

COGNITIVE ABILITY, UNION MEMBERSHIP, AND VOTER TURNOUT

Daniel Stegmueller

Duke University

Michael Becher

Institute for Advanced Study in Toulouse, UT1

This version: January 2019

ABSTRACT

Labor unions are said to influence elections and public policy by increasing their members' electoral turnout. But existing research likely overestimates the turnout effect of union membership by ignoring sorting in the labor market. In the presence of a union wage premium, both membership and turnout are shaped by the same (unobserved) factors, such as cognitive ability. To disentangle the union effect from positive selection, we use unique 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 about one-third of the observed union turnout effect is due to selection, more than what previous studies suggest.

JEL classification: A12, 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 meetings of APSA, EPSA and MPSA.

The political power of labor unions is based to an important extent on unions' ability to increase voter turnout among their members, thereby influencing election outcomes and a broad range of economic policies at the local and national level (Anzia 2011, 2012; Bartels 2008; Becher, Stegmueller, and Kaepfner 2018; Feigenbaum, Hertel-Fernandez, and Williamson 2018; Freeman and Medoff 1984; Masters and Delaney 2005; Schlozman, Verba, and Brady 2012). In line with this view, numerous studies by political scientists, economists and sociologists have documented a positive relationship between individual union membership and political participation in contemporary democracies (Delaney, Masters, and Schwochau 1988; Freeman 2003; Flavin and Radcliff 2011; Leighley and Nagler 2007; Norris 2002; Rosenfeld 2014). However, economic theory suggests that in the presence of a union wage premium, union membership is influenced by the same (unobserved) personal characteristics that influence voter turnout. In particular, cognitive ability matters for both selection into union jobs and the decision to turn out to vote. While selection into formal groups based on normative or political motivations is a well-recognized identification problem in the literature on turnout (Abrams, Iversen, and Soskice 2011), selection driven by sorting in the labor market has been neglected. In fact, scholars often assume that the importance of economic incentives for union membership in the United States ensures that selection is not a problem for the study of voting (Kerrissey and Schofer 2013: 918; Rosenfeld 2014: 145).¹

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 and spell out an approach to model it. Sorting in the labor market implies that union members and non-members differ systematically on fundamental determinants of political behavior. Exploiting the unique features of the NLSY, we implement an empirical strategy that uses auxiliary information to allow us to 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 for higher ability job applicants. The result, then, is sorting on ability between union and nonunion workers (Freeman and Medoff 1984: 45; Robinson 1989: 643). Complementarily, research in political science has

¹The 2018 Supreme Court decision (*Janus v. American Federation of State, County, and Municipal Employees*) rules out union fees for non-members in the public sector.

established a link between cognitive ability and electoral turnout. Individuals with 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 taken into account in research on 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.² 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 US states without “right-to-work” legislation. While such union shops may mitigate endogeneity concerns based on explicit political motivations (Kim and Margalit 2017; 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 (Abbring and Heckman 2007), we exploit the rich individual-level data of the NLSY. It allows 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 insights from labor economics, our empirical approach captures employers’ incentives to carefully screen candidates for more costly unionized jobs, as well as employees’ 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

²The common practice of controlling for political attitudes and other proximate predictors of turnout invites post-treatment bias (e.g., see Samii 2016).

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 rich 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 of the observed descriptive difference in turnout between members and non-members. Accounting for the selection process, there is nonetheless 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 in the particular election we study—a politically significant effect.

This paper contributes to several strands of research. First, it adds to the multi-disciplinary literature on the political effects of unions. Due to data limitations, most studies of union membership and political preferences or voting rely on cross-sectional regression analysis with covariate adjustment. Recent exceptions exploit quasi-random variation to estimate union effects. The study of Kim and Margalit (2017) 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. Leveraging the differential adoption of right-to-work laws, which weaken unions' shop protections, in neighboring countries across state borders, Feigenbaum, Hertel-Fernandez, and Williamson (2018) find that the laws reduce Democratic votes shares in presidential, congressional and gubernatorial elections by about 3.5 percentage points. 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, our results suggest that the selection effect for turnout is larger than the one found for trade preferences by Kim and Margalit (2017).

Second, our approach and findings are also relevant for the broader turnout literature. Recent formal theories of voting have turned their attention to group dynamics to explain

voting in large elections (Feddersen and Sandroni 2006; Herrera, Morelli, and Nunnari 2016). 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 in political behavior and thereby providing a clear basis for empirical research.

