The Effect of Union Membership on Turnout*

Michael Becher†
Daniel Stegmueller‡

This version: October, 2016

Abstract

Do labor unions have a causal effect on voter turnout? Economic theory suggests that union membership and voting are shaped by the same (unobserved) factors, such as ability. To disentangle the union effect from selection, we use data from the U.S. National Longitudinal Survey of Youth. It allows us to specify a latent factor potential outcome model with matching on both observable and unobservable individual characteristics. We find that in the 2006 mid-term election about one-third of the observed union turnout premium is due to selection. Accounting for this, union members are 10 percentage points more likely to vote.

JEL classification: D72; J51

Keywords: Voting; Trade union; Political process

---

*We are grateful to helpful comments and suggestions from Michael Donnelly, Thomas Gschwend, Kyle L. Saunders, Roland Zullo and participants at the annual meeting of the Midwest Political Science Association.
†University of Konstanz, michael.becher@uni-konstanz.de
‡Duke University, daniel.stegmueller@duke.edu

I. INTRODUCTION

Many economists and political scientists believe that the political power of labor unions is based to an important degree on unions' ability to increase voter turnout among their members, thereby influencing election outcomes and a broad range of economic policies (Anzia 2011; Freeman and Medoff 1984; Masters and Delaney 2005; Schlozman, Verba, and Brady 2012; Uhlaner 1989). Consistent with this view, multiple studies have documented a positive relationship between individual union membership and political participation in many contemporary democracies (Delaney, Masters, and Schwochau 1988; Freeman 2003; Flavin and Radcliff 2011; Leighley and Nagler 2007; Norris 2002; Rosenfeld 2014). However, standard economic theory suggests that union membership is influenced by the same (unobserved) personal characteristics that influence voter turnout. In particular, individual ability should matter for both selection into union jobs and the decision to turn out to vote. While selection into groups based on normative or political motivations is a well-recognized identification problem in the turnout literature, selection based on sorting in the labor market has been neglected. In this paper, we use data from the U.S. National Longitudinal Study of Youth (NLSY) to demonstrate the empirical relevance of the economic endogeneity problem. Exploiting the unique features of the NLSY, we implement an empirical strategy that allows us to nonetheless estimate the causal effect of union membership on turnout under comparatively weak assumptions.

We show empirically that union members in the NLSY are characterized by higher levels of cognitive ability than non-members with the same socio-demographic profile. Research in labor economics suggests that this ability gap is the result of sorting on both the employer and employee side. The union wage gap induces more workers to apply for unionized jobs. Employers faced with union wages above the competitive wage and collective bargaining agreements that restrict firing have strong incentives to screen job applicants. The result, then, is sorting on ability between union and nonunion workers (Freeman and Medoff 1984: 45; Robinson 1989: 643). Research in political science has established a link between cognitive ability and electoral turnout. Individuals of higher cognitive capacity are more likely to engage with the political sphere and its abstract concepts and symbols. Ability has been shown to influence individual turnout propensities in behavioral studies using survey data, as well as in genetic studies using twin data (Dawes et al. 2015; Denny and Doyle 2008; Hauser 2000; Luskin 1990; Nie, Junn, and Stehlik-Barry 1996; Verba, Schlozman, and Brady 1995). Furthermore, ability shapes turnout indirectly, since it influences factors closely related to turnout, such as education, and political interest and sophistication (Denny and Doyle 2008: 294).

The imbalance in ability between union members and non-members is usually not observed by researchers studying union membership and turnout. National election surveys simply do not include the required data. The result is an endogeneity problem that makes it difficult to assess the micro-foundations of the mobilizing effect of organized labor in the electoral arena.[1] The problem also exists where jobs are tied to becoming a union member (or at least paying a union fee). This includes a majority of U.S. states without "right-to-work" legislation. While such union shops may mitigate endogeneity concerns based on explicit political motivations (Kim and Margalit 2016; Rosenfeld 2014), they do not rule out the more subtle but important problem based on economic selection.

In contrast to previous studies, our analysis explicitly models selection of individuals into a unionized job and their decision vote on election day as function of both *observable* and *unobservable* characteristics. Following recent advances in the analysis of treatment effects using observational data (cf. Abbring and Heckman 2007), the rich individual-level data of the NLSY allow us to exploit different sources of causal identification and impose weaker assumption compared to approaches exclusively relying on control variables or standard instrumental variables regression. In line with standard labor economics, our empirical approach captures employers' incentives to carefully screen candidates for more costly unionized jobs, as well as employee's economic incentives to obtain a union job. In particular, we specify and analyze a latent factor potential outcome model of union membership and turnout in the 2006 congressional election. The model exploits three distinct sources of causal identification: explicit economic incentives to become (or remain) a union member, a latent variable structure that allows for unobserved (by the researcher) correlations between sorting into union membership and voting, and high-quality cognitive tests that vary independently of treatment status. To capture non-political incentives, we match individual records with industry data, from which we compute instruments for the economic incentives of obtaining a union job: the wage differential between members and non-members in a particular industry, and the level of concentration in a given industry. In addition, the model captures selection bias due to endogenous union membership via a latent factor structure that allows the unobserved factors driving union membership and the decision to vote to be correlated. To identify this latent factor, we exploit the unique data structure of the NLSY and use a measurement system, which is not subject to selection bias. An extensive battery of cognitive performance tests (conducted before respondents' entry into the labor market) allows us to measure an underlying latent variable, cognitive ability, that comprises part of the (otherwise) unobservable latent factor.

We find that there is sorting into unions by individuals with higher ability, who are also more likely to vote. Selection accounts for about one third and perhaps as much as one half of the observed descriptive difference in turnout between members and non-members. Accounting for the selection process, there is robust evidence that unions increase the propensity of their members to participate in elections. On average, union membership increases the probability of an individual to vote by about 10 percentage points – a politically significant effect. In sum,

[1] The common practice of controlling for political attitudes and other proximate predictors of turnout invites post-treatment bias.

our analysis demonstrates that prospective voters are more likely to be selected into unions, but it also demonstrates that, after accounting for selection, unions effectively encourage voting.

This paper contributes to the literature of union effects on political behavior. Due to data limitations, most studies rely on cross-sectional regression analysis with covariate adjustment. A few recent exceptions exploit quasi-random variation to estimate union effects. Kim and Margalit (2016) uses an innovative survey of union workers that allows for matching by industry and exploit a shift in the position of a national union to estimate the effect of union membership on trade policy preferences. Similarly, Ahlquist, Clayton, and Levi (2014) examine the effect of unions on trade policy preferences by studying dock workers within the same industry across unions with different policy positions. Taking a different approach, we provide what we think is the first study of the union-turnout link that jointly models selection into union membership and turnout based on observable as well as unobservable factors. Focusing on voting rather than policy preferences, we find that the selection effect for turnout is larger than the one found for trade preferences by Kim and Margalit (2016).

Our approach and findings are also relevant for the general literature on turnout. Recent formal models have turned their attention to group dynamics to explain voting in large elections (Feddersen and Sandroni 2006; Uhlaner 1989). The electoral role of labor unions is one central motivation for this line of research (Uhlaner 1989). The fundamental challenge facing group-based explanations of voting is to account for why individuals join certain groups in the first place (Feddersen 2004). Our empirical approach tackles the problem and quantifies the magnitude of the selection problem for a large group. It also illustrates the relevance of economic theory for identifying potential sources of selection bias and thereby providing a clear basis for empirical research. One may hope that the dominance of economic incentives in the context of American labor law imply that endogeneity is not a problem for the study of voting (Kerrissey and Schofer 2013: 918; Rosenfeld 2014: 145). This paper shows otherwise. Economic selection implies that union members and non-members differ systematically on fundamental determinants of political behavior.

II. ENDOGENOUS UNION MEMBERSHIP AND VOTING

To motivate the empirical analysis to follow, let us consider the identification problem a bit more carefully. Sorting may cause prospective voters to be more likely to become union members than prospective non-voters – even in the context of union shops common in many U.S. states.

