to both  $g(\cdot)$  and  $m(\cdot)$  by including a large number ( $p$ ) of control terms  $x_i = P(z_i) \in \mathbb{R}^p$ . These control terms can be transforms of covariates, higher order interaction terms, products of covariates with state-specific time trends 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 et al. 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):

$$y_i = u_i\alpha + x_i'\beta_{g0} + r_{gi} + \zeta_i \quad (\text{A.5.3})$$

$$u_i = x_i'\beta_{m0} + r_{mi} + v_i \quad (\text{A.5.4})$$

Here,  $r_{gi}$  and  $r_{mi}$  are approximation errors (or residuals).

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  $x$  variables in equation (A.5.3) fails. It relies on perfect model selection and can lead to biased inferences and

misleading confidence intervals (see Leeb and Pötscher 2008). Thus, we express our problem as one of prediction by substituting the auxiliary treatment equation (A.5.4) for  $u_i$  in (A.5.3) yielding the reduced form

$$y_i = x_i'(\alpha\beta_{m0} + \beta_{g0}) + (\alpha r_{mi} + r_{gi}) + (\alpha v_i + \zeta_i) \quad (\text{A.5.5})$$

$$= x_i'\bar{\beta}_0 + \bar{r}_i + \epsilon_i \quad (\text{A.5.6})$$

$$u_i = x_i'\beta_{m0} + r_{mi} + v_i, \quad (\text{A.5.7})$$

where  $\bar{r}_i$  is a composite approximation error (cf. Belloni et al. 2013). Both equations in this system represent predictive relationships and are thus amenable to high-dimensional selection techniques.

Note that using this two 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  $x$ , 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  $\beta_{m0}$ , can lead to large omitted variable bias in the estimate of  $\alpha$  even when the size of their coefficient in  $\bar{\beta}_0$  is moderate. The Post-double selection estimator suggested by Belloni et al. (2013) addresses this problem, by basing selection on *both* reduced form equations. Let  $\hat{I}_1$  be the  $y$ -control set selected by LASSO of  $y_i$  on  $x_i$  in (A.5.6); and let  $\hat{I}_2$  be the  $d$ -control set selected by LASSO of  $u_i$  on  $x_i$  in (A.5.7). Then, parameter estimates for the effect of union concentration,  $\alpha$ , and the regularized control set,  $\beta$ , are obtained by

$$(\check{\alpha}, \check{\beta}) = \underset{\alpha \in \mathbb{R}, \beta \in \mathbb{R}^p}{\operatorname{argmin}} \left\{ \frac{1}{n} \sum_{i=1}^n [(y_i - u_i\alpha - x_i'\beta)^2] : \beta_j = 0, \forall j \notin \hat{I} = \hat{I}_1 \cup \hat{I}_2 \right\}, \quad (\text{A.5.8})$$

where  $\hat{I}$  is the union of the  $d$  and  $y$  control sets. This estimator has low bias and yields accurate confidence intervals even under moderate selection mistakes (Belloni and Chernozhukov 2009; Belloni et al. 2014).<sup>3</sup> Responsible for this robustness is the indirect LASSO step selecting the  $d$ -control set in (A.5.7). 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  $u_i$  and  $y_i$ , which decidedly limits the scope of omitted variable bias (Chernozhukov et al. 2015).

---

<sup>3</sup>For a very general discussion see Belloni et al. (2017).

## A.6 Additional robustness tests

We present a number of robustness tests in Table A.6.1.

*Redistricting.* In specification (1) we exclude districts in Georgia and Texas affected by court-ordered redistricting. We find that the removal of 53 observations affected by this has no substantive influence on our results.

*108th Congress.* In specification (2) we include the 108th Congress. It was excluded from our analysis in the main text because several control variables we obtain from the American Community Survey are not available for that time period. To include the 108th congress, we impute these missing observations using state-specific linear and quadratic time trends. This specification shows little change in results compared to our baseline model.

*AFL-CIO scores.* In specification (3), we use a different dependent variable: pro-union voting scores for each legislator as provided by the AFL-CIO. We reverse the scale of this variable so that higher values indicate more “conservative” (anti-union) positions. Using this alternative measure we also find clear evidence for the role of union concentration and union membership.

*Share of union members.* In specification (4), we substitute a different measure for (logged) union membership: the share of union members among the total population in a district (calculated using ACS population estimates for each district). It confirms our results.

*Outliers.* In specification (5), we examine if our results are sensitive to influential cases (“outliers”). We calculate the *dfbeta* influence statistic for each district-year and variable. We find rather small influence statistics for all cases. Nonetheless, we exclude 27 cases, whose influence statistic was larger than a threshold of 0.05. The findings are robust to excluding these observations.