ENDOGENOUS UNION MEMBERSHIP AND VOTING

To motivate the empirical analysis to follow, let us clarify the problem with some notation. 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:

$$Y_{0i} = \mu_0(X_i) + U_{0i} \quad \text{if } D_i = 0 \tag{1}$$

$$Y_{1i} = \mu_1(X_i) + U_{1i} \quad \text{if } D_i = 1. \tag{2}$$

Crucially, unobservables U_{0i} and U_{1i} may be correlated with unobservable factors explaining sorting into union membership. We can think of sorting into union membership D_i as a function

of observable and unobservable individual characteristics:

$$\begin{aligned} D_i^* &= \mu_D(Z_i) + U_{Di} \\ D_i &= 1 \text{ if } D_i^* \geq 0, D_i = 0 \text{ otherwise.} \end{aligned} \tag{3}$$

Here, D_i^* is a latent index representing the net utility or gain of union membership, Z_i are observed individual characteristics (including factors that make union jobs more attractive, such as union wage differentials), whereas U_{Di} represents unobservables shifting the propensity of union membership.³

Unobserved individual characteristics affecting turnout and union membership are collected in the random vector $(U_{0i}, U_{1i}, U_{Di})'$. 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 of (U_0, U_1, U_D)

$$\Sigma = \begin{bmatrix} \sigma_0^2 & \sigma_{01} & \sigma_{0D} \\ \sigma_{01} & \sigma_1^2 & \sigma_{1D} \\ \sigma_{0D} & \sigma_{1D} & \sigma_D^2 \end{bmatrix}. \tag{4}$$

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: σ_{jk} parametrizes the covariance between U_j, U_k . Notably, σ_{01} represents the covariance between unobservables in potential outcome states for union members and non-union members. Since we can never observe the same individual in two treatment states at once, σ_{01} is not identified from the data (Vijverberg 1993: 74). This is the root of the well-known “fundamental problem of causal inference” (Holland 1986: 947).

There are clear reasons why one would expect sorting into unions based on unobservables that also impact voting. Theories of labor markets suggest that sorting into union jobs is 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; see Borjas 2013: 444 for a recent textbook treatment). On the worker side, better wages or benefits induce more workers to apply for a unionized job, increasing the pool of attractive candidates. Even if employers

³The latent index structure is a common setup for many econometric choice models (e.g., Maddala 1986) and is also used in ‘classical’ selection models (Heckman 1976). It is quite instructive to note the equivalence between the (nonparametric) latent index model and the LATE approach of Imbens and Angrist (1994): the assumption that there is an unobserved index crossing a threshold (above which we observe an individual being a union member) is equivalent to the independence and monotonicity assumptions made for LATE. See Vytlačil (2002) for a detailed discussion of the equivalence result.

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 higher ability types out of the queue of workers applying for unionized jobs (Abowd and Farber 1982). These arguments (and their empirical support) strongly suggest that we should expect sorting on ability between union and nonunion workers.

The threat to valid inference stems from the fact that cognitive ability is likely to also influence turnout. Ability is a fundamental trait discussed in the literature on turnout, but it is not measured 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)⁴, 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). It includes a latent factor capturing unobserved individual heterogeneity that introduces a correlation between union membership and voting (i.e., $\sigma_{jk} \neq 0$). Moreover, our setup allows us to exploit additional information in order 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. Note that cognitive tests are not employed as an instrumental variable. Rather, they serve as a proxy for unobserved characteristics relevant for both union membership and political participation that helps to pin down the structure of *unobserved* heterogeneity (the section

⁴The 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.

“Latent variable potential outcome model” provides more details).⁵

This identification strategy is complemented with the more ‘traditional’ approach of including instruments that only shift the probability of union membership. Equation (3) above contains two sets of covariates 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 (conditional on X_i and θ_i). These serve as instrumental variables. We include two instruments that both capture economic aspects of selection into union membership. First, the higher the union wage or benefit premium, the more likely it is that we find a worker being a union member (Lee 1978, Schnabel 2003: 14).⁶ A second factor influencing union membership is industry concentration: union organization is less costly in sectors with four firms than with 50 (Lee 1978: 416, Hirsch and Berger 1984, Stephens and Wallerstein 1991: 943). 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. When introducing our model below, we discuss possible threats to the validity of these instruments and how we address them.⁷ After describing the data set that enables our analysis, we explain how we transform this conceptual framework into an empirically estimable setup.

DATA

We use the National Longitudinal Study of Youth (NLSY), a longitudinal panel study directed by the U.S. Department of Labor’s Bureau of Labor Statistics.⁸ Its widespread use in economics 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

⁵Over 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.

⁶Wage differentials may be due to monopoly power, firm-worker matching on productivity, or represent compensatory payments for work conditions.

⁷It is important to remember that the model is identified without any instrumental variables. Absent any valid instruments, however, we need a distributional assumption on the latent variable. We return to this issue below.

⁸The 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.

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 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.⁹ 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. 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.

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 2007.¹⁰ The high quality of this data allows us to use detailed disaggregated concentration ratios for 243 industries.

Figure I plots the distribution of union wage differentials and industrial concentration.¹¹ It

⁹We also exclude NLSY's military subsample, as members of the military are not union members.

¹⁰The Economic Census is conducted in 5-year intervals and samples around 4 million firms. Participation is required by law.

¹¹Appendix A contains more detailed descriptive information on these two variables.