In any given election, we can only ever observe an individual in one of two possible states: being a member of a trade union or not. Thus, the propensity to turn out on election day for each individual is a *potential* outcome (Rubin 1978). Denote the two potential outcomes (turnout propensity) in our two counterfactual states (union member, non-member) by Y_{1i}

and Y_{0i}. For each individual, we assume that the pair (Y_{0i}, Y_{1i}) exists, but we can only ever observe one possible state per individual so that $Y_i = D_i Y_{1i} + (1 - D_i) Y_{0i}$. The core quantity of interest of this paper, the effect of union membership on turnout, is the average treatment effect $\Delta = E(Y_{1i} - Y_{0i})$. This is the *ceteris paribus* effect on turnout of moving an otherwise identical individual into union membership. Thus, for each union membership state, $D_i = (0, 1)$ we need to identify the potential outcome in the alternative state. This *counterfactual* outcome is unobserved. Both potential outcomes are a (possibly non-linear) function of a vector of observed individual characteristics, $\mu(X_i)$, such as age or education. Furthermore, we need to account for unobserved confounders. The influence of such unobservables is captured by including individual random variables U_i. This yields the following two potential outcome equations:

$$\begin{aligned} Y_{0i} &= \mu_0(X_i) + U_{0i} \qquad \text{if } D_i = 0 \\ Y_{1i} &= \mu_1(X_i) + U_{1i} \qquad \text{if } D_i = 1. \end{aligned} \tag{1}$$

Crucially, unobservables U_{0i} and U_{1i} may be correlated with unobservable factors explaining sorting into union membership. We can think of sorting into D_i as a function of observable individual characteristics, Z_i, and unobservables, U_{Di}: $D_i = \mu_D(Z_i) + U_{Di}$. Unobserved individual characteristics affecting union membership and turnout are collected in the random vector $(U_{Di}, U_{1i}, U_{0i})'$. Since factors influencing union membership also shape potential (turnout) outcomes, one has to allow for correlations between all unobservables. For a sample of individuals, this yields the following 3×3 variance-covariance matrix:

$$\text{Cov}\begin{bmatrix} U_D \\ U_1 \\ U_0 \end{bmatrix} = \begin{bmatrix} \sigma_D^2 & \rho_{D1}\sigma_D\sigma_1 & \rho_{D0}\sigma_D\sigma_0 \\ \rho_{D1}\sigma_D\sigma_1 & \sigma_1^2 & \rho_{10}\sigma_1\sigma_0 \\ \rho_{D0}\sigma_D\sigma_0 & \rho_{10}\sigma_1\sigma_0 & \sigma_0^2 \end{bmatrix}. \tag{2}$$

Its diagonal entries represent the variances of unobservables in union and turnout equations. Off-diagonal entries capture the relationship between unobservables in turnout and union membership. Here, ρ_{jk} parametrizes the correlation between U_j, U_k, e.g., ρ_{D0} represents the correlation between unobservables affecting union membership and unobservables affecting turnout of non-union members. Since we can never observe the same individual in two treatment states at once, the correlation between both potential outcomes, ρ_{10}, is not identified (Vijverberg 1993: 74). This is the root of the well-known fundamental problem of causal inference (Holland 1986: 947).

Why would there be sorting on unobservables into unions that also affect voting? Standard arguments in labor economics imply that there is sorting into union jobs based on ability (related to expected productivity) as long as union jobs provide higher wages, benefits, or job security (Freeman and Medoff 1984: 45; Robinson 1989: 643).[2] On the worker side,

[2] For a textbook discussion, see Borjas (2013: 444).

better wages or benefits induce more workers to apply for a unionized job, increasing the pool of attractive candidates. Even if employers are unwilling to screen their job applicants, there will be sorting if higher ability types have a higher reservation wage (e.g., better outside options in self-employment). Employers, on the other hand, are faced with union wages above the competitive wage and collective bargaining agreements that make it more difficult to lay off unionized workers (e.g., by enforcing seniority rules, Abraham and Medoff 1984). This produces incentives to screen job applicants for ability. As a consequence, the prediction is that we should observe sorting on ability between union and nonunion workers.

Cognitive ability is also a fundamental trait discussed in the literature on turnout but not included in most nationally representative election surveys (Luskin 1990; Verba, Schlozman, and Brady 1995; Nie, Junn, and Stehlik-Barry 1996; Hauser 2000). It systematically influences turnout by shaping education, civic skills, and political interest and sophistication (Denny and Doyle 2008: 294). Sorting in the labor market is not perfect because screening is costly and job matching is probabilistic. But given that ability is also a fundamental determinant of political participation, a correlation between union membership and ability leads to an endogeneity problem for empirical research. While the existence of union shops reduces concerns about sorting based on adherence to civic norms that are frequently voiced in the empirical literature and can be derived from theories of social customs (Akerlof 1980; Corneo 1997)[3], it does not block the economic sorting mechanism.

To address this endogeneity problem, we draw on a growing literature in econometrics that extends the potential outcomes framework for causal inference to non-random treatment assignment with complex data structures, which can provide additional identifying information (Abbring and Heckman 2007: 5166). We use a latent factor potential outcomes setup (see, e.g., Aakvik, Heckman, and Vytlacil 2005; Heckman, Lopes, and Piatek 2013) to allow for a correlation between unobservables influencing union membership and voting (i.e., $\rho_{jk} \neq 0$). Crucially, our model allows us to exploit additional information to impose less restrictive identification assumptions. For our problem, the use of cognitive tests serves this purpose. It captures the theoretical intuition, discussed above, that labor market sorting leads to a positive correlation between ability and union membership. Cognitive tests are not employed as an instrumental variable. Loosely speaking (section iv is more formal), they serve as a proxy for unobserved traits relevant for both union membership and political participation that helps to pin down the structure of unobserved heterogeneity.[4]

These two sources of identification complement the more traditional approach (which we also follow) of trying to find plausible instrumental variables for an endogenous variable

[3]The structure of normative models of union membership is similar to group-based turnout models: Only individuals with a cost of participation below a threshold will join or vote.

[4]Over the last twenty years the use of randomized field experiments has revolutionized the study of electoral mobilization (Gerber and Green 2000), but ethical and practical issues have by and large precluded their application to the central question of group membership and voting.

like union membership. One can think of two sets of variables shaping sorting into unions, $Z_i = (X_i, W_i)$. First, X_i contains basic individual background characteristics (or pre-treatment covariates), such as education, age, and gender. Second, W_i contains one or more variables which make union membership an economically attractive option, but are unrelated to election turnout, given X_i. These serve as instrumental variables. Intuitively, the higher the union wage or benefit premium, the more likely it is that we find a worker being a union member (Schnabel 2003: 14).[5] A second factor influencing union membership is industry concentration, which has been found to be a relevant determinant of unionization levels (i.e., due to economies of scale union organization is easier in sectors with four firms than with 50; Stephens and Wallerstein 1991: 943) and the resulting benefits for workers (Hirsch and Berger 1984). Higher levels of industrial concentration are connected with higher wages for union members (Kwoka 1983) as well as higher provision of fringe benefits (Freeman 1980). Net of observed and unobserved worker characteristics, industry concentration is predicted to encourage unionization.[6]

We discuss how we transform this conceptual framework into an empirically estimable setup below, after describing the unique data set that enables our analysis.

III. DATA

We use the widely used National Longitudinal Study of Youth (NLSY), a longitudinal panel study directed by the U.S. Department of Labor's Bureau of Labor Statistics.[7] Its widespread use is due to the high quality of its sample design, data collection, and the availability of cognitive measurements (e.g., Lang and Manove 2011). Due to its mission the NLSY does not include political questions. However, in an exceptional collaboration the American National Election Study (ANES) was able to place a short set of political items in the NLSY 2008 wave, including the turnout question asked in each ANES survey (Krosnick and Lupia 2006). We make use of this unique data-set (which includes rich information on individuals) to study the effect of union membership on turnout.

The key design characteristic of the NLSY is that it is a nationally representative sample of certain birth cohorts. Currently there are two NLSY panels: a recent one started in 1997, comprised of cohorts born between 1980 and 1984, and a long-run panel started in 1979, which is made up of cohorts born between January 1, 1957, and December 31, 1964 (and who resided in the US in 1979). We use the latter for our analysis, since it focuses on individuals

[5] For instance, wage differentials may be due to monopoly power, firm-worker matching on productivity, or represent compensatory payments for work conditions.

[6] If these instruments are valid, the structure used in our model does not require more stringent assumptions than those imposed in the LATE framework of Imbens and Angrist (1994). See Vytlacil (2002).

[7] The NLSY79 survey is sponsored and directed by the U.S. Bureau of Labor Statistics and conducted by the Center for Human Resource Research at The Ohio State University. Interviews are conducted by the National Opinion Research Center at the University of Chicago. See www.bls.gov/nls for more details.

who participated in the labor market for a substantial number of years. Due to the cohort design of the NLSY, they are between 41 and 50 years old in 2006. We focus on male respondents only, in order to work with a sample from a population generated by a reasonably homogeneous data generating process. A complete analysis of women's union membership would have to include an explicit model of their decision to participate in the labor market, which is beyond the scope of this paper.[8] This yields a sample size of 2,460 respondents. We match each individual in this micro-data set with industry characteristics (industry concentration and industry union-nonunion wage differentials) calculated from administrative data sources.

Industry characteristics. Industrial concentration has a long history in economics. It is usually measured by the ratio of the combined market share of the four largest firms to the whole market size of that industry (Pryor 1972), the so called CR_4 concentration ratio. We use concentration ratios based on the Census Bureau's Economic Census of American businesses in 2007.[9] The high quality of this data allows us to use detailed disaggregated concentration ratios for 243 industries.

