Do Voters Join Unions or Do Unions Encourage Voting?

Daniel Stegmueller
University of Essex
mail@daniel-stegmueller.com

Michael Becher
University of Konstanz
michael.becher@uni-konstanz.de

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 (2014).

Abstract

What is the causal effect of group membership on voter turnout? Extensive research on voting has documented a positive association between membership in groups like unions and participation in elections. But it is widely recognized that existing studies provide limited evidence on whether this reflects a causal relationship or is due to self-selection into union membership. We address this question using panel data from the US, the National Longitudinal Survey of Youth, which allows us to exploit credible strategies to deal with unobserved confounders. Building on recent advances in econometrics, we use a potential outcomes model with matching on both observable and unobservable individual characteristics. Unobservable characteristics are captured using a latent factor structure, which allows for proxies measured with error. We implement the model in a Bayesian framework. We find a clear average treatment effect of union membership. All else equal, union member are 10 percentage points more likely to turn-out on election day.

1. Introduction

Does group membership have a causal effect on voting or does it merely reflect citizens' prior predispositions to join groups or be active in politics? Recent theoretical research in the rational choice framework has converged on group models to explain why citizens often vote in large election despite a minuscule chance of influencing the result (e.g., Aldrich 1993; Coate and Conlin 2004; Cox 1999; Cox *et al.* 1998; Feddersen and Sandroni 2006; Morton 1991; Shachar and Nalebuff 1999; Uhlaner 1989). As groups with a large membership have a considerably larger chance of being pivotal in an election than individuals, group leaders often have strategic incentives to mobilize their (non-strategic) members in order to influence election outcomes. Assuming that members base their turnout decision on the cost of voting and considerations of civic duty, leaders can affect those parameters through mobilization efforts. Group altruism may also lead to group-based turnout (Coate and Conlin 2004; Feddersen and Sandroni 2006). Group models can explain apparently strategic behavior at the macro level, such as the the positive correlation between turnout and the marginality of elections. One fundamental challenge of this theoretical perspective, as the review of Feddersen (2004: p. 448) points out, is to account for why individuals join certain groups to begin with. Existing models take group membership as given. The obvious problem for researchers interested in estimating the impact of group membership on turnout is that membership may be endogenous to underlying values that also shape voting behavior, such as civic duty or policy preferences, and are difficult to measure. This endogeneity problem has long been recognized in the empirical literature on groups and political participation. For example, Leighley (1996: p. 448) writes that "mobilization effects identified in previous studies may simply reflect the fact that individuals who belong to and are active in groups also generally participate more in politics, regardless of group dynamics." This problem cannot be fully solved using control variables. One reason is that possible selection variables, like ability or preferences, are typically not measured in surveys on the issue. And more easily measured factors, such as political interest, may be a consequence rather than a cause of membership or reflect a complex dynamic relationship.

In this paper, we address the endogeneity issue by examining membership in a particular group. Labor unions provide the opportunity to exploit individuals' non-political incentives to join a group and thereby address the endogeneity problem that has hampered empirical (and theoretical) research on groups and turnout. More specifically, we combine the National Longitudinal Study of Youth (NLSY), which contains a unique battery of

cognitive variables that help model individual heterogeneity, with industry data capturing economic incentives of joining a union and analyze them with the appropriate statistical model that has been ignored by the literature on turnout (and perhaps political science more general). The statistical approach simultaneously models the decisions to join (or remain member of) a union and whether to vote or abstain in the 2006 congressional election and implements a potential outcomes model with matching on observable and unobservable individual characteristics. To capture non-political incentives to become a union member, we rely on the wage differential between members and non-members in a particular industry as well as industrial concentration.[1] Building on recent advances in econometrics and statistics, the model also specifies a latent factor structure that allows the unobserved (by the researchers) errors concerning the decision to join a union and the decision to vote to be correlated (Aakvik *et al.* 2005; Abbring and Heckman 2007). This approach is significantly strengthened by drawing on the rich data of the NLSY. In particular, we rely on a battery of cognitive ability tests to measure the underlying latent variable, cognitive ability, that explains part of the (otherwise) unobserved heterogeneity. Conceptually, cognitive ability is closely linked to the motivations and resources that have been shown to shape political participation by a large body of research in political science (e.g., Verba *et al.* 1995). Altogether, the statistical approach exploits three distinct sources of causal identification (economic incentives, factor structure, cognitive tests). To the best of our knowledge, this type of model has not yet been applied in this context.[2]

To preview, we find 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 9-10 percentage points. Thus in the population we have studied, an important part of the well-documented union gap in turnout is not driven by self-selection. This magnitude is politically significant even in times of declining union membership. It is consistent with group based models of mobilization. The analysis also reveals, however, that there is selection into unions by individuals with higher abilities that are also more likely to vote otherwise. Our results suggests that 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. Thus, failing to account for this selection

[1] Some scholars argue that union membership is plausibly exogenous to political participation as the decision to join is mainly economic (e.g., Kerrissey and Schofer 2013: p. 918). We completely agree that exploiting non-political incentives to join groups is a good idea, but we think one has to explicitly model them rather than assume that they trump selection effects.