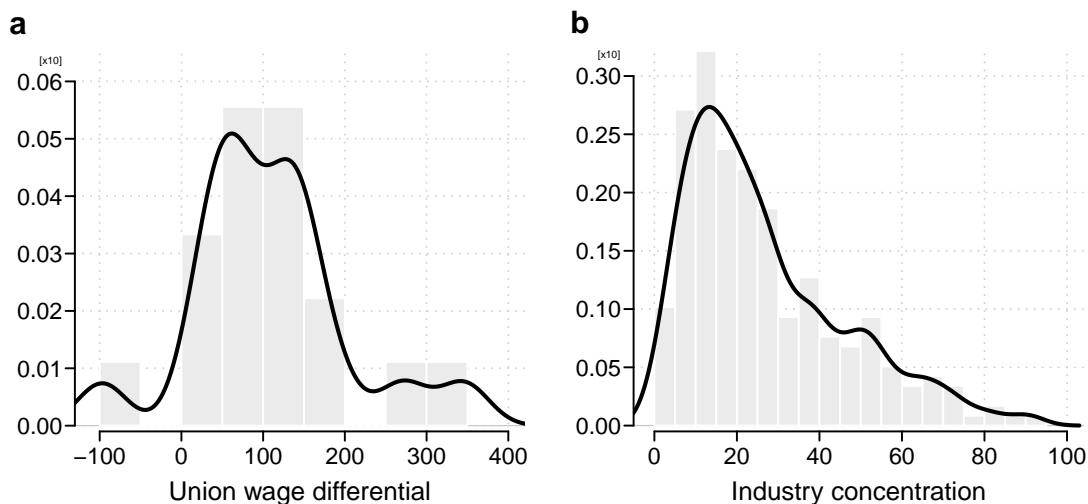


FIGURE I
Distribution of industry characteristics.

Panel (a) shows union wage differentials measured via differences in median weekly earnings (in \$) between union members and non-members. Panel (b) shows levels of industrial concentration measured via the 4-firm concentration ratio. Histograms with kernel density plots.

shows the existence of substantial variation in both variables. 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). 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.

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, see Jensen 1985) to the civilian and military NLSY79 samples.¹² The ASVAB consists of several 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.

¹²The 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.

Turnout. After the November election in 2006 respondents were queried if they voted.¹³ 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”.¹⁴ 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 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 indicator variables capturing systematic cohort differences.

Table I provides descriptive means of our central variables for union members and non-members. It shows that the (unadjusted) difference in turnout is 14 percentage points.¹⁵ In terms of observable characteristics, union members experience fewer unemployment spells and are less likely to be from the South, where right-to-work legislation is predominant. Union members work in industries that are characterized by larger wage differentials and are more concentrated. Most notably, we also find that union members are characterized by higher levels of cognitive ability than non-members (despite having rather similar levels of education). Insofar as cognitive ability also shapes turnout—we show later that a standard deviation increase in ability is associated with a more than seven percentage point increase in turnout—this raises clear concerns about the presence of selection bias.

¹³The 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?”

¹⁴Over-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%.

¹⁵This is nearly identical to the unadjusted difference in the employed population for mid-term elections based on CPS data which is 13 percentage points (Freeman 2003). Averaging across all elections between 1984-2008, the difference is closer to 10 percentage points (Rosenfeld 2014).

TABLE I
Sample characteristics. Means and standard errors.

	Union members [N=456]	Non-members [N=2004]
Turnout rate	0.73 (0.02)	0.59 (0.01)
<i>Individual characteristics</i>		
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)
Cognitive ability ($\hat{\theta}$) ^b	0.79 (0.25)	-0.14 (0.21)
<i>Industry characteristics</i>		
Industry concentration ^c	28.58 (1.08)	23.21 (0.47)
Wage differential [\$ / week] ^d	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 Estimated cognitive ability. See equation (11) for a discussion of its construction in our model and estimation. Measure is normalized in sample to have mean zero and standard deviation of 10.

^c 4-firm concentration ratio CR_4 .

^d Difference in median weekly earnings (in \$) of full-time employed union members and non-members at 2-digit industry level.

LATENT VARIABLE POTENTIAL OUTCOME MODEL

In this section we discuss a solution to the endogeneity problem outlined previously. 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) .¹⁶ 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 (4) using

¹⁶From here on we assume that we have access to an equiprobability sample and suppress individual subscripts for easier notation.

an underlying low-dimensional set of random factors (Heckman 1981). Thus, following Aakvik, Heckman, and Vytlačil (2005), we decompose unobservables using the following factor structure:

$$U_D = \alpha_D \theta + \epsilon_D \tag{5}$$

$$U_0 = \alpha_0 \theta + \epsilon_0 \tag{6}$$

$$U_1 = \alpha_1 \theta + \epsilon_1 \tag{7}$$

Here, θ is a latent factor or random effect (Cameron and Heckman 1998; Skrondal and Rabe-Hesketh 2004), which represents unobserved individual heterogeneity, such as differences in 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.

Assuming this factor structure solves the core identification problem (Aakvik, Heckman, and Vytlačil 2005; Carneiro, Hansen, and Heckman 2003).¹⁷ To see this, note that the non-identified parameter, $\sigma_{10} \equiv \text{Cov}(Y_0, Y_1)$, can now be obtained as $\alpha_0 \alpha_1$. The latent factor generates 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.