We calculate union-nonunion wage differentials from the Bureau of Labor Statistics' CPS-LU series, which, based on Current Population Survey data, provides wages for workers (not) covered by union wage contracts. To avoid small sample bias, we use a lower resolution than for our concentration measure and calculate wage differentials for 19 major industrial sectors. Our measure of union-nonunion wage differentials is the difference in median weekly earnings of full-time employed union members and non-members at the 2-digit industry level.

Figure 1 plots the distribution of industrial concentration and union wage differentials.[10] It shows the existence of substantial variation in both variables. The market share held by the four largest firms ranges from less than 10 percent in some industries to over 80 in others. For example, within the non-durable goods manufacturing sector, the four largest firms in the textile mills industry hold only about 16 percent of the total market share, while tire production is highly concentrated, with over 70 percent in the hands of the four largest firms. Wage differentials are similarly spread out. In the majority of industries union differentials are around 100 dollars a week (for example, in health care and social assistance), but they range from −100 (in finance and insurance) to almost 350 dollars (in construction).

Cognitive ability tests. In 1980, the Department of Defense and the Department of Labor jointly sponsored the administration of the Armed Services Vocational Aptitude Battery (ASVAB, cf. Jensen 1985) to the civilian and military NLSY79 samples.[11] The ASVAB consists of several

[8] We also exclude NLSY's military subsample, as members of the military are not union members.

[9] The Economic Census is conducted in 5-year intervals and samples around 4 million firms. Participation is required by law.

[10] Appendix A contains more detailed descriptive information on these two variables.

[11] The DoD uses a subset of the ASVAB to create an Armed Forces Qualifications Test score (AFQT) as a general measure of trainability used in Armed Forces enlistment.

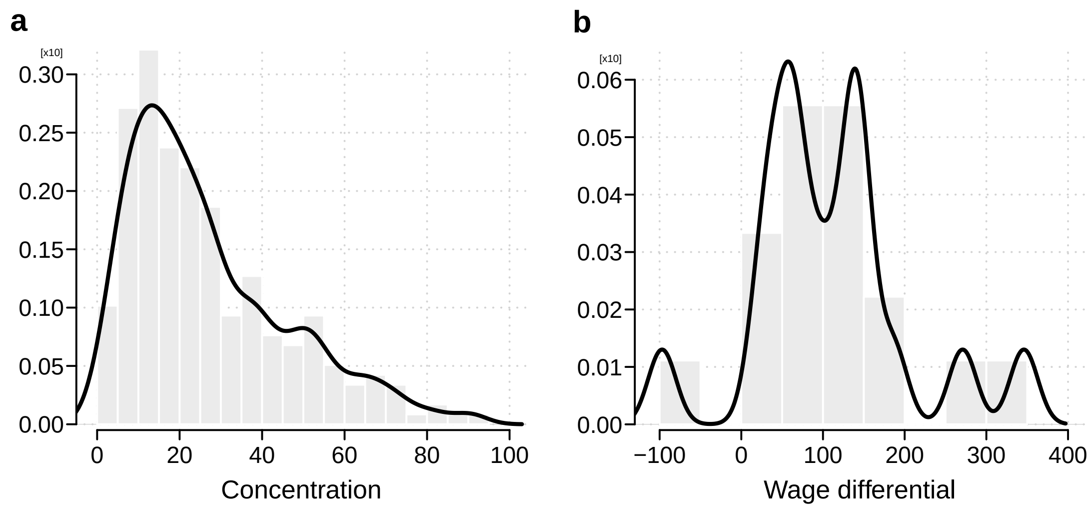

Figure 1: Distribution of industry characteristics. Panel (a) shows levels of industrial concentration measured via the 4-firm concentration ratio , panel (b) shows union wage differentials measured via differences in median weekly earnings (in $) between union members and non-members.

subtests that measure aptitude in areas such as arithmetic reasoning, coding speed, mathematics, and word knowledge. We follow recent innovations in the economics literature and construct a measurement model, which posits an underlying latent variable—cognitive ability—that produces observed test scores, thus accounting for the fact that an individual's scores on a test and his or her general cognitive ability are not the same thing.

Turnout. After the November election in 2006 respondents were queried if they voted.[12] To reduce over-reporting respondents had several options to indicate abstention: "I did not vote in the November 2006 election", "I thought about voting in 2006, but didn't", "I usually vote, but didn't in 2006".[13] Turnout was indicated by the response option "I am sure I voted". We create an indicator variable equal to one if a respondent chose this option and zero otherwise.

Controls. We include a number of additional variables to capture heterogeneity between individuals. A respondent's income is measured as total wage and salary income before taxes

[12] The text of this questions reads: "In talking to people about elections, we often find that a lot of people were not able to vote because they were sick or they just didn't have time or for some other reason. Which of the following statements best describes you?"

[13] Over-reporting of turnout is a well-known problem. While we argue that over-reporting *per se* is not necessarily a problem for our inferences (because the model works with differences in turnout outcomes), we compared our data to the American National Election Study. Mean turnout in our data set (for union members and non-members combined) is 62.8%. This is at the lower end of the 95% confidence bound of turnout among the same age group obtained from the "gold standard" ANES (Aldrich and McGraw 2011), which ranges from 61 to 81%.

Table 1: Sample characteristics. Means and standard errors.

	Union members [N=456]		Non-members [N=2004]	
Turnout	0.73	(0.02)	0.59	(0.01)
Income [1000$]	52.98	(1.44)	50.69	(1.22)
Education	13.11	(0.10)	13.26	(0.06)
Black	0.31	(0.02)	0.31	(0.01)
Hispanic	0.21	(0.02)	0.19	(0.01)
Family size	2.96	(0.07)	2.83	(0.03)
Married	0.65	(0.02)	0.58	(0.01)
Unemployment exp.[a]	0.06	(0.01)	0.12	(0.01)
Rural area	0.78	(0.02)	0.73	(0.01)
South	0.22	(0.02)	0.45	(0.01)
Industry concentration[b]	28.58	(1.08)	23.21	(0.47)
Wage differential[c]	150.17	(4.14)	128.23	(2.46)

Note: Cohort dummies and variables in test equations not shown to save space. Details on the construction of all variables are available in appendix A.
[a] Unemployed in past calendar year (indicator variable).
[b] 4-firm concentration ratio CR_4.
[c] Difference in median weekly earnings (in $) of full-time employed union members and non-members at 2-digit industry level.

and deductions. Education is captured by years of schooling. Besides accounting for family size, we include indicator variables for being married, unemployment spells in the previous calendar year, living in a rural area, and a Southern state dummy. In order to capture well-known turnout differences of minority groups, we also include indicator variables for Black and Hispanic. To account for the cohort design of the NLSY, we also create a set of dummy variables to capture systematic cohort differences in unobservables.

Table 1 provides descriptive means of our central variables for union members and non-members. Most notably, the (unadjusted) difference in turnout is 14 percentage points.[14] In terms of observables, union members experience fewer unemployment spells and are less likely to be from the South, where right-to-work legislation is predominant. Most notably, union members work in industries that are more concentrated and characterized by larger wage differentials.

[14]This is nearly identical to the unadjusted difference in the employed population for mid-term elections based on CPS data: 13 percentage points (Freeman 2003). Averaging across all elections between 1984-2008, the difference is closer to 10 pp (Rosenfeld 2014).

IV. A LATENT VARIABLE POTENTIAL OUTCOME MODEL

In this section we discuss our solution to the endogeneity problem outlined in section ii. The latent variable potential outcome model we specify exploits the rich information available in our data to add statistical structure allowing us to identify potential outcomes and derive the relevant treatment effects. First, as noted above, a joint model of potential outcomes and union membership does not contain any information about the correlation between potential outcomes, (Y_0, Y_1).[15] What is needed for identification are the joint distributions (U_D, U_0) and (U_D, U_1) of unobservables in treatment and outcome equations (Chib 2007; Heckman 1990). These can be obtained by parameterizing the structure of $\text{Cov}(U_D, U_1, U_0)$ in equation (2) using an underlying low-dimensional set of random factors (cf. Heckman 1981). Thus, following Aakvik, Heckman, and Vytlacil (2005), we decompose unobservables using the following factor structure:

$$U_D = \alpha_D \theta + \epsilon_D \tag{3}$$

$$U_0 = \alpha_0 \theta + \epsilon_0 \tag{4}$$

$$U_1 = \alpha_1 \theta + \epsilon_1 \tag{5}$$

Here, θ is a latent factor or random effect (Cameron and Heckman 1998; Skrondal and Rabe-Hesketh 2004), which represents unobserved individual characteristics, such as cognitive ability, which systematically shape both union membership and the propensity to turn out on election day. Note that θ is allowed to affect union choice and potential outcomes differently. In the current application, we specify $\theta \sim N(0, 1)$—a distributional assumption which is convenient and quite robust against misspecification (cf. Bartholomew 1988; Neuhaus, Hauck, and Kalbfleisch 1992; Wedel and Kamakura 2001).[16]

This structure solves the core identification problem by inducing dependency between potential outcomes and treatment equation (Carneiro, Hansen, and Heckman 2003; Aakvik, Heckman, and Vytlacil 2005). To see this, note that the non-identified parameter, $\rho_{10} \equiv \text{Cov}(Y_0, Y_1)$, can be recovered from the factor coefficients as $\rho_{10} = \alpha_0 \alpha_1$. In other words, the latent factor is assumed to generate the correlation between potential outcomes and treatment choices. Assuming that the factor structure captures a relevant part of unobserved individual characteristics, such as ability, which is approximately normally distributed in the population, the fundamental identification problem is removed. Below we add two more sources of information, providing more robust identification of the effect of union membership on turnout. Before doing so, we detail our specifications of turnout and union membership equations.