[2] For applications in economics, see, e.g., Aakvik *et al.* 2005, Carneiro *et al.* 2003, and Cunha *et al.* 2005. See Abbring and Heckman 2007 for an extensive overview.

effect would generate upward biased estimates of the impact of unions on turnout.

The plan of the paper is as follows. Section 2 briefly discusses the related empirical literature on turnout and why it matters for political representation. Section 3 conceptualizes the two decision problems we study, whether to become a union member and whether to vote, to motivate the statistical specification. Before we go into the details of the model in Section in Section 5, Section 4 describes the data. The results are presented in Sections 6 and 7. The concluding remarks sum up the results and their relevance.

2. Empirical studies of unions, turnout, and representation

While landmark studies of turnout have not paid much attention to unions (Rosenstone and Hansen 1993; Powell 1986), by now multiple studies in the US have found that individual union membership is positively associated with a higher propensity to vote (Freeman 2003; Kerrissey and Schofer 2013; Leighley and Nagler 2007; Radcliff 2001; Rosenfeld 2010). Studies examining aggregate data at the the state or county level also find a positive relationship between unionization and turnout (Radcliff and Davis 2000; Zullo 2008). It is well documented that unions try to get out the vote and there is some direct evidence based on contact data that they do so successfully (Lamare 2010; Zullo 2004).[3] There also is growing evidence for a union-turnout link in other advanced industrial democracies (D'Art and Turner 2007; Flavin and Radcliff 2011; Gray and Caul 2000; Radcliff and Davis 2000). Thus, the growing body of empirical studies has significantly enhanced our knowledge of unions and turnout. Yet it has not fully addressed the endogeneity problem. Existing-individual level studies do not account for unobserved individual characteristics that may drive both union membership and turnout. Controlling for observables already indicates that part of the raw differences in turnout between union members and non-members is driven by selection. Using Current Population Survey data covering the 1990-2000 period, Freeman (2003) finds that descriptive turnout differences are large, with members being about 10 percent more likely to vote than non-members. Adding demographic variables, family income and job characteristics reduces the turnout gap to about 4 percent, which

[3]For example, Freeman (2003: p. 1) reports that in the 2000 presidential election "AFL-CIO unions made 8 million phone calls to members, sent out 12 million pieces of mail, distributed 14 million leaflets at union workplaces once a week from September to Election day and spent more than $43 million to help win a popular vote victory for the Democratic presidential candidate."

may nonetheless still be politically significant. Freeman (2003: p. 22) concludes that "Much of the union/ nonunion voting gap is due to the differential characteristics of union and nonunion workers." This underlines the question we examine in this paper: To what degree does the observed relationship between union membership and turnout reflect a causal effect of unions rather than self-selection?

Answering this question does not only further clarify the empirical micro-foundations for group based theories of turnout. It also sheds light on important debates about the state of American democracy. First, the decline in union membership may account for a part of the decline in turnout in the US, which is one of the enduring puzzles in American politics (Aldrich 1993). If union membership has a causal effect on turnout, then the decline in union membership contributes rather than merely correlates to low turnout. Second, lower union mobilization may lead to more unequal turnout and more unequal responsiveness by politicians. The resulting policies increase economic inequality and perhaps reduce social mobility. Admittedly, this is a longer causal change whose empirical examination is far beyond the scope of this paper. But recent research that it cannot be rejected out of hand. There is scholarly debate about whether unequal turnout actually matters for representation as some studies find that the preferences of voters and non-voters are quite similar (but see the comprehensive study of Leighley and Nagler 2013, which finds that voters are economically more conservative than non-voters). As Schlozman *et al.* (2012: p. 567) point out, unions are one of the few groups that has consistently mobilized those less likely to turn out. Hence even if mobilization by parties or ordinary citizens in general does not reduce inequality in participation (or may actually increase it) due to rational prospecting (Schlozman *et al.* 2012: ch. 15; Enos *et al.* 2014), mobilization by unions is distinct in that it has naturally focused on a segment of the population that a priori is less likely to participate. Moreover, recent studies find evidence that higher country-level turnout in American elections tends to benefit presidential candidates of the Democratic Party (Hansford and Gomez 2010), though some studies find no similar effect at the state level (Erikson 1995), and that it is associated with more egalitarian economic policies (Hill *et al.* 1995).[4] Marginal voters, who are most likely to be swayed by mobilization efforts, tend to be more likely to be supporters of the Democratic Party compared to regular voters (Fowler 2012). It is well established that Democratic legislators are more likely than Republican legislators to support egalitarian policies (McCarty *et al.* 2006; Lee *et al.*

[4]For a review, see Martinez (2010). There are some related findings in the comparative politics literature (Fowler 2013; Mahler 2008).