We specify $\theta \sim N(0, 1)$. This is an a priori distributional assumption. While it is quite robust against misspecification (Bartholomew 1988; Neuhaus, Hauck, and Kalbfleisch 1992; Wedel and Kamakura 2001), it is not necessary for identification and can be relaxed.¹⁸ To fix the sign of the latent factor (Anderson and Rubin 1956), we fix α_D to 1.¹⁹ Thus, we anchor θ to

¹⁷Another way to deal with this fundamental non-identification is to provide bounds instead of a point estimate, by leveraging the positive definiteness constraint of the variance covariance matrix (Vijverberg 1993; Heckman, Smith, and Clements 1997) or the prior dependence between identified and unidentified parameters (Koop and Poirier 1997; Poirier 1998). However obtained bounds are often quite wide. Furthermore, results are highly influenced by specific prior choices (Poirier and Tobias 2003), and formulating informed a priori values is difficult.

¹⁸Cunha, Heckman, and Navarro (2005) and Cunha, Heckman, and Schennach (2010) discuss nonparametric identification of θ and $\epsilon_D, \epsilon_0, \epsilon_1$. In appendix F we use an extended model with a nonparametric distributional specification for θ and show that assuming a normal distribution is a sensible approximation.

¹⁹This ‘rotation problem’ of latent factor models occurs since elements of $\alpha = (\alpha_D, \alpha_0, \alpha_1, \alpha_{M1}, \dots, \alpha_{MP})'$ 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.

the union membership equation, such that higher values of θ induce union membership.

Union membership. We specify sorting into union membership via a latent index model (Heckman and Vytlačil 1999, 2007) with a linear-in-parameters formulation.²⁰

$$D^* = \beta_D'Z + \alpha_D\theta + \epsilon_D. \quad (8)$$

$$D = \mathbf{1}(D^* > 0) \quad (9)$$

Here β_D is a parameter vector associated with covariates in Z (more below), while α_D is the latent factor coefficient. Errors ϵ_D are white noise (normalized to have variance one) and assumed orthogonal to Z and θ . We observe union membership D whenever the latent index D^* crosses a threshold (set to zero without loss of generality).

We partition covariates Z into two subvectors $Z = (X, W)$, where X are individual characteristics that also enter the turnout equations while W are variables only used in the turnout equation (referred to as instruments or exclusions). In our application W captures the economic benefits of union membership and, together with the latent factor θ , shifts the propensity score of union membership $P(D = 1|W = w)$.

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 in our model. This, of course, depends crucially on the strength and validity of the chosen instruments. While we show below that both industrial wage differentials and concentration are strongly related to turnout, their exclusion restriction cannot be verified empirically and thus remains subject to (potentially contested) plausibility arguments. For our instruments to be valid, there should be no direct effect of industrial concentration or wage structure on turnout beyond union membership after conditioning on education, race, location, birth cohort and heterogeneity (θ). Note that our model accounts for sorting into industries or regions based on socio-demographic background and earnings potential. While the literature on turnout suggests no obvious direct effect of these industry characteristics on voting, we identify one threat to the exclusion restriction of industrial concentration. If hurdles to overcome free-riding problems for firms interested in political protection are lower in more concentrated industries, one expects more political activity by business in such industries, such as higher campaign contributions (Pittman 1977; Grier, Munger, and Roberts 1994). This may also foster electoral mobilization, including higher levels of turnout. Even though the

²⁰In 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.

empirical evidence linking concentration to corporate political activity is somewhat mixed (see the review in Mao and Zaleski 2001), this argument makes at least plausible the possibility that the exclusion restriction for industrial concentration does not hold.

We address this problem from two angles. First, we provide robustness tests where we study our results when industrial concentration is not included in W . Second, we provide results without both instruments (i.e., $Z = X$). Note that the presence of valid instruments is not required to identify the model—this is achieved by estimating the distribution of the latent factor affecting both union membership and turnout potential outcomes. In the absence of a valid instrument we are reliant on the functional form assumption for θ .²¹ We provide empirical evidence for the plausibility of this functional form assumption, by estimating the distribution of θ nonparametrically (Appendix F).

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

$$\begin{aligned} Y_d^* &= \beta_d' X + \alpha_d \theta + \epsilon_d & d = 0, 1 \\ Y_d &= \mathbf{1}(Y_d^* > 0). \end{aligned} \tag{10}$$

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 .²²

So far, our latent factor potential outcome model contains two independent sources of identification, the factor structure θ and instruments W . Our third source of more robust model identification is created by utilizing the panel-structure of our data to extract auxiliary information, which is independent of treatment status, and which we use to identify variation in θ .

Cognitive tests. 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

²¹In turn, having access to a valid instrument means that θ can be identified nonparametrically. See Heckman (1990) and Heckman and Vytlacil (2007) for conditions for nonparametric identification of θ when instruments are available.

²²To 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 \mathcal{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 , $1 > Pr(D = 1|X) > 0$.

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 ($p = 1, \dots, P$) 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 or the fact that individuals take the test at different ages).²³ Random variables ϵ_{M_p} denote idiosyncratic variation in tests scores that are not explained by background covariates or the latent factor.²⁴

$$M_p = X_p' \omega_p + \alpha_{M_p} \theta + \epsilon_{M_p} \quad p = 1, \dots, P. \quad (11)$$

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 (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 (other treatment parameters can be derived similarly).²⁵ Appendix C provides a formal characterization.

²³Table B.1 gives an overview of all variables used in choice, turnout, and test equations.