---

[15]From here on we assume that we have access to an equiprobability sample and suppress individual subscripts for easier notation.

[16]Note that assuming normality is convenient but not necessary for identification. Cunha, Heckman, and Navarro (2005) and Cunha, Heckman, and Schennach (2010) discuss nonparametric identification of θ and $\epsilon_D, \epsilon_0, \epsilon_1$.

We write sorting into union membership as a latent index model (Heckman and Vytlacil 1999, 2007) with a linear-in-parameters formulation.[17] Covariates are $Z = (X, W)$, comprised of observed individual characteristics (confounders), X, and a set of variables W capturing the benefits of union membership. Variables in W are excluded from turnout equations (more below). They, together with the latent factor θ, shift D^*, the propensity of union membership:

$$D^* = \beta_D' Z + \alpha_D \theta + \epsilon_D. \tag{6}$$

$$D = \mathbf{1}(D^* > 0) \tag{7}$$

Here β_D is a parameter vector associated with covariates and exclusions in Z, while α_D is the latent factor coefficient. Errors ϵ_D are white noise (normalized to have variance one) and assumed independent of Z and θ.

The fact that we have non-political variables W in Z that encourage union membership (ceteris paribus), but which are unrelated to turnout decisions, provides an additional source of identification. Of course, variables for which exclusion restrictions hold are hard to find, and almost always hotly contested. But notice that the validity of this exclusion is not strictly necessary to identify the model (this is achieved by the latent factor structure). However, having a valid instrument means that we are less reliant on the exact functional form of the latent factor θ.[18]

For each potential outcome Y_d $(d = 0, 1)$ we assume that it is generated by an underlying latent outcome Y_d^* (say, propensity to turnout) using the following specification:

$$\begin{aligned} Y_d^* &= \beta_d' X + \alpha_d \theta + \epsilon_d \qquad d = 0, 1 \\ Y_d &= \mathbf{1}(Y_d^* > 0). \end{aligned} \tag{8}$$

In this setup, each latent potential outcome is shaped by observed individual characteristics X and their associated parameter vectors β_d and by the latent factor θ with associated coefficients α_d. Errors ϵ_d are assumed to be independent of X and θ.[19]

---

[17] In other words, we set $\mu(Z) = Z\beta_D$ (for a discussion of linear-in-parameters specifications in latent index models, see Eckstein and Wolpin 1989). Then $D = \mathbf{1}(D^* > 0)$, where $\mathbf{1}(\cdot)$ is an indicator function evaluating to one if its argument is true and zero otherwise.

[18] Heckman (1990) and Heckman and Vytlacil (2007) prove conditions for nonparametric identification of θ when instruments are available.

[19] To be explicit, we employ the following technical assumptions (next to the ones listed in the text). (1) $\mu(Z)$ is a non-degenerate random variable conditional on X, i.e., we have a valid exclusion restriction, such that a variable determines union choice but not turnout. (2) The joint distributions of unobservables (U_D, U_1) and (U_D, U_0) are absolutely continuous (w.r.t. Lebesgue measure on $\mathscr{R}^2$). (3) Independence of covariates, $(U_D, U_1) \perp (X, Z)$ and $(U_D, U_0) \perp (X, Z)$ (a standard instrumental variable assumption, which could be relaxed by conditioning on X). (4) Finally, the existence of treated and untreated individuals for each set of confounders X, $1 > Pr(D = 1|X) > 0$.

So far, our latent factor potential outcome model contains two independent sources of identification, the factor structure θ and instruments in Z. Our third source of more robust model identification is created by utilizing the panel-structure of our data to generate auxiliary information, which is independent of treatment status, and which we use to identify variation in θ. We use a battery of aptitude tests, which proxy individuals' cognitive ability. As we have argued above, ability is an important unobservable that likely affects both turnout and union membership through employer selection or self-selection.

These tests were administered to *all* respondents in our sample when they were young adults. Therefore, test scores vary exogenously, i.e., they are not influenced by sorting into union membership (taking place a decade later). Assuming that observed test scores are systematically related to θ their variation provides an additional source of identification. Technically, we have a measurement system M that is independent of individuals' treatment status D, and which is adjoined to the latent factor θ (Carneiro, Hansen, and Heckman 2003). Our measurement system is comprised of P observed variables, namely several ASVAB measures of ability. Each measurement p ($p = 1, \ldots, P$) is generated by a linear combination of θ and controls, X_p, which we include to account for their possibly biasing influence on observed test scores (such as family background).[20] Random variables ϵ_{M_p} denote idiosyncratic variation in tests scores that are not explained by covariates or the latent factor.[21]

$$M_p = \lambda_p X_p + \alpha_{M_p}\theta + \epsilon_{M_p} \quad p = 1, \ldots, P. \tag{9}$$

Attaching this measurement system to θ achieves three things. First, it provides meaning to the latent factor, i.e., it yields evidence (via statistical tests of α_{M_p}) to what extent θ captures unobserved cognitive ability. Second, it anchors θ and eases its interpretation in relation to a tangible object, such as test performance. Third, it provides more robust identification. Having measurements on θ that are independent of D—i.e., individuals were administered cognitive tests irrespective of future union membership—provides an additional source of identification in the model (cf. Carneiro, Hansen, and Heckman 2003).

Treatment effects Using this statistical structure we can identify our central quantity of interest: the effect of union membership on the probability of turnout. We focus on the average treatment effect, which can be derived from the estimated coefficients of our model (the treatment effect on the treated can be calculated similarly). Appendix C provides a formal characterization.

---

[20] Table B.1 gives an overview of all variables used in choice, turnout, and test equations.

[21] They are assumed to be distributed mean zero with finite variance, $\sigma^2_{M_p}$, and independent of all covariates and the latent factor. This is a standard conditional independence assumption made in measurement models (see e.g. Jackman 2008), stating that conditional on the latent variable errors are independent.

Estimation We estimate the model in a Bayesian framework.[22] A key advantage of the Bayesian approach is that we recover the full posterior distribution of the average treatment effect as part of the model. To resolve the common rotation problem of latent factor models (Anderson and Rubin 1956), we fix α_D to 1.[23] Thus we anchor it to the union membership equation, such that higher values of θ induce union membership. To complete the Bayesian model setup we assign priors to all model parameters. We choose "non-informative" priors so that all inference in our model is dominated by the data. Details on the parametrization of our prior parameter distributions are given in Appendix D. There we also conduct sensitivity checks (Gill 2008: 199f.) to show that our results do not depend on particular prior choices.[24]

V. RESULTS

In this section we provide a detailed discussion of our model estimates. We first discuss our model parameter estimates and then the resulting treatment effect of union membership in subsection B.

A. Parameter estimates

Union membership. Table 2 shows estimates from the union membership equation (6). It shows a summary of the posterior distribution for each parameter—its mean and standard deviation, as well as the 95% highest posterior density (HPD) region, which can be understood as Bayesian analogue to the frequentist confidence interval.[25] For easier interpretation of effect sizes, the final column of Table 2 displays first differences in predicted probabilities. As discussed above, the coefficient of the latent factor θ is normalized to unity in the selection equation. It affects union membership substantially: even after accounting for observable differences between union members (such as being black or living in a rural area), a standard deviation change in θ raises the probability of union membership by 33 percentage points (holding

[22]Bayesian models for potential outcomes are comparatively rare (but see Heckman, Lopes, and Piatek (2013)). Note that the model is identified under classical criteria. Note further, that estimating the system of equations using maximum likelihood (using Gauss-Hermite quadrature to integrate over θ) yields comparable results. In fact, we use ML estimates as starting values for our Gibbs sampler.