2004) and that the partisan control of the presidency is linked to economic police and outcomes (Bartels 2008). In the last decades, economic inequality has increases while union membership has declined (e.g., see Schlozman *et al.* 2012: ch. 3). This may not be coincidental. Economists, sociologists and political scientists debate the causes of rising income inequality. While the literature has not identified a single smoking gun (we think there is unlikely to be one), unions may be an important factor. The study of Western and Rosenfeld (2011) suggests that declining unionization has led to weaker norms of wage inequality. Declining mobilization of lower income citizen and subsequent biases in representation may be another channel.

3. Union membership and turnout

To understand the effect of union membership on turnout, we need to model two choices simultaneously. The decision to join, or remain member of, a union, and the decision to turn out on election day.

Union membership We model union membership as a choice that is mainly influenced by economic factors. Let D_i^* be the *latent utility* of union membership for an individual, which is determined by a vector of observed individual characteristics, Z_i, and by an unobserved random variable, U_{Di},

$$D_i^* = \mu_D(Z_i) - U_{Di}. \tag{1}$$

The vector of individual characteristics, Z_i, is composed of two parts. First, it includes basic individual background characteristics (or pre-treatment covariates), such as age and gender (collected in X_i). Second, Z_i contains one or more variables, which make union membership an economically attractive option, but are unrelated to election turnout. We discuss two such variables below. The term $\mu_D(Z_i) - U_{Di}$ can be interpreted as the *net utility* of union membership for an individual with characteristics (Z_i, U_{Di}). Only if net utility is positive will he or she join a union. Thus, the decision rule becomes

$$D_i = 1 \text{ if } D_i^* \geq 0,\ D_i = 0 \text{ otherwise.} \tag{2}$$

The assumptions we employ when using this latent index structure are the same as those imposed in the LATE framework of Imbens and Angrist (1994).[5]

Which systematic factors induce individuals to join unions? To put it differently, think of two individuals with identical sets of X_i characteristics and whose values of U_{Di} make them just indifferent between joining a union or not. An increase in Z_i for the first individual means that he becomes a union member, all else equal. Research in economics and political science has produced a number of factors; among the most relevant are firm and industry characteristics (Hirsch and Berger 1984). We focus on industry union–non-union wage differentials (e.g., Budd and Na 2000) and on concentration levels in a worker's industry (e.g., Belman 1988).

In a basic model of an optimizing worker, union membership is treated as simply as an "asset in the portfolio of the worker" (Pencavel 1971: 180) with associated (private and/or collective) benefits, such as higher wages, better grievance procedures etc., as well as costs (in the form of membership dues). In this framework, the higher the union–non-union wage differential, the higher the benefits of union membership (all else equal), and the more likely it is that we find a worker being a union member (Schnabel 2003: 14).[6] However, wage differentials are not easy to measure and do not represent the full picture of union benefits.[7] Therefore, we also include industry concentration, which is a central determinant of unionization levels (union organization is easier in sectors with four firms than with 50; Stephens and Wallerstein 1991: 943) and its resulting benefits for workers (Hirsch and Berger 1984). Higher levels of industrial concentration are thus connected with higher wages for union members (Kwoka 1983) as well as higher provision of fringe benefits (Freeman 1980). Thus, we argue that net of observed and unobserved worker characteristics, employment in higher concentrated industries induces union membership.

Turnout We model the effect of union membership on turnout in an explicit potential outcomes framework (Rubin 1978; see Morgan and Winship 2007 for an introduction).[8] We can only ever observe an individual in one of two possible states: being a member or not being a member of a trade union in a given election. Thus, the propensity to turn out on election day for each individual is *potential* rather than a realized outcome.

[5] Vytlacil (2002) provides an equivalence proof between such a latent index structure with $U_D \perp Z$ and the assumptions imposed in the LATE framework.

[6] For our argument we remain agnostic about *why* wage differentials exist. They might be due to straightforward union rent extraction (Johnson 1975), firm-worker matching on productivity (Lee 1978), or might represent compensatory payments for work conditions (Duncan and Stafford 1980).

[7] We describe our measurements in more detail in the following section.

[8] This setup is also related to the switching regression framework of Quandt (1972).

Denote the two potential outcomes – turnout propensity – in our two counterfactual states – union member, non-member – by Y_{1i} and Y_{0i}. In other words, if an individual were a union member ($D_i = 1$), we would observe Y_{1i}; if he were not a member ($D_i = 0$), we would observe Y_{0i}. The core quantity of interest of this paper, the effect of union membership on turnout, is the average treatment effect $\Delta = E(Y_{1i} - Y_{0i})$.