²⁴They are assumed to be distributed mean zero with finite variance, $\sigma_{M_p}^2$, 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.

²⁵Note that the structure outlined here can identify *distributions* of treatment effects, a possibility we do not explore in this paper.

Estimation We jointly estimate all treatment, potential outcome, and measurement equations using the Bayesian framework. Note that the model is identified under classical criteria.²⁶ 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 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 to show that our results do not depend on particular prior choices (Lopes and Tobias 2011).²⁷

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.

Parameter estimates

Union membership. Table II shows estimates from the union membership equation (8). 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.²⁸ For easier interpretation of effect sizes, the final column of Table II 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. Conditional on covariates and the latent factor, 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

²⁶Bayesian models for potential outcomes are comparatively rare (but see Heckman, Lopes, and Piatek (2013)). Note 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.

²⁷We 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.

²⁸More 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 II
 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 probabilities of unit change in z
^b Fixed parameter

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 III shows estimates from our measurement system of ASVAB test items, given in equation (11). 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 their ability constant). These distorting influences do indeed exist and are accounted for in our measurement model.

Figure II 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 II

TABLE III
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 ω_p set equal across test items. See appendix A for variable definitions.

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 IV, which shows estimates of the mean of the latent factor 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

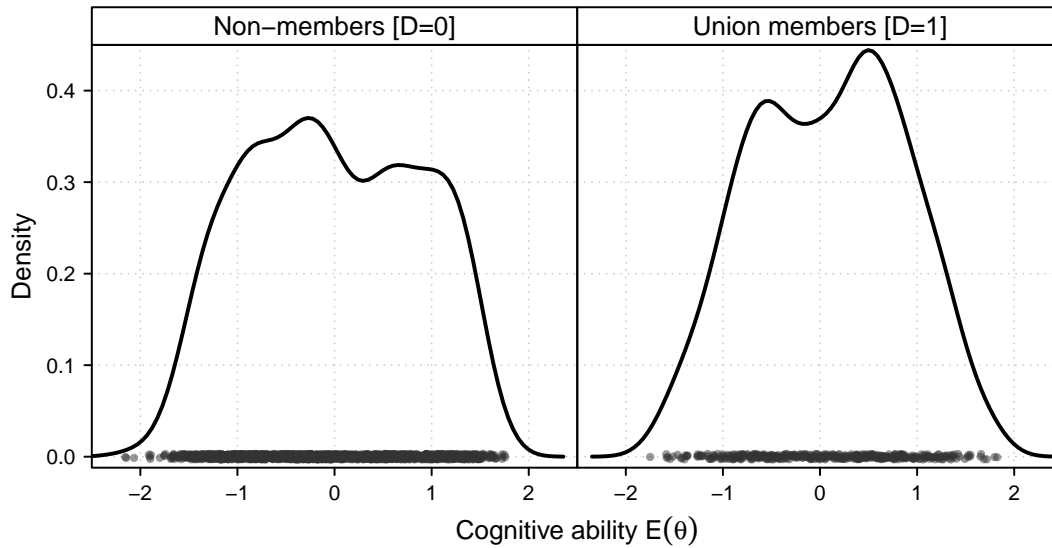


FIGURE II

Distribution of cognitive ability (posterior means of θ) by treatment status

TABLE IV

Means and quantiles of cognitive ability distribution by treatment status

	Union members		Non-members		Difference
Mean	0.079	(0.025)	-0.014	(0.021)	0.093 (0.017)
20th percentile	-0.702	(0.039)	-0.924	(0.029)	0.221 (0.038)
80th 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 α_0 and α_1 in both potential outcome equations. Table V shows posterior parameter summaries for equations (10), 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 unobserved individual heterogeneity θ does indeed substantially affect turnout in both potential outcome states. Both coefficients are of sizable 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.

Another way to illustrate the role played by unobservables is to calculate the covariance or

TABLE V
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.

correlation between them.²⁹ The correlation between unobservables shaping union membership and unobservables shaping turnout for union members, $Corr(U_1, U_D)$, is 0.21, while the their correlation for non-members, $Corr(U_0, U_D)$, is 0.14. This, again, illustrates the importance of accounting for unobservables, which shape both union membership and the propensity to vote on election day.

The role of observed confounders in Table V is as expected from previous research. In particular, higher socio-economic status (income, education) is associated with a higher propensity

²⁹The correlation between U_D and U_0 is given by $\frac{\alpha_0}{\sqrt{2}\sqrt{1+\alpha_0^2}}$ and *mutatis mutandis* for U_1 .

TABLE VI
Effect of union membership on turnout. Average Treatment
Effect estimates from baseline model and under several
alternative specifications.

	Mean	SD	95% HPD	
Baseline model	0.104	0.018	0.069	0.139
<i>Specification checks</i>				
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
<i>Exclusion restrictions</i>				
Only one instrument	0.074	0.030	0.015	0.130
No instruments	0.082	0.030	0.022	0.137

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

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 a specification with potential-outcome-specific covariate effects.

Resulting 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 VI 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 . 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 I) 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, 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 VI. 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).³⁰ 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 VI 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 strongly overlap, suggesting that results do not differ statistically.