[23]The rotation problem of latent factor models occurs since elements of $\alpha = (\alpha_D, \alpha_0, \alpha_1, \alpha_{M1}, \ldots, \alpha_{MP})'$ can switch sign. To see the problem more precisely, let R be a matrix such that $R'R = I$ and note that $\alpha^{(R)} = \alpha R'$. In other words, α is rotation invariant, we obtain the same likelihood when we "flip" it (Anderson and Rubin 1956). There are several solution strategies, and we choose the simplest one, by fixing one α coefficient.

[24]We estimate our model using Markov Chain Monte Carlo simulation. We use data augmentation to sample latent index variables D^*, Y_1^*, and Y_0^* (Albert and Chib 1993). Conditional on samples from these, all other parameters can be sampled via Gibbs sampling steps.

[25]More precisely a region R is a $100(1-\alpha)$ percent HPD region (not necessarily contiguous) for parameter θ if (1) $P(\theta \in R) = 1 - \alpha$ and (2) $P(\theta_1) \geq P(\theta_2)$ for all $\theta_1 \in R$ and $\theta_2 \notin R$, i.e., it yields an interval estimate with the added requirement that each value in the interval is larger than those outside of it.

Table 2: Union membership equation ($D = 1$) parameter estimates.

	Mean	SD	95% HPD		Prob.[a]
Latent factor	1.000[b]				0.332
Industry concentration	0.133	0.024	0.085	0.177	0.032
Union wage diff.	0.195	0.025	0.147	0.243	0.049
Education	−0.476	0.027	−0.530	−0.424	−0.082
Black	1.025	0.061	0.903	1.141	0.134
Hispanic	0.450	0.068	0.309	0.577	0.044
Rural area	0.157	0.059	0.040	0.270	0.016
South	−0.692	0.055	−0.800	−0.584	−0.064

Note: Cohort dummies and intercept not shown
[a] First difference in predicted probability of unit change in Z
[b] Fixed parameter

everything else constant). Conditional on covariates and the latent factor (i.e., unobservables), we find that both of our instruments induce union membership in the expected way. The higher the differential between union and non-union wages, the higher the probability of union membership. Similarly, working in a more highly concentrated industry raises the probability of being a union member by around 3 percentage points. The confidence bounds for both coefficients are far away from zero.

Test scores. Given the clear importance of θ in our model, we now investigate if it captures meaningful differences between individuals. If θ does represent (to some extent) cognitive abilities, we should find that it significantly shapes observed cognitive test scores. Table 3 shows estimates from our measurement system of ASVAB test items, given in equation (9). We find that θ has a substantial influence on achieved test scores. Higher values of θ are associated with higher coding speed, improved arithmetic reasoning, and more knowledge of language and mathematics. To a lesser extent it also influences basic reading comprehension. Inspecting 95% HPD intervals shows that all relationships are highly statistically reliable. The availability of these additional measurements allows us to give meaning to the latent factor in our model. These relationships hold while adjusting for individual background variables, which might bias test results. For example, one would expect that an individual which came from a broken home (defined as living with a single parent), or from a low-resource familial background (as indicated by many siblings or low family income), would do worse on a test (holding all else, including θ, equal). These distorting influences do indeed exist and are accounted for in our measurement model.

Figure 2 plots the distribution of θ for union members and non-members. We construct this plot by drawing 500 samples from the posterior distribution of θ_i, calculate the posterior expectation, $E(\theta_i)$ for each individual, and then calculate a kernel density estimate. Figure 2

Table 3: ASVAB test equations parameter estimates.

	Mean	SD	95% HPD	
Factor effects				
Arithmetic reasoning	6.735	0.097	6.547	6.929
Word knowledge	7.195	0.116	6.956	7.409
Paragraph comprehension	2.892	0.053	2.787	2.991
Coding speed	10.784	0.234	10.304	11.219
Math knowledge	5.406	0.083	5.251	5.575
Test covariates[a]				
Age at test	0.156	0.050	0.061	0.256
Broken family	−0.364	0.103	−0.561	−0.163
Education mother	0.532	0.062	0.418	0.657
Education father	0.314	0.066	0.183	0.441
Number siblings	−0.336	0.052	−0.435	−0.233
Family income	0.232	0.057	0.125	0.350

Note: Intercepts and variances not shown
[a] Covariate effects λ_p set equal across test items. See appendix A for variable definitions.

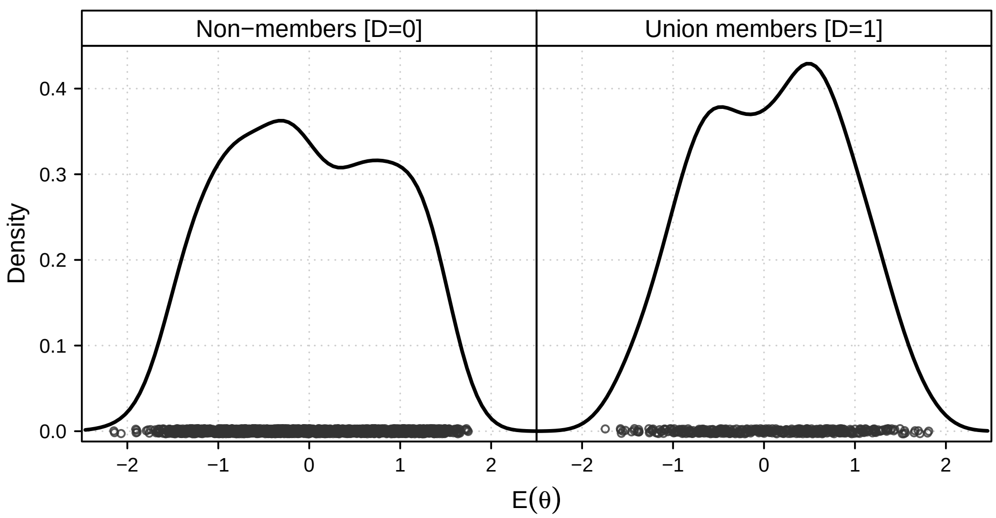

Figure 2: Distribution of latent factor values (ability) by treatment status

indicates that union members differ from non-members in that they have somewhat higher levels of (latent) ability. There is a larger portion of union members with ability above the mean (remember that θ is normalized to zero in the population) than non-members. This point is made more formally in Table 4, which shows estimates of the mean of the latent factor

Table 4: Distribution of θ for union members and non-members.

	Union members		Non-members		Difference	
Mean	0.079	(0.025)	−0.014	(0.021)	0.093	(0.017)
20*th* percentile	−0.702	(0.039)	−0.924	(0.029)	0.221	(0.038)
80*th* percentile	0.842	(0.042)	0.915	(0.029)	−0.073	(0.038)

Note: Uncertainty of estimates in parentheses. Calculated using Monte Carlo integration (2,000 draws from posterior distribution of θ_i).

for union members and non-members, as well as the 20th and 80th quantile. It confirms that union members do, on average, have higher levels of ability than non-members, consistent with the economic argument that employer screening as well as self-selection leads to sorting into higher paid union jobs. It also shows that the distribution is more compressed among union members, i.e., at the 20th percentile of the distribution union members have substantially fewer low θ values than non-members. The same finding obtains (somewhat less pronounced) at the top of the distribution. This result underscores, once again, the importance of accounting for differences in unobservables between individuals.

Turnout. The previous paragraph has shown clear evidence for selection into union membership based on unobservables. If these influence turnout as well, then ignoring them leads to biased inferences. Our setup provides for a straightforward test of this issue: if unobservables driving union membership also influence turnout we will find significant parameter estimates for α_0 and α_1 in both potential outcome equations. Table 5 shows posterior parameter summaries for equations (8), as well as effect sizes via first differences in predicted probabilities. Panel (A) displays coefficient estimates for union members, panel (B) for non-members. We find that θ does indeed substantially affect turnout in both potential outcome states. Both coefficients are of sizeable magnitude and their posterior uncertainty intervals are far away from zero. Thus, we clearly reject the hypothesis that there are no selection effects. The latent factor has a slightly stronger influence on the turnout choice of union members. A standard deviation change raises their probability of turnout by 8.5 percentage points, while the corresponding figure among non-members is 7.5 percentage points.

The role of confounders in Table 5 is as expected from previous research. In particular, higher socio-economic status (income, education) is associated with a higher propensity to vote. Notably, income has a much more pronounced effect on potential outcomes among union members than among non-members. The fact that several more control variables have different effects for union members and non-members underscores the importance of using our flexible model setup.