Both potential outcomes can be written as a mean given a vector of observed individual characteristics or controls, $\mu(X_i)$, such as age or education. Furthermore, we need to control for unobserved effects. Even with a fully specified matrix of covariates, X_i, the majority of the variance of an individual's turnout decision is due to unobservables. For example, one factor mentioned prominently in the literature on turnout is cognitive ability (Luskin 1990; Verba *et al.* 1995; Nie *et al.* 1996; Hauser 2000). It systematically influences turnout by shaping education, civic skills, and political interest and sophistication (Denny and Doyle 2008: 294). The influence of such unobservables is captured by including individual random variables U_i. Then, the two equations for potential turnout outcomes are:

$$\begin{aligned} Y_{0i} &= \mu_0(X_i) + U_{0i} \qquad \text{if } D_i = 0 \\ Y_{1i} &= \mu_1(X_i) + U_{1i} \qquad \text{if } D_i = 1. \end{aligned} \tag{3}$$

Note that unobservables U_{0i} and U_{1i} are allowed to be correlated with unobservables in the union choice equation, U_{Di}. This captures the fact that unobserved individual characteristics, say cognitive ability, might push individuals to join unions and participate in elections. We discuss how we transform this conceptual model into an empirically estimable setup below, after we describe the unique data set that enables our analysis.

4. 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.[9] It is widely used in economics due to the high quality of its sample design, data collection, and the availability of cognitive measurements (e.g., Lochner and Monge-Naranjo 2011; Lang and Manove 2011). Due to its mission the NLSY does not include political questions. However, in a unique collaboration the American National Election Study was able to place a short set of political items in the NLSY 2008 wave, including the election turnout item asked in each ANES survey (Krosnick and Lupia 2006). We make use of this unique data-set

[9] See www.bls.gov/nls.

(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: one recent panel 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 (we discard the military subsample). To work with a homogeneous sample we focus on male respondents. Due due to cohort design of the NLSY, they are aged between 41 and 50 in 2006. We matched 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 Industry concentration has a long history in economics and comparative politics. It is usually measured by how much of a sectors revenue is generated by the four largest firms (Pryor 1972; Stephens and Wallerstein 1991), the so called CR_4 concentration ratio. Concentration ratios are coded from the US Bureau of Labor's Economic census of 2007, which provides a large random sample of US firms. The high quality of the data allows us to use disaggregated concentration ratios for 243 industries. Figure 1 plots histograms of concentration ratios within major industrial sectors.

We calculate union-nonunion wage differences from the US Bureau of Labor's 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 used a lower resolution than for our concentration measure and calculated wage differentials for 19 major sectors. See appendix Table A.1 for wage differences by industry.

Cognitive ability tests In 1980, the Departments of Defense and Labor jointly sponsored the administration of the Armed Services Vocational Aptitude Battery (ASVAB, cf. Jensen 1985) to the civilian and military NLSY79 samples.[10] The ASVAB consists of several subtests that measure vocational aptitude in areas such as arithmetic reasoning, coding speed, and 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. In other words, we properly account

[10]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.

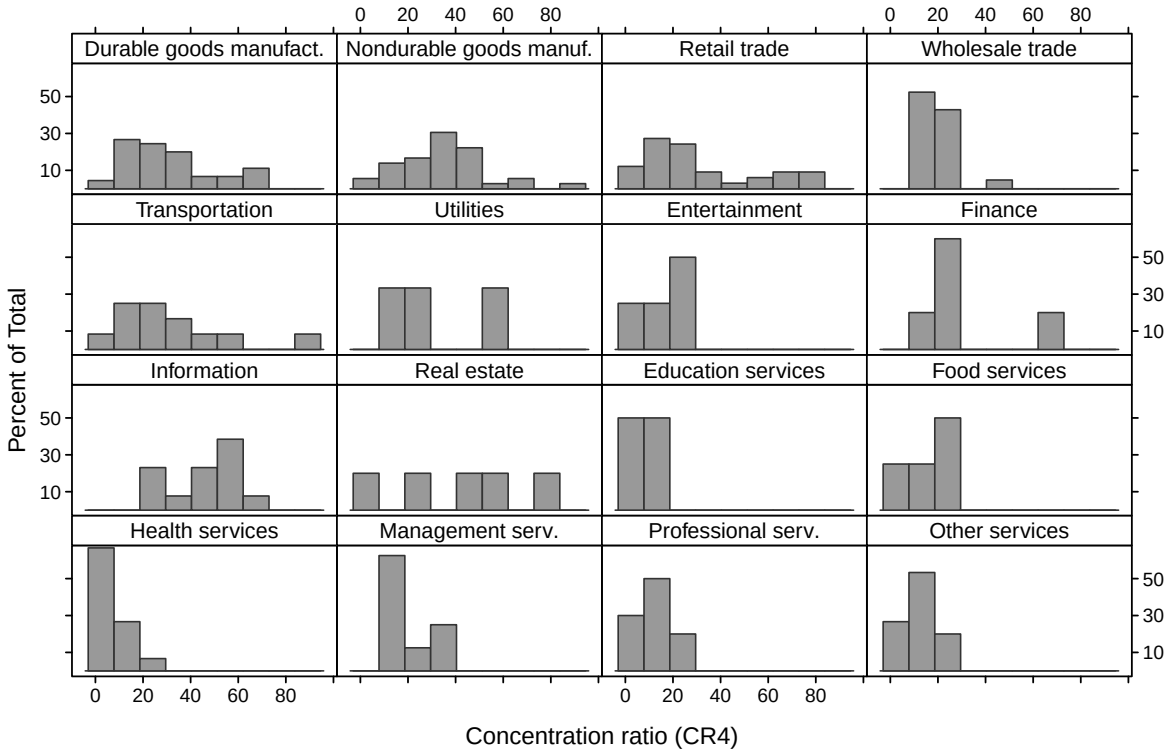

Figure 1: Industry concentration by major sectors

for the fact that realized test scores and general cognitive capability are not the same thing.