While the two instruments provide a useful source of arguably exogenous variation, in the model specification section we have discussed a potential violation of the exclusion restriction with respect to industry concentration. Following a logic of collective action, concentrated industries may see higher political mobilization even in the absence of unions. However, excluding industry concentration from the analysis leads to similar results. The average treatment effect is slightly smaller but remains politically relevant and precisely estimated. While we cannot think of a theory suggesting the industry wage is not a valid instrument, given the model and auxiliary information from test scores we can also identify the effect of union membership on turnout when both instruments are excluded.

CONCLUSION

The effect of labor unions on voting concerns the fundamental relationship between the economic sphere and democratic politics. In this paper, we have used a unique survey data set to provide robust estimates of the causal effect of union membership on turnout in the presence of positive selection. Our empirical approach accounts for the problem labor markets may sort prospective voters into union jobs. To jointly model endogenous union membership and

³⁰We use the North American Industry Classification System (NAICS), 2002 revision, at the 1-digit level.

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 otherwise comparable union members and non-members. This stands in contrast to previous studies, which have mostly assumed that economic sorting is not a problem. 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, and shows how it can be addressed empirically.

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.
- Abowd, J. M., and H. S. Farber. 1982. "Job queues and the union status of workers." *Industrial & Labor Relations Review* 35 (3): 354–367.
- 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.
- 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.
- 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.
- Anzia, S. F. 2012. "The Election Timing Effect: Evidence from a Policy Intervention in Texas." *Quarterly Journal of Political Science*. 7 (3): 209-248.
- Bartels, L. 2008. *Unequal Democracy*. Princeton and Oxford: Princeton University Press.
- 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.
- Becher, M., D. Stegmüller, and K. Kaepfner. 2018. "Local Union Organization and Law Making in the US Congress." *Journal of Politics* 80 (2): 39-554.
- 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., 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.
- Feigenbaum, J., A. Hertel-Fernandez, and V. Williamson. 2018. "From the Bargaining Table to the Ballot Box: Political Effects of Right to Work Laws." NBER Working Paper 24259.
- Ferguson, T. S. 1983. "Bayesian density estimation by mixtures of normal distributions." In *Recent Advances in Statistics*, ed. M. Rizvi, J. Rustagi, and D. Siegmund. Vol. 24. New York: Academic Press.
- 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.
- Grier, K. B., M. C. Munger, and B. E. Roberts. 1994. "The determinants of industry political activity, 1978–1986." *American Political Science Review* 88 (4): 911–926.

- Hauser, S. M. 2000. "Education, ability, and civic engagement in the contemporary United States." *Social Science Research* 29 (4): 556–582.
- Heckman, J. J. 1976. "The common structure of statistical models of truncation, sample selection and limited dependent variables and a simple estimator for such models." *Annals of Economic and Social Measurement* 5 (4): 475–492.
- 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. Vytlačil. 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. Vytlačil. 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.
- Heckman, J. J., J. Smith, and N. Clements. 1997. "Making the most out of programme evaluations and social experiments: Accounting for heterogeneity in programme impacts." *Review of Economic Studies* 64 (4): 487–535.
- Herrera, H., M. Morelli, and S. Nunnari. 2016. "Turnout Across Democracies." *American Journal of Political Science* 60 (3): 607–624.
- 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. 2017. "Informed Preferences? The Impact of Unions on Workers' Policy Views." *American Journal of Political Science* 61: 728-743.
- Koop, G., and D. J. Poirier. 1997. "Learning about the across-regime correlation in switching regression models." *Journal of Econometrics* 78 (1): 217-227.
- 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.
- Lee, L.-F. 1978. "Unionism and Wage Rates: A Simultaneous Equations Model with Qualitative and Limited-Dependent Variables." *International Economic Review* 19 (2): 415-433.
- 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.
- Lopes, H. F., and J. L. Tobias. 2011. "Confronting prior convictions: On issues of prior sensitivity and likelihood robustness in Bayesian analysis." *Annual Review of Economics* 3 (1): 107-131.
- Luskin, R. C. 1990. "Explaining political sophistication." *Political Behavior* 12 (4): 331-361.
- Maddala, G. 1986. *Limited-dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.
- Mao, W., and P. Zaleski. 2001. "The effect of industry concentration on free riding." *Review of Industrial Organization* 19 (3): 295-303.
- 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.

- Pittman, R. 1977. "Market structure and campaign contributions." *Public Choice* 31 (1): 37–52.
- Poirier, D. J. 1998. "Revising beliefs in nonidentified models." *Econometric Theory* 14 (4): 483–509.
- Poirier, D. J., and J. L. Tobias. 2003. "On the predictive distributions of outcome gains in the presence of an unidentified parameter." *Journal of Business & Economic Statistics* 21 (2): 258–268.
- 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.
- Samii, C. 2016. "Causal Empiricism in Quantitative Research." *The Journal of Politics* 78 (3): 941–955.
- 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.
- 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.
- Vytlačil, 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.

APPENDIX

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 1980.³¹ *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 the *Union wage differential* are calculated from administrative sources. We calculate union-nonunion wage differentials from the Bureau of Labor Statistics’ CPS-LU series, for 19 major industrial sectors. Our measure 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. Industrial concentration is captured using the CR4 concentration ratio from the Census Bureau’s Economic Census of American businesses conducted in 2007 for 243 industries. 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.