Table 5: Potential outcome equations parameter estimates

	Mean	SD	95% HPD		Prob.[a]
(A) Union members (Y_1)					
Latent factor	0.312	0.088	0.146	0.490	0.085
Income	0.232	0.109	0.029	0.456	0.065
Education	0.352	0.081	0.198	0.514	0.094
Black	0.324	0.144	0.032	0.599	0.044
Hispanic	0.136	0.137	−0.124	0.409	0.016
Family size	0.070	0.067	−0.062	0.201	0.021
Married	0.122	0.133	−0.132	0.390	0.018
Unemployment exp.	−0.367	0.193	−0.759	−0.002	−0.038
Rural area	−0.210	0.131	−0.461	0.053	−0.030
South	0.306	0.131	0.057	0.565	0.043
(B) Non-members (Y_0)					
Latent factor	0.207	0.039	0.131	0.285	0.075
Income	0.076	0.031	0.018	0.137	0.028
Education	0.355	0.034	0.288	0.421	0.124
Black	0.268	0.066	0.138	0.398	0.046
Hispanic	−0.303	0.067	−0.431	−0.170	−0.046
Family size	0.115	0.030	0.056	0.172	0.042
Married	0.323	0.060	0.202	0.439	0.058
Unemployment exp.	−0.271	0.072	−0.413	−0.134	−0.033
Rural area	0.015	0.054	−0.094	0.120	0.003
South	−0.075	0.049	−0.171	0.017	−0.014

Note: Cohort dummies and intercept not shown.
[a] First difference in predicted probability of unit change in covariate.

B. Treatment effects

In contrast to the wealth of tables produced in the previous section, the summary of our treatment effects is straightforward. Following equation (C.1) in the appendix, we calculate the (population) average treatment effect of union membership on turnout. The first line of Table 6 shows a summary of the posterior distribution of this quantity. The average treatment effect of union membership on turnout is estimated as 0.104 ± 0.018. In other words, even after accounting for selection on observables and unobservables, union membership increases the likelihood of turnout by 10.4 percentage points. 95% of the posterior density of the average treatment effect lies between roughly 7 and 14% – clearly quite a way from being zero. Since the difference between members and non-members in the raw data (recall Table 1) is fourteen points, our results suggest that selection accounts for about one third and perhaps as much as one half of the observed difference. In other words, unions do increase voter participation,

Table 6: Effect of union membership on turnout. Average Treatment Effect estimates from baseline model and under several robustness checks.

	Mean	SD	95% HPD	
Baseline model ATE	0.104	0.018	0.069	0.139
Robustness checks				
State fixed effects	0.110	0.025	0.061	0.159
Public sector employment	0.108	0.019	0.071	0.144
Industry fixed effects	0.102	0.018	0.069	0.138
Random subsamples	0.119	0.023	0.076	0.167

Note: Based on 10,000 MCMC samples. Values are probability differences.

though the type of person who becomes a union member is quite different, on average, from one who does not.

Robustness tests We conduct several robustness checks, which are summarized in the lower half of Table 6. In our first specification, we include state fixed effects. These capture time-constant state-level confounders omitted from our model. The most relevant among those is probably "right-to-work" (RTW) legislation. Under such a law, employees in unionized workplaces may opt out of union membership without foregoing collective benefits. In terms of our model, this systematically affects unobserved costs of union membership in some states. Since RTW legislation is time-constant in our sample, including state fixed effects captures its effect. Furthermore, we estimated models including an indicator variable for public sector employment, as well as industry fixed effects (thus using only within-industry changes in concentration levels).[26] Finally, we used a more radical random subsample approach to gauge the stability of our inference. We re-estimate our models 5 times, while each time randomly deleting one third of observations and then average our estimates with an added penalty for variability (following the rules of Little and Rubin 2002). As Table 6 shows, we find our central results confirmed: the substantive magnitude of ATE estimates is very similar. In fact, the credible intervals of all robustness models overlap, indicating that results don't differ statistically.

VI. CONCLUSION

The effect of labor unions on voting concerns the fundamental relationship between economics and democratic politics. In this paper, we have used a unique survey data set to

[26]We use the North American Industry Classification System (NAICS), 2002 revision, at the 1-digit level.

provide robust estimates of the causal effect of union membership on turnout. Our empirical approach accounts for the problem that prospective voters may be sorted into unions through the labor market or self-selection. To jointly model endogenous union membership and vote choice, we have drawn on three distinct sources of causal identification in the presence of unobserved confounders: economic incentives captured via industry-specific variables, a random factor structure, and explicit measures of cognitive ability. We find that sorting into union membership based on workers' ability accounts for a significant part of the observed turnout gap between union members and non-members. Accounting for sorting, however, there remains a statistically and politically significant union membership turnout premium.

One limitation of this study is that it only considers one election for the cohort that makes up the NLSY. This reflects data constraints. While the panel survey we analyze is exceptionally rich in economic and psychological items, it rarely measures turnout. There is obviously no easy statistical fix for dealing with unobserved heterogeneity in the study of political participation. One main advantage of the approach we have taken is that it exploits high-quality data on individual abilities that are not featured in surveys frequently used to study voting (like the American National Election Study). It clearly illustrates the potential of including similar items, possibly in an abridged version, in election surveys. Taken together, our analysis confronts the problem of endogenous membership raised by both theoretical and empirical scholars of groups and voting (Feddersen 2004). It shows how it can be addressed empirically, finding that unions do matter for voting.

REFERENCES

Aakvik, A., J. J. Heckman, and E. Vytlacil. 2005. "Estimating treatment effects for discrete outcomes when responses to treatment vary: an application to Norwegian vocational rehabilitation programs." *Journal of Econometrics* 125 (1-2): 15–51.

Abbring, J. H., and J. J. Heckman. 2007. "Econometric evaluation of social programs, part III: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation." In *Handbook of Econometrics*, ed. J. J. Heckman and E. E. Leamer. Vol. 6. Elsevier.

Abraham, K. G., and J. L. Medoff. 1984. "Length of service and layoffs in union and nonunion work groups." *Industrial & Labor Relations Review* 38 (1): 87–97.

Ahlquist, J. S., A. B. Clayton, and M. Levi. 2014. "Provoking Preferences: Unionization, Trade Policy, and the ILWU Puzzle." *International Organization* 68: 33–75.

Akerlof, G. A. 1980. “A theory of social custom, of which unemployment may be one consequence.” *Quarterly Journal of Economics* 94 (4): 749–775.

Albert, J. H., and S. Chib. 1993. “Bayesian Analysis of Binary and Polychotomous Response Data.” *Journal of the American Statistical Association* 88 (422): 669–679.

Aldrich, John H. and K. M. McGraw, eds. 2011. *Improving Public Opinion Surveys. Interdisciplinary Innovation and the American National Election Studies*. Princeton: Princeton University Press.

Anderson, T. W., and H. Rubin. 1956. “Statistical inference in factor analysis.” In *Proceedings of the third Berkeley symposium on mathematical statistics and probability*, ed. J. Neyman. Vol. 5, Berkeley pp. 111–150.

Anzia, S. F. 2011. “Election Timing and the Electoral Influence of Interest Groups.” *The Journal of Politics* 73 (2): 412–427.

Bartholomew, D. 1988. “The sensitivity of latent trait analysis to choice of prior distribution.” *British Journal of Mathematical and Statistical Psychology* 41 (1): 101–107.

Borjas, G. J. 2013. *Labor Economics*. Boston: McGraw-Hill.

Cameron, S. V., and J. J. Heckman. 1998. “Life Cycle Schooling and Dynamic Selection Bias: Models and Evidence for Five Cohorts of American Males.” *Journal of Political Economy* 106 (2): 262–333.

Carneiro, P., K. T. Hansen, and J. J. Heckman. 2003. “Estimating Distributions of Treatment Effects with an Application to the Returns to Schooling and Measurement of the Effects of Uncertainty on College Choice.” *International Economic Review* 44 (2): 361–422.

Chib, S. 2007. “Analysis of treatment response data without the joint distribution of potential outcomes.” *Journal of Econometrics* 140 (2): 401–412.

Corneo, G. G. 1997. “The theory of the open shop trade union reconsidered.” *Labour Economics* 4 (1): 71–84.

Cunha, F., J. J. Heckman, and S. M. Schennach. 2010. “Estimating the Technology of Cognitive and Noncognitive Skill Formation.” *Econometrica* 78 (3): 883–931.

Cunha, F., J. J. Heckman, and S. Navarro. 2005. “Separating uncertainty from heterogeneity in life cycle earnings.” *Oxford Economic Papers* 57 (2): 191–261.

Dawes, C. T., J. E. Settle, P. J. Loewen, M. McGue, and W. G. Iacono. 2015. “Genes, psychological traits and civic engagement.” *Philosophical Transactions of the Royal Society Series B* 370 (1683).

Delaney, J. T., M. F. Masters, and S. Schwochau. 1988. "Unionism and voter turnout." *Journal of Labor Research* 9 (3): 221–236.

Denny, K., and O. Doyle. 2008. "Political Interest, Cognitive Ability and Personality: Determinants of Voter Turnout in Britain." *British Journal of Political Science* 38 (02): 291–310.