Turnout After the November election in 2006 respondents were asked if they voted in that election. To reduce over-reporting respondents had more than one option to indicate non-turnout: "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".[11] Turnout was indicated by the response "I am sure I voted". We create an indicator variable equal to one if a respondent chose the last option and zero otherwise.[12]

Controls We include a number of additional variables to capture heterogeneity between individuals. A respondent's education is measured by years of schooling. We include a measure of family size, as well as an indicator variable equal to one if a respondent is married. Two further indicator variables capture if a respondent lives in the South and/or a rural area. We also include indicator variables for Black and Hispanic, in order to capture well known turnout differences for minority groups. Finally, we create a measure of previous unemployment experience equal to one if a respondent experienced a spell of unemployment in the previous calendar year. To account for the cohort design of the NLSY, we also create a set of cohort dummies to capture systematic cohort differences between individuals in our sample. Table 1 provides descriptive means of our central variables for union members and non-members.

5. Model

5.1. The identification problem

As discussed above, we have two potential outcomes: turnout for union members, Y_1, and turnout for non-members, Y_0. For each individual, we assume that the pair (Y_0, Y_1)

[11] 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%.

[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".

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

[a] Weeks unemployed in past calendar year

[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.

exists, but we can of course only ever observe one possible state per individual.[13] Thus, our actually observed outcome, Y, is

$$Y = DY_1 + (1-D)Y_0. \tag{4}$$

The key parameter we are interested in is the turnout enhancing (or depressing) effect of union membership $\Delta \equiv Y_1 - Y_0$. In other words we are interested in the *ceteris paribus* effect of moving an otherwise identical individual into union membership. Thus for each individual in state $d = (0,1)$, i.e, being either union member or non-member, we need to identify his or her potential outcome in the alternative state. This *counterfactual* outcome is unobserved.

A straightforward approach would be to simply compare outcomes between groups of union members and non-members, be it by using simple group means or regression models. The problem with this naive strategy is that it ignores the underlying selection

[13] In the following we assume that we have access to an equiprobability sample of individuals and thus suppress i subscripts.

process. Union members and non-members are not the same. As is well-known to practitioners, even models containing large numbers of observed individual characteristics explain only a small fraction of the observed variance in individual behavior. The bulk of unobserved individual characteristics affecting union membership choice and turnout are unobserved to the researcher. These unobserved characteristics are captured by the random vector (U_D, U_1, U_0). If we allow for correlations between unobservables, i.e., allowing that unobserved factors that influence an individual's union memberships also shape turnout, we obtain a 3×3 variance-covariance matrix:

$$\text{Cov} \begin{bmatrix} U_D \\ U_1 \\ U_0 \end{bmatrix} = \begin{bmatrix} \sigma_D^2 & \rho_{D1}\sigma_1 & \rho_{D0}\sigma_0 \\ \rho_{D1}\sigma_1 & \sigma_1^2 & \rho_{10}\sigma_1\sigma_0 \\ \rho_{D0}\sigma_0 & \rho_{10}\sigma_1\sigma_0 & \sigma_0^2 \end{bmatrix}. \tag{5}$$

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, where ρ_{jk} represents the correlation between U_j, U_k. Since we can never observe the same individual in two different states at once, the correlation between both potential outcomes, ρ_{10} is not identified (Vijverberg 1993). This is the fundamental problem of causal inference.

5.2. Latent variable potential outcome model

We will now discuss a possible solution to this identification problem, which is based on a combination of additional statistical structure and available auxiliary information. First, note that a joint model of potential outcomes and union membership does not contain any information about the correlation between potential outcomes, (Y_0, Y_1). All that 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 (5) using an underlying low-dimensional set of random factors (cf. Heckman 1981). Thus, following Aakvik *et al.*

(2005), we decompose unobservables using the following factor structure:[14]

$$U_D = \alpha_D \theta + \epsilon_D \tag{6}$$

$$U_0 = \alpha_0 \theta + \epsilon_0 \tag{7}$$

$$U_1 = \alpha_1 \theta + \epsilon_1 \tag{8}$$

Here, θ is a latent factor or random effect, which represents unobserved individual characteristics, such as ability, which systematically shape both utility of union membership and the propensity to turn out at election day. Note that θ affects union choice and outcomes differently. It influences unobservables in both union membership as well as (potential) turnout equations via coefficients α_D and α_0, α_1, respectively (see McFadden 1984; Cameron and Heckman 1998; see Skrondal and Rabe-Hesketh 2004 for a general introduction to random effect and factor models). We assume that θ is normalized to have zero mean with finite variance. In the current application, we specify $\theta \sim N(0,1)$ – a distributional choice which is convenient and quite robust against misspecification (cf. Bartholomew 1988; Neuhaus *et al.* 1992; Wedel and Kamakura 2001).[15]