*Bootstrap.* Finally, specification (6) replaces our analytical cluster-robust standard errors with a cluster bootstrap and a wild cluster bootstrap (Cameron et al. 2008), which are more robust under small number of clusters. Again, our central results are robust.

TABLE A.6.1  
Additional robustness tests

	Members	Concentration
(1) Redistricting	−0.078 (0.023)	0.346 (0.124)
(2) 108th Congress included	−0.098 (0.022)	0.307 (0.111)
(3) LHS: AFL-CIO voting scores	−0.059 (0.015)	0.289 (0.097)
(4) Share of union members	−0.659 (0.107)	0.450 (0.131)
(5) Influential cases	−0.083 (0.018)	0.461 (0.119)
(6) Bootstrap confidence intervals		
Cluster bootstrap	−0.075 [−0.12, −0.03]	0.368 [0.13, 0.61]
Wild cluster bootstrap	−0.075 [−0.14, −0.01]	0.364 [0.09, 0.65]

Note: OLS with standard errors robust to heteroscedasticity and within-state correlation (specifications 0 to 5). Specification (6) uses bootstrap 95% confidence intervals based on 1,000 replicates via a non-parametric bootstrap and a wild cluster bootstrap (Cameron et al. 2008).

## A.7 The structure of union membership

In this section, we consider two alternative ways to capture the structure of union membership. First, we replace our measure of union concentration (CR4) with a measure of the *effective number of unions*. It follows the same idea as the effective number of parties proposed by Laakso and Taagepera (1979), and counts the number of unions while weighting them by their membership size. Thus, for each district, we calculate:

$$N = \left( \sum_{i=1}^n u_i^2 \right)^{-1}, \quad (\text{A.7.1})$$

where  $n$  is the number of unions with at least one member and  $u_i$  is the fractional share of members of the  $i$ -th union. Table A.7.1 shows how the effective number of unions influences legislative ideology. To make the direction of this measure comparable, we reverse it so that higher values indicate a lower effective number of unions. We find a significant relationship between the effective number of unions and legislative ideology. A standard deviation decline in the number of effective unions is associated with a 0.08 ( $\pm 0.02$ ) unit increase in

right-leaning ideology. Note that this effect is virtually of the same magnitude as the one based on the CR4 concentration ratio.

TABLE A.7.1  
Alternative ways of capturing the structure of union membership

	Coef.	SD change
(0) Original specification	0.368 (0.116)	0.073 (0.023)
(1) Effective number of unions	0.099 (0.024)	0.077 (0.018)
(2) Ratio concentration to members	0.197 (0.060)	0.077 (0.024)

*Note:* OLS with standard errors robust to heteroscedasticity and within-state correlation.  
Column SD change shows effect of standard deviation increase from median.

Second, as suggested by one of our reviewers, we conceptualize the structure of union membership using a single variable: the ratio of union concentration to union membership numbers in each district. This approach captures the idea that the effect of union membership depends on concentration and vice versa. In particular, the ratio tends to zero as the number of union members becomes very small regardless of their concentration. One drawback is that it is more difficult to interpret. Results for specification (2) in Table A.7.1 show that this ratio also has a significant relationship to legislative ideology. Expressed as change in standardized units, the size of its effect is quite similar to our original specification (using measures of membership numbers and concentration): a standard deviation increase in the ratio of concentration to the number union members increases right-leaning ideology by 0.08 ( $\pm 0.02$ ) points.

## A.8 Interactive specifications

We have argued in the main text that the selection of different types of partisan representatives is one central channel through which local unions can influence policy in Congress. This argument reflects the impact of political polarization and is consistent with recent evidence on whether voters affect or elect policy (e.g., see Lee et al. 2004). Our findings are consistent with this view. They suggest that both union membership and concentration matter for legislative voting in the House. They do so to an important degree by increasing the probability that Democratic rather than Republican politicians are elected. This fact is not lost on unions: over the last decades they have focused more and more attention on electoral mobilization (Lichtenstein 2013; Masters and Delaney 2005). Nonetheless, one may ask whether there is a residual direct effect of union organization on lawmaking, and whether it varies by party (partisanship need not be the only channel of influence).

TABLE A.8.1  
Marginal effects conditional on party membership

	Average marginal effect:	
	Membership	Concentration
Republican	−0.012 (0.011)	0.187 (0.049)
Democrat	−0.012 (0.008)	−0.018 (0.039)
Difference	0.000 (0.012)	0.205 (0.071)

*Note:* Average conditional marginal effects. OLS with standard errors robust to heteroscedasticity and within-state correlation.