Eckstein, Z., and K. I. Wolpin. 1989. "The specification and estimation of dynamic stochastic discrete choice models: A survey." *The Journal of Human Resources* 24 (4): 562–598.

Feddersen, T., and A. Sandroni. 2006. "A Theory of Participation in Elections." *The American Economic Review* 96 (4): 271-1282.

Feddersen, T. J. 2004. "Rational Choice Theory and the Paradox of Not Voting." *The Journal of Economic Perspectives* 18 (1): 99-112.

Flavin, P., and B. Radcliff. 2011. "Labor Union Membership and Voting Across Nations." *Electoral Studies* 30 (4): 633–641.

Freeman, R. B. 1980. "The Effect of Unionism on Fringe Benefits." *Industrial and Labor Relations Review* 34: 489.

Freeman, R. B. 2003. "What Do Unions Do ... To Voting?" Working Paper 9992 National Bureau of Economic Research.

Freeman, R. B., and J. Medoff. 1984. *What Do Unions Do?* New York: Basic Books.

Gerber, A. S., and D. P. Green. 2000. "The Effects of Canvassing, Telephone Calls, and Direct Mail on Voter Turnout: A Field Experiment." *The American Political Science Review* 94 (3): 653-663.

Gill, J. 2008. *Bayesian Methods. A Social and Behavioral Sciences Approach*. Boca Raton: Chapman & Hall.

Hauser, S. M. 2000. "Education, ability, and civic engagement in the contemporary United States." *Social Science Research* 29 (4): 556–582.

Heckman, J. J. 1981. "Statistical models for discrete panel data." In *Structural Analysis of Discrete Data with Econometric Applications*, ed. C. Manski and D. McFadden. Cambridge: MIT Press.

Heckman, J. J. 1990. "Varieties of selection bias." *American Economic Review* 80 (2): 313–318.

Heckman, J. J., and E. J. Vytlacil. 1999. "Local instrumental variables and latent variable models for identifying and bounding treatment effects." *Proceedings of the National Academy of Sciences* 96 (8): 4730–4734.

Heckman, J. J., and E. J. Vytlacil. 2007. "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments." In *Handbook of Econometrics*, ed. J. J. Heckman and E. E. Leamer. Vol. 6, Part B of *Handbook of Econometrics*. Elsevier.

Heckman, J. J., H. F. Lopes, and R. Piatek. 2013. "Treatment Effects: A Bayesian Perspective." *Econometric Reviews* 33 (1-4): 36–67.

Hirsch, B. T., and M. C. Berger. 1984. "Union Membership Determination and Industry Characteristics." *Southern Economic Journal* 50 (3).

Holland, P. W. 1986. "Statistics and Causal Inference." *Journal of the American Statistical Association* 81 (396): 945–960.

Imbens, G. W., and J. D. Angrist. 1994. "Identification and estimation of local average treatment effects." *Econometrica* 62 (2): 467–475.

Jackman, S. 2008. "Measurement." In *Oxford Handbook of Political Methodology*, ed. J. M. Box-Steffensmeier, H. E. Brady, and D. Collier. Oxford: Oxford University Press.

Jensen, A. R. 1985. "Armed Services Vocational Aptitude Battery." *Measurement and Evaluation in Counseling and Development* 18 (1): 32–37.

Kerrissey, J., and E. Schofer. 2013. "Union Membership and Political Participation in the United States." *Social Forces* 91 (3): 895-928.

Kim, S. E., and Y. Margalit. 2016. "Informed Preferences? The Impact of Unions on Workers' Policy Views." *American Journal of Political Science*: Forthcoming.

Krosnick, J. A., and A. Lupia. 2006. "A Report on the First Collaboration between the American National Election Studies and the National Longitudinal Surveys." ANES Newsletter, January 19, 2006.

Kwoka, J. E. J. 1983. "Monopoly, plant, and union effects on worker wages." *Industrial and Labor Relations Review*: 251–257.

Lang, K., and M. Manove. 2011. "Education and Labor Market Discrimination." *American Economic Review* 101 (4): 1467–1496.

Leighley, J. E., and J. Nagler. 2007. "Unions, Voter Turnout, and Class Bias in the U.S. Electorate, 1964-2004." *The Journal of Politics* 69 (2): 430-441.

Little, R. J. A., and D. B. Rubin. 2002. *Statistical Analysis with Missing Data*. Hoboken: Wiley.

Luskin, R. C. 1990. “Explaining political sophistication.” *Political Behavior* 12 (4): 331–361.

Masters, M. F., and J. T. Delaney. 2005. “Organized Labor’s Political Scorecard.” *Journal of Labor Research* 26 (3): 365-392.

Neuhaus, J. M., W. W. Hauck, and J. D. Kalbfleisch. 1992. “The effects of mixture distribution misspecification when fitting mixed-effects logistic models.” *Biometrika* 79 (4): 755–762.

Nie, N., J. Junn, and K. Stehlik-Barry. 1996. *Education and citizenship in America*. Chicago: Cambridge University Press.

Norris, P. 2002. *Democratic Phoenix: Reinventing Political Activism*. Cambridge: Cambridge University Press.

Pryor, F. L. 1972. “An international comparison of concentration ratios.” *Review of Economics and Statistics* 54 (2): 130–140.

Robinson, C. 1989. “The Joint Determination of Union Status and Union Wage Effects: Some Tests of Alternative Models.” *Journal of Political Economy* 97 (3): 639-667.

Rosenfeld, J. 2014. *What Unions No Longer Do*. Cambridge, MA; London, England: Harvard University Press.

Rubin, D. B. 1978. “Bayesian inference for causal effects: The role of randomization.” *The Annals of Statistics* 6 (1): 34–58.

Schlozman, K. L., S. Verba, and H. E. Brady. 2012. *The Unheavenly Chorus*. Princeton and Oxford: Princeton University Press.

Schnabel, C. 2003. “Determinants of trade union membership.” In *International Handbook of Trade Unions*, ed. J. T. Addison and C. Schnabel. Cheltenham: Edward Elgar.

Skrondal, A., and S. Rabe-Hesketh. 2004. *Generalized latent variable modeling: Multilevel, longitudinal and structural equation models*. Boca Raton: Chapman & Hall.

Stephens, J. D., and M. Wallerstein. 1991. “Industrial Concentration, Country Size, and Trade Union Membership.” *American Political Science Review*: 941–953.

Uhlaner, C. J. 1989. “Rational Turnout: The Neglected Role of Groups.” *American Journal of Political Science* 33 (3): 390-422.

Verba, S., K. L. Schlozman, and H. E. Brady. 1995. *Voice and Equality*. Harvard: Harvard University Press.

Vijverberg, W. P. 1993. "Measuring the unidentified parameter of the extended Roy model of selectivity." *Journal of Econometrics* 57 (1): 69–89.

Vytlacil, E. 2002. "Independence, monotonicity, and latent index models: An equivalence result." *Econometrica* 70 (1): 331–341.

Wedel, M., and W. A. Kamakura. 2001. "Factor analysis with (mixed) observed and latent variables in the exponential family." *Psychometrika* 66 (4): 515–530.

A. [ONLINE] APPENDICES

A. Variable definitions

Here we list all variables used in our model together with their definitions. Until noted otherwise these are based on NLSY data. *Income* is measured as total wage and salary income before taxes and deductions in contemporary US dollars. *Education* is years of schooling completed. *Black* and *Hispanic* are indicator variables (based on self-assessed race). *Married* is an indicator variable for being married. *South* is an indicator variable for living in a Southern state (as defined by Census region). *Rural area* is an indicator for living in a rural area (defined following the Census definition of living in an 'urbanized area' or in a place with greater than 2,500 population). *Family size* is the number of persons living in a household (based on household enumeration data). *Unemployment experience* is an indicator variable equal to one if a respondent was unemployed for any period of time in the past calendar year.

The following variables were used as controls in our cognitive test equations. *Education father* and *Education mother* are the highest grade completed by father and mother (based on the respondent's information). Missing information on these variables is imputed from predictions based on family income in 1980.[27] *Broken family* is an indicator equal to 1 if a respondent lived with a single parent at age 14. *Number of siblings* is the number of siblings in the respondent's household at age 14 (from household enumeration information). *Family income 1980* is a respondent's family's income in 1980 (from household interview data). *Age at test* is a respondent's age when taking the cognitive test.

Industry concentration and *Union wage differential* are calculated from administrative sources. We use the CR4 concentration ratio from the Census Bureau's Economic Census of American businesses conducted in 2007 for 243 industries. We calculate union-nonunion wage differentials from the Bureau of Labor Statistics' CPS-LU series, for 19 major industrial sectors. Our measure of union-nonunion wage differentials is the difference in median weekly earnings in contemporary dollars. Values refer to sole or principal job of full-time wage and salary workers. Excluded are all self-employed workers regardless of whether or not their businesses are incorporated.