This structure solves the core identification problem by inducing dependency between potential outcomes and treatment equation (Carneiro *et al.* 2003; Aakvik *et al.* 2005).[16] To see this, note that the non-identified parameter, $\text{Cov}(Y_0, Y_1) \equiv \rho_{10}$, can be recovered from the factor loadings as $\rho_{10} = \alpha_0 \alpha_1$. In other words, the latent factor is assumed to generate the correlation between observed and 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. However, we prefer a more robust identification strategy, and therefore describe two other pieces of information, which help us identify the effect of union membership on turnout. Before doing that, we need to

[14] See Carneiro *et al.* (2003: 369) for an application using a multidimensional factor structure.

[15] We discuss sources of robustness against this distributional assumption on θ below. Note that assuming normality is convenient but not necessary for identification. Cunha *et al.* (2005) discuss nonparametric identification of θ and $\epsilon_D, \epsilon_0, \epsilon_1$.

[16] Another 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 *et al.* 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 often unrealistic. Thus, these approaches make inference in an applied setting, such as ours, problematic.

detail our our specifications of turnout and union membership equations.

We write the choice of union membership as latent index model (Heckman and Vytlacil 1999, 2007) with a linear-in-parameters formulation.[17] We specify observed individual characteristics (confounders), X, a set of variables Z capturing the pure benefits of union membership and excluded from turnout equations (more below), and the latent factor θ as influencing the continuous latent utility of union membership, D^* (Maddala 1986):

$$D^* = \beta_D' Z + \alpha_D \theta + \epsilon_D \tag{9}$$

$$D = \mathbf{1}(D^* > 0) \tag{10}$$

Here β is a parameter vector associated with covariates and exclusions in Z, while α_D is a latent factor loading in the treatment equation. Errors ϵ_D are white noise (normalized to have variance one) and assumed $\epsilon_D \perp X, Z, \theta$. Latent utility D^* is created via a threshold-crossing or latent index model (Heckman and Vytlacil 1999), $\mathbf{1}(D^* > 0)$, where $\mathbf{1}(\cdot)$ is an indicator function evaluating to one if its argument is true and zero otherwise.

The fact that we have variables in Z that encourage an individual to become a union membe (ceteris paribus), but which are unrelated to turnout decisions, provides an additional source of identification. Of course, good exclusion restrictions are hard to find, and almost always hotly contested. But notice that the validity of this exclusion (or instrument) 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. Heckman (1990) and Heckman and Vytlacil (2007) prove conditions for nonparametric identification of θ when instruments are available.

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:

$$\begin{aligned} Y_d^* &= \beta_d X + \alpha_d \theta + \epsilon_d \qquad d = 0, 1 \\ Y_d &= \mathbf{1}(Y_d^* > 0) \end{aligned} \tag{11}$$

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. Latent errors ϵ_d are assumed to be independent of observed covariates

[17] In other words, we set $\mu(Z) = Z\gamma$. For a discussion of linear-in-parameters specifications in choice models, see, e.g, Eckstein and Wolpin (1989)

and unobserved characteristics, i.e., $\epsilon_d \perp X, \theta$.[18]

So far, our latent factor potential outcome model is 'robustly' identified by the factor structure θ and exclusions Z. Our third source of a more robust model identification is provided by using auxiliary information (cognitive test scores) from our individuals to measure θ. More precisely, we have a measurement system M that is independent of an individual's treatment status D, and which is adjoined to the latent factor θ (Carneiro *et al.* 2003). In other words, although θ represents latent ability, which is *unobserved* by us, we nonetheless expect it to systematically generate *observable* outcomes such as cognitive test scores. Our measurement system is made up of P observed variables, which are several ASVAB measures of cognitive ability. Each measurement p in $p = 1, \ldots, P$ is generated by the latent factor θ, while we control for observed covariates, X_p, which might bias test scores (such as family wealth).[19] Random variables ϵ_{M_p} denote idiosyncratic variation in tests scores not explained by covariates or the latent factor.[20]

$$M_p = \lambda_p X_p + \alpha_{M_p} \theta + \epsilon_{M_p} \quad p = 1, \ldots, P. \tag{12}$$

Attaching this measurement system to θ achieves three things. First, it provides meaning to the latent factor, i.e., it yields evidence (via tests on α_{M_p}) to what extent it really does capture unobserved cognitive ability. Second, it anchors θ. 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 *et al.* 2003).

With this statistical structure in hand we can identify our central quantities: the treatment effect of union membership.

[18] To be explicit, we employ the following 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, $1 > Pr(D = 1|X) > 0$.

[19] Table A.2 gives an overview of all variables used in choice, turnout, and test equations.