³¹We 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. Difference in median
 weekly earnings (in contemporary \$)

Industry	Difference
Finance and insurance	-97
Agriculture, forestry, fishing, and hunting	0 ^a
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 Not calculated since sample size less than 50,000.

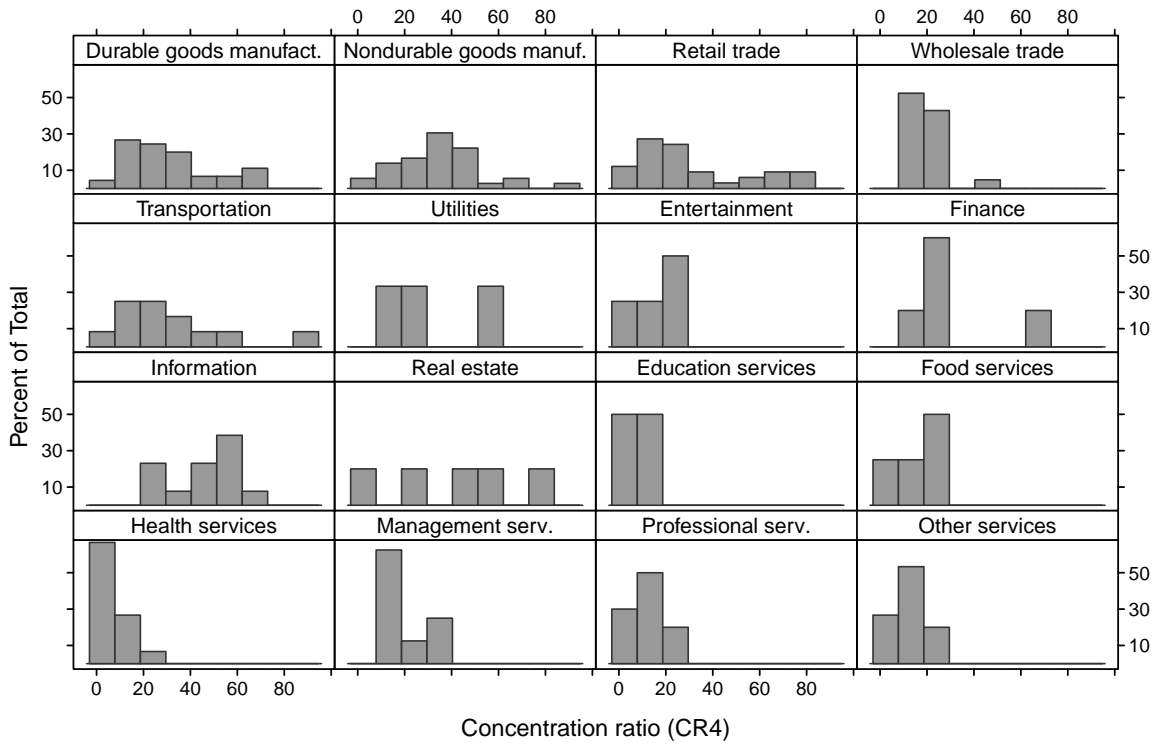


FIGURE A.1
Industry concentration by major economic sectors

B. MODEL EQUATIONS

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

$$\begin{aligned} ATE(X, \Gamma) &= Pr(Y_1 = 1|X, \Gamma) - Pr(Y_0 = 1|X, \Gamma) \\ &= \Phi\left(\frac{\beta'_1 X}{\sqrt{1 + \alpha'_1 \alpha_1}}\right) - \Phi\left(\frac{\beta'_0 X}{\sqrt{1 + \alpha'_0 \alpha_0}}\right), \end{aligned} \quad (C.1)$$

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 $D = 1$:

$$\begin{aligned} TT(X, \Gamma, D = 1) &= Pr(Y_1 = 1|X, D = 1, \Gamma) - Pr(Y_0 = 1|X, D = 1, \Gamma) \\ &= \left(\frac{\Phi(\beta'_D Z)}{\sqrt{1 + \alpha'_D \alpha_D}}\right)^{-1} \int [\Phi(\beta'_1 X + \alpha_1 \theta) - \Phi(\beta'_0 X + \alpha_0 \theta)] \\ &\quad \times \Phi(\beta'_D Z + \alpha_D \theta) \phi(\theta) d\theta \end{aligned} \quad (C.2)$$

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. PRIOR DISTRIBUTIONS

We assume independent priors for factor coefficients in potential outcome equations $\alpha_j \sim N(\tilde{\alpha}_j, \tilde{v}_j)$, $j = 0, 1$. 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 $\tilde{B}_j = \mathbf{I}_j \tilde{b}_j$. Finally, we use normal priors for θ -coefficients in our measurement equations, $\omega_p \sim N(\tilde{\omega}_p, \tilde{v}_p)$, and we use normal priors for covariates in these equations as well: $\alpha_{M_p} \sim N(\tilde{\alpha}_{M_p}, \tilde{v}_{M_p})$. 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
$\tilde{\nu}_j$	$j = 0, 1$	10	10	100
$\tilde{\beta}_{jk}$	$j = 0, 1, D; k = 1, \dots, K$	0	0	0
\tilde{b}_{jk}	$j = 0, 1, D; k = 1, \dots, K$	10	10	100
$\tilde{\omega}_p$	$p = 1, \dots, P$	0	0	0
$\tilde{\nu}_p$	$p = 1, \dots, P$	10	10	100
$\tilde{\alpha}_{M_p}$	$p = 1, \dots, P$	0	0	0
$\tilde{\nu}_{M_p}$	$p = 1, \dots, P$	10	10	100
\tilde{a}_p	$p = 1, \dots, P$	1	1	1
\tilde{b}_p	$p = 1, \dots, P$	2	0.005	2