Exploring this issue, table A.8.1 shows average marginal effects from interactive regression specifications where both union membership and concentration are interacted with a dummy variable indicating representatives' partisan affiliation. Otherwise the specification is identical to our baseline model (specification 1 of Table I). Our results suggest that there is a significant interaction between partisan affiliation and concentration, but not between partisanship and membership size. Even after accounting for representatives' partisanship, concentration appears to have a direct effect on legislative ideology. This effect is substantively and statistically significant only for Republican representatives, and its direction is the same as the total effect of concentration: higher concentration is linked to more conservative voting scores.

## REFERENCES

- Ahlquist, J. S., A. B. Clayton, and M. Levi (2014). Provoking preferences: Unionization, trade policy, and the ilwu puzzle. *International Organization* 68, 33–75.
- Ahlquist, J. S. and M. Levi (2013). *In the Interests of Others*. Princeton and Oxford: Princeton University Press.
- Belloni, A. and V. Chernozhukov (2009). Least squares after model selection in high-dimensional sparse models. *Bernoulli* 19(2), 521–547.
- 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, 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, and C. B. Hansen (2013). Inference for high-dimensional sparse econometric models. In D. Acemoglu, M. Arellano, and E. Dekel (Eds.), *Advances in Economics and Econometrics: Tenth World Congress*, Volume 3, pp. 245–295. Cambridge: Cambridge University Press.
- Berinsky, A. J. (2006). American public opinion in the 1930s and 1940s. the analysis of quota-controlled sample survey data. *Public Opinion Quarterly* 70(4), 499–529.
- Bonica, A. (2014). Ideology and interests in the political marketplace. *American Journal of Political Science* 57(2), 294–311.
- Cameron, A. C., J. B. Gelbach, and D. L. Miller (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics* 90(3), 414–427.
- Chernozhukov, V., C. Hansen, and M. Spindler (2015). Valid post-selection and post-regularization inference: An elementary, general approach. *Annual Review of Economics* 7(1), 649–688.
- Flavin, P. and M. 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.
- Goplerud, M. (2015). Crossing the boundaries: An implementation of two methods for projecting data across boundary changes. *Political Analysis* 24(1), 121.
- Hirsch, B., D. Macpherson, and W. Vroman (2001). Estimates of union density by state. *Monthly Labor Review* 124(7), 51–55.
- Hirsch, B. T. and D. A. Macpherson (2003). Union membership and coverage database from the current population survey: Note. *Industrial & Labor Relations Review* 56(2), 349–354.
- Holmes, T. J. (1998). The effect of state policies on the location of manufacturing: Evidence from state borders. *Journal of Political Economy* 106(4), 667–705.



- Kim, S. E. and Y. Margalit (2016). Informed preferences? the impact of unions on workers' policy views. *American Journal of Political Science*, (forthcoming).
- Laakso, M. and R. Taagepera (1979). Effective number of parties: A measure with application to west europe. *Comparative Political Studies* 12(1), 3–27.
- Lee, D. S., E. Moretti, and M. J. Butler (2004). Do voters affect or elect policies? evidence from the u. s. house. *Quarterly Journal of Economics* 119(3), 807–859.
- Leeb, H. and B. M. Pötscher (2008). Can one estimate the unconditional distribution of post-model-selection estimators? *Econometric Theory* 24(2), 338–376.
- Lichtenstein, N. (2013). *State of the Union: A Century of American Labor* (2 ed.). Princeton and Oxford: Princeton University Press.
- Masters, M. F. and J. T. Delaney (2005). Organized labor's political scorecard. *Journal of Labor Research* 26(3), 365–392.
- Mennis, J. and T. Hultgren (2006). Intelligent dasymetric mapping and its application to areal interpolation. *Cartography and Geographic Information Science* 33(3), 179–194.
- Mitchener, K. J. and I. W. McLean (2003). The productivity of us states since 1880. *Journal of Economic Growth* 8(1), 73–114.
- Poole, K. T. and H. Rosenthal (1997). *Congress: A Political-Economic History of Roll Call Voting*. New York: Oxford University Press.
- Ratkovic, M. and D. Tingley (2017). Sparse estimation and uncertainty with application to subgroup analysis. *Political Analysis* 25(1), 1–40.
- Schlozman, D. (2015). *When Movements Anchor Parties*. Princeton and Oxford: Princeton University Press.
- Southworth, C. and J. Stepan-Norris (2009). American trade unions and data limitations: A new agenda for labor studies. *Annual Review of Sociology* 35, 297–320.
- Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society B* 58(1), 267–288.
- Wallerstein, M. and B. Western (2000). Unions in decline? what has changed and why. *Annual Review of Political Science* 3, 355–377.