[20] They are assumed to be distributed mean zero with finite variance, $\sigma^2_{M_p}$, and independent of all covariates and the latent factor. More precisely, $\epsilon_M \perp X, Z, \theta$ and $\epsilon_{M_i} \perp \epsilon_{M_j}, i \neq j, i, j = 1, \ldots, P$. This is the typical conditional independence assumption made in measurement models (see e.g. Jackman 2008) – that conditional on the latent variable errors are independent.

5.3. Treatment effects

We focus on the two most common treatment parameters, the average treatment effect and the effect of the treatment on the treated (Heckman and Robb 1985; Heckman 1997). 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

$$\begin{aligned} ATE(X,\Gamma) &= Pr(Y_1 = 1|X,\Gamma) - Pr(Y_0 = 1|X,\Gamma) \\ &= \Phi\left(\frac{\beta_1' X}{\sqrt{1+\alpha_1'\alpha_1}}\right) - \Phi\left(\frac{\beta_0' X}{\sqrt{1+\alpha_0'\alpha_0}}\right), \end{aligned} \tag{13}$$

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

$$\begin{aligned} TT(X,\Gamma,D=1) &= Pr(Y_1 = 1|X,D=1,\Gamma) - Pr(Y_0 = 1|X,D=1,\Gamma) \\ &= \left(\frac{\Phi(\beta_D' Z)}{\sqrt{1+\alpha_D'\alpha_D}}\right)^{-1} \int \left[\Phi(\beta_1' X + \alpha_1\theta) - \Phi(\beta_0' X + \alpha_0\theta)\right] \\ &\quad \times \Phi(\beta_D' Z + \alpha_D\theta)\phi(\theta)d\theta \end{aligned} \tag{14}$$

Here, ϕ denotes the normal distribution PDF. Since below we are interested in describing population average treatment effects of union membership (unconditional of individual characteristics), we integrate over the (empirical) distribution of X.[21]

5.4. Priors and estimation

We follow Heckman *et al.* (2013) and estimate the model in a Bayesian framework. While the philosophical advantages of the approach are well documented (e.g., Gill 2008), Bayesian models for potential outcomes are rarely used.[22] One of the key advantages of a Bayesian approach is that we recover the full distribution of all parameters in the model. We are thus able to generate posterior distributions for all estimated treatment effects and

[21] In other words, $E(ATE) = \int ATE(x)dF_X(x)$ and mutatis mutandis for $E(TT)$.

[22] But note that estimating the system of equations using maximum likelihood (using Gauss-Hermite quadrature to integrate over the latent variable) yields comparable results. In fact, we use ML estimates as starting values for our Gibbs sampler.

communicate the full picture of their uncertainty.[23]

We estimate the 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. One issue that needs resolving is the common identification problem of latent factor models, which allows elements of $\alpha = (\alpha_D, \alpha_0, \alpha_1, \alpha_{M1}, \ldots, \alpha_{MP})'$ to 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. We identify the orientation of α by fixing $\alpha_D = 1$. 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 (prior (hyper-)parameters are marked by a tilde). 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 $\tilde{B}_j = \boldsymbol{I}_j \tilde{b}_j$. 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 as well as alternative values used for our robustness checks are given in appendix Table A.3.[24]

6. Results: parameter estimates

In this section we provide a detailed discussion of our estimation results. We discuss the treatment effect of union membership in Section 7.

[23] The "price" paid is that we need to specify priors for all parameters in the model. We discuss our prior choices below and also provide an appendix where we conduct sensitivity analyzes for our prior parameterizations.

[24] To summarize our prior robustness tests: An alternative IG prior with shape=1 and scale=0.005 yields estimates that are numerically close and substantively identical. Normal priors with variances twice as large also show that our results are not insensitive to this change (see also Table 7.

Table 2: Choice equation ($D = 1$)

	Mean	SD	95% HPD		Marg.[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] Marginal effect (on probability scale) of unit change in X,Z
[b] Fixed parameter

Union choice Table 2 shows estimates from our equation modeling selection into union membership. 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] As discussed above our latent factor θ is normalized to unity in the selection equation. It affects union membership substantially: a standard deviation change raises the probability of union membership by 33 percentage points (holding everything else constant). Conditional on the latent factor (i.e., holding cognitive ability constant), we find that both of our two central variables induce union membership in the expected way. The higher the differential between union and non-union wages, the higher the likelihood of union membership. Similarly, having a job in a more highly concentrated industry raised the probability of being a union member.

Turnout If there is systematic selection on unobservables (or, in our case, more accurately, hitherto unobservables) into union membership and turnout, we will find that θ significantly affects our potential outcomes. Table 3 shows the posterior summary from our two turnout equations. Panel (A) displays estimates for union members, panel (B) for non-members. We find that θ does indeed substantially affect turnout in both potential outcome states. The Bayesian posterior density intervals are far away from zero in both

[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 3: Turnout equations

	Mean	SD	95% HPD		Marg.[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] Marginal effect of unit change in X

cases, indicating the statistical "significance" of our result. The latent factor has a stronger influence on turnout choice of union members. A standard deviation change raises the 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 3 is as expected from previous research. In particular, higher socio-economic status (income, education) is associated with a higher propensity to vote. Individuals who previously experienced spells of unemployment are less likely to turn out.