TABLE D.2
Prior robustness checks. Posterior summary of average treatment effects under alternative prior parameter values.

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

E. SIGNAL-TO-NOISE RATIO OF ASVAB TEST ITEMS

The core idea of the model in (11) is that each observed ASVAB item contributes information to identify the latent factor. We don't use the total variation of each cognitive measurement since a portion of the variation in test items over individuals is due to measurement error (thus, creating a simple summed measure of all items mistakes the test for the construct being measured). We can calculate the fraction of the variance of each item that is due to measurement error (noise) and underlying latent factor (signal), respectively (see, e.g., Cunha, Heckman, and Schennach 2010: 908). Table E.1 shows these quantities together with resulting signal-to-noise ratios.

TABLE E.1
Signal-to-noise ratio of ASVAB test items

	Signal	Noise	SNR
Arithmetic reasoning	0.813	0.187	4.364 (0.321)
Word knowledge	0.724	0.276	2.640 (0.188)
Paragraph comprehension	0.667	0.333	2.009 (0.151)
Coding speed	0.472	0.528	0.895 (0.065)
Math knowledge	0.745	0.255	2.927 (0.198)

Note: Standard error of SNR estimate in parentheses

We find that four items have large signal-to-noise ratios, indicating that their variation provides substantial information for estimating the latent factor. More than two thirds of the total variation in arithmetic reasoning, math and word knowledge, and paragraph comprehension is due to variation in θ . Slightly less than half of individual differences in coding speed are due to θ leading to a signal-to-noise ratio of less than one.

F. NONPARAMETRIC ESTIMATES OF θ

A critical reader might argue that assuming a normal distribution for θ is an arbitrary choice. While there is research arguing for this choice on the grounds of its robustness against misspecification note that this assumption is not necessary for our model. Jettisoning the

normal distribution assumption, we also estimate $f(\boldsymbol{\theta})$ semi-parametrically by approximating it using a finite mixture of normals (Ferguson 1983):

$$f(\boldsymbol{\theta}) = \sum_{k=1}^K \pi_k \phi(\boldsymbol{\theta} | \mu_k, \sigma^2)$$

Here $\phi(\cdot)$ is the normal density, and $\boldsymbol{\pi} = (\pi_1, \dots, \pi_K)'$ is discrete with mass at (μ_k, σ^2) . To insure identification $\sum_k \pi_k = 1$. Even with very few mass points (K being as little as two) finite normal mixtures are flexible enough to approximate a wide number of densities. (see Skrondal and Rabe-Hesketh (2004: ch.6) for an extended discussion). In our empirical application we set $K = 3$ to produce a non-parametric estimate of $f(\boldsymbol{\theta})$ in addition to the one based on the normal distribution described above.

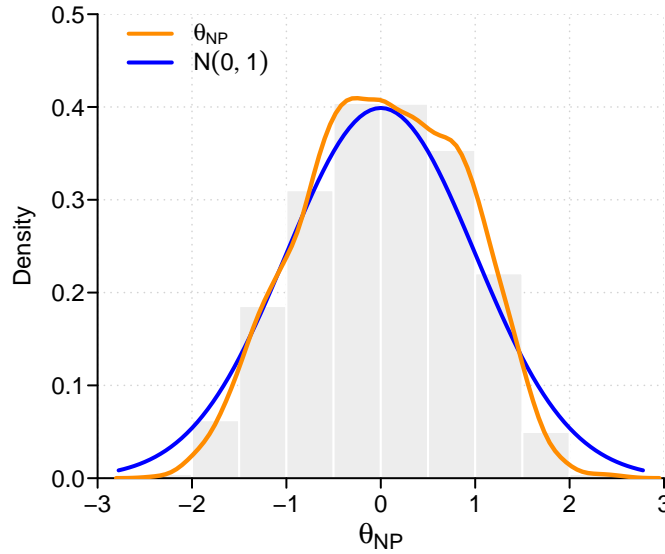


FIGURE F.1

Comparison of parametric and nonparametric specification of θ .

This figure shows a histogram of the posterior distribution of θ estimated using a finite mixture of normals with a kernel density estimate (using a Gaussian kernel with bandwidth 0.14 evaluated over 100-point grid) added. For comparison, we also add the standard normal distribution implied by our main model specification.

Figure F.1 plots the distribution of θ estimated non-parametrically using a histogram and a kernel density estimate. It also contains a standard normal distribution as reference point. Comparing both, we find the distributional assumption we make in our main specification to be quite reasonable.