Table A.1 shows union-non-union wage differentials; Figure A.1 plots histograms of industry concentration ratios separately for the 16 major sectors of the economy.

[27]We also conducted robustness tests showing that excluding one or both variables from our model does not substantively alter results.

Table A.1: Union-nonunion wage difference

Industry	Difference[a]
Finance and insurance	−97
Agriculture, forestry, fishing, and hunting	0[b]
Mining, quarrying, and oil and gas extraction	21
Wholesale trade	36
Retail trade	37
Nondurable goods manufacturing	54
Professional and technical services	65
Arts, entertainment, and recreation	65
Durable goods manufacturing	66
Real estate and rental and leasing	93
Health care and social assistance	104
Educational services	130
Accommodation and food services	136
Information	140
Management, administrative, and waste services	144
Utilities	151
Transportation and warehousing	187
Other services	271
Construction	346

Source: US Bureau of Labor Statistics, series LU.
[a] Difference in median weekly earnings (contemporary $). [b] Not calculated since sample size less then 50,000.

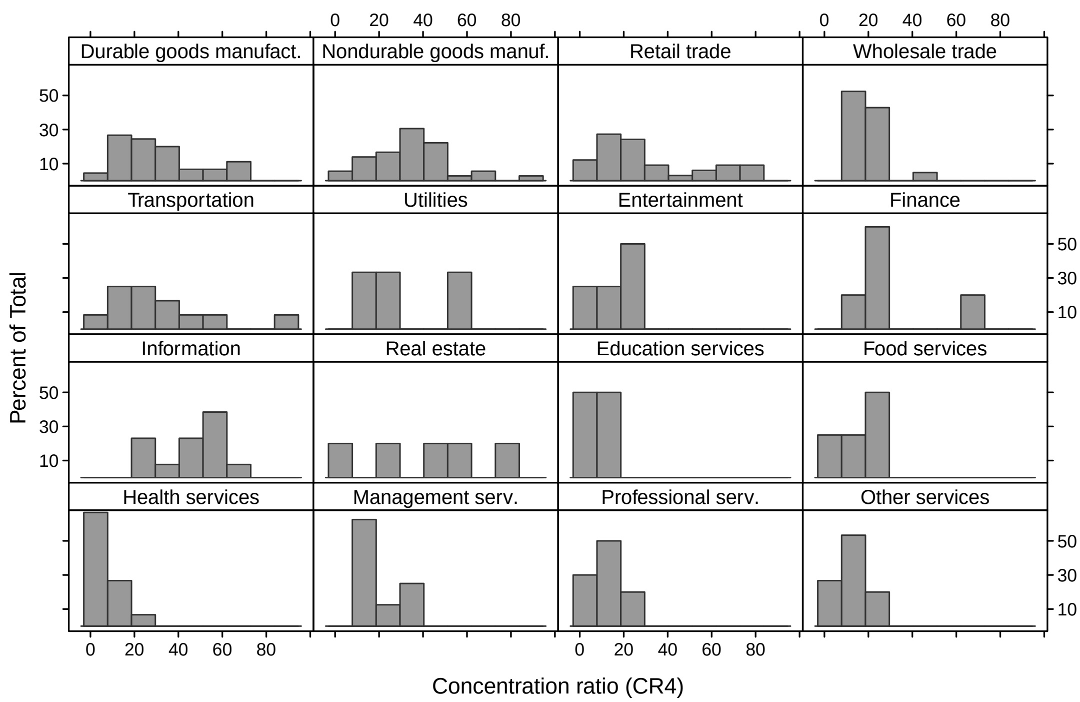

Figure A.1: Industry concentration by major economic sectors

Table B.1: Variables in membership, turnout, and ability test equations

	(Y_0, Y_1)	D	M_p
Income	x		
Education	x	x	
Black	x	x	
Hispanic	x	x	
Married	x		
Family size	x		
South	x	x	
Rural area	x	x	
Cohort dummies	x	x	
Unemployment experience	x		
Industry concentration		x	
Union wage differential		x	
Education father			x
Education mother			x
Broken family at 14			x
Number of siblings			x
Family income 1980			x
Age at test			x

C. Treatment effects

To simplify notation, denote by Γ the vector of all model parameters. The average treatment effect, conditional on covariates, represents the effect of union membership for a randomly chosen individual with characteristics X. It is given by

$$\begin{aligned} ATE(X,\Gamma) &= Pr(Y_1 = 1|X,\Gamma) - Pr(Y_0 = 1|X,\Gamma) \\ &= \Phi\left(\frac{\beta_1' X}{\sqrt{1+\alpha_1'\alpha_1}}\right) - \Phi\left(\frac{\beta_0' X}{\sqrt{1+\alpha_0'\alpha_0}}\right), \end{aligned} \tag{C.1}$$

where Φ is the CDF of the normal distribution. The corresponding treatment effect on the treated represents the effect of union membership on turnout among union members. It is obtained by conditioning on $D = 1$:

$$\begin{aligned} TT(X,\Gamma,D=1) &= Pr(Y_1 = 1|X,D=1,\Gamma) - Pr(Y_0 = 1|X,D=1,\Gamma) \\ &= \left(\frac{\Phi(\beta_D' Z)}{\sqrt{1+\alpha_D'\alpha_D}}\right)^{-1} \int \left[\Phi(\beta_1' X + \alpha_1\theta) - \Phi(\beta_0' X + \alpha_0\theta)\right] \\ &\quad \times \Phi(\beta_D' Z + \alpha_D\theta)\phi(\theta)d\theta \end{aligned} \tag{C.2}$$

Here, ϕ denotes the normal distribution PDF. Since here we are interested in describing population average treatment effects of union membership (unconditional of individual characteristics), we integrate over the (empirical) distribution of X. In other words, $E(ATE) = \int ATE(x)dF_X(x)$ and *mutatis mutandis* for $E(TT)$.

D. Details on prior distributions

We assume independent priors for factor coefficients in potential outcome equations $\alpha_j \sim N(\tilde{\alpha}_j, \tilde{\nu}_j), j = 0, 1$. We use common inverse Gamma priors for error variances: $\sigma_p^{-2} \sim G(\tilde{a}_p, \tilde{b}_p)$, where a and b are shape and scale parameters of the Gamma distribution, respectively. For slopes in potential outcome and choice equations we use regression-type priors $\beta_j \sim N(\tilde{\beta}_j, \tilde{B}_j)$, $j = 0, 1, D$, with $\tilde{B}_j = \boldsymbol{I}_j \tilde{b}_j$. Finally, we use normal priors for θ-coefficients in our measurement equations, $\lambda_p \sim N(\tilde{\lambda}_p, \tilde{\nu}_p)$, and we use normal priors for covariates in these equations as well: $\alpha_{M_p} \sim N(\tilde{\alpha}_{M_p}, \tilde{\nu}_{M_p})$. The actual numerical values for these priors are chosen such that they are "uninformative", i.e., they express a priori ignorance (for example by having mean zero and large prior variance of, say, 100). Numerical values used are given in Table D.1.

We conduct a range of prior sensitivity analyses (see Gill (2008: 199f.) for an overview). Table D.1 lists hyper-parameter values used in the main text (S1) and for two different prior sensitivity simulations. Specification 2 used alternative parameters for the inverse Gamma distribution. Specification 3 use prior variances 10 times larger for loadings and all effect parameters. In all specifications prior mean values were kept at zero to signal our *a priori* ignorance about the true effect. The result of this exercise yields estimates that are numerically close and substantively identical to the ones used in the main text of our paper. Table D.2 shows that the resulting average treatment effects are all very close.

Table D.1: Prior parameters specifications

Prior hyperparameters		Values		
		S1	S2	S3
$\tilde{\alpha}_j$	$j = 0, 1$	0	0	0
$\tilde{v}_j$	$j = 0, 1$	10	10	100
$\tilde{\beta}_{jk}$	$j = 0, 1, D;\ k = 1, \ldots, K$	0	0	0
$\tilde{b}_{jk}$	$j = 0, 1, D;\ k = 1, \ldots, K$	10	10	100
$\tilde{\lambda}_p$	$p = 1, \ldots, P$	0	0	0
$\tilde{v}_p$	$p = 1, \ldots, P$	10	10	100
$\tilde{\alpha}_{M_p}$	$p = 1, \ldots, P$	0	0	0
$\tilde{v}_{M_p}$	$p = 1, \ldots, P$	10	10	100
$\tilde{a}_p$	$p = 1, \ldots, P$	1	1	1
$\tilde{b}_p$	$p = 1, \ldots, P$	2	0.005	2

Table D.2: Prior robustness checks. Posterior summary of average treatment effects under alternative prior parametrizations.

Specification	Mean	SD	95% HPD	
S1	0.104	0.018	0.069	0.139
S2	0.105	0.018	0.069	0.140
S3	0.104	0.018	0.069	0.140