Test scores Given the clear importance of θ in our model and in the results discussed above, we should investigate if it captures meaningful differences between individuals. If

Table 4: ASVAB test equations

	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 are held equal across test items

θ does represent (to some extent) cognitive abilities, we expect to find that it significantly shapes observed test scores of an individual. Table 4 shows estimate from our system of equations (12).

We clearly see that θ substantially influences achieved ability test scores. Higher values of θ are associated with higher coding speed, more correct 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 the all these relationship are highly statistically reliable. In other words, the availability of D-independent measurements allows us to give meaning to the (hitherto unobserved) latent factor in our model. This relationship holds while controlling 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 test (holding all else, including θ, equal). These distorting influences do exist and are controlled for in our measurement model.

Figure 2 plots the distribution of our latent factor for union members and non-members. We constructed the plot by drawing 500 samples from the posterior distribution of θ_i,

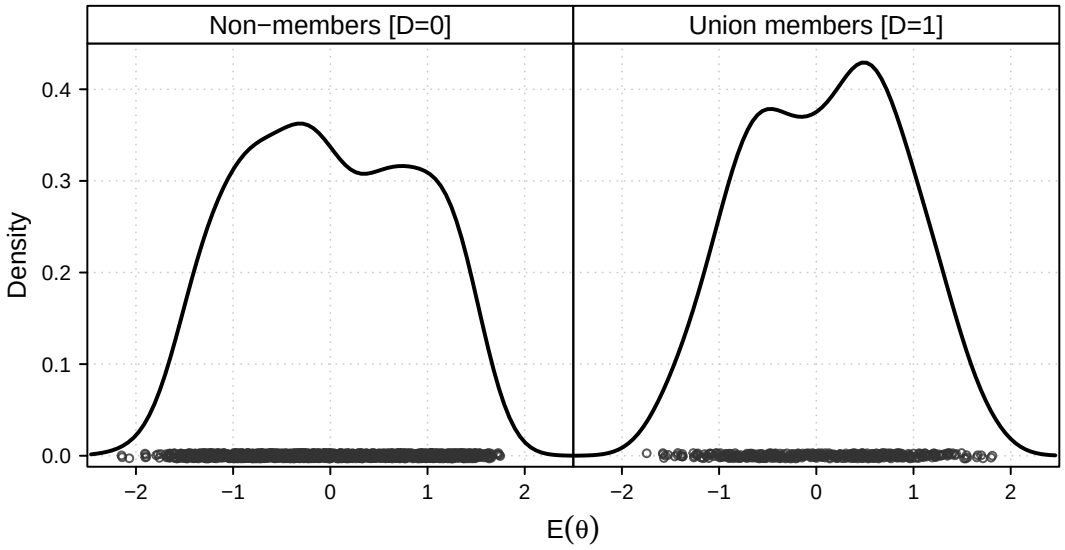

Figure 2: Distribution of latent factor values

calculate the posterior expectation, $E(\theta_i)$ for each individual, and then calculating a kernel density estimate. Figure 2 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 5, 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 shows that union members do indeed, on average, have higher ability than non-members. 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 less low θ values than non-members. The same finding obtains (somewhat less pronounced) at the top of the distribution.

7. Results: treatment effects

In contrast to the wealth of tables produces in the previous section, the summary of our treatment effects is straightforward. Following equations (13) and (14), we calculate the (population) average treatment effect of union membership on turnout, as well as the treatment effect among union members. Table 6 shows a summary of the posterior

Table 5: Summary statistics for distribution of $\boldsymbol{\theta}$ 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).

Table 6: Treatment effects

	Mean	SD	95% HPD	
Average treatment effect (ATE)	0.104	0.018	0.069	0.139
Treatment effect on treated (TT)	0.093	0.021	0.050	0.134

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

distribution of both quantities. The average treatment effect of union membership on turnout is estimated as 0.104 ± 0.018. In other words, union membership increases the likelihood of turnout by 10.4 percentage points. Because the Bayesian approach employed here yields the full distribution of uncertainty of this estimate, we know that 95% of the posterior "mass" 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, though the type of person who joins a union is quite different, on average, from one who does not. The treatment effect among the treated (where $D = 1$, i.e, union members) is 0.093 ± 0.021. The likelihood of turnout is raised by 9.3 percentage points, with model uncertainty ranging from 5 to 13 percentage points.

We conducted several robustness checks displayed in Table 7. Most noteworthy, we used a random subsample approach to capture the influence of unobserved confounders. 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). Furthermore, we include regional effects as well as an indicator variable for public sector employment. We find our central results confirmed (or

Table 7: Robustness checks of average treatment effect

Specification	Mean	SD	95% HPD	
Random subsamples	0.119	0.023	0.076	0.167
Region fixed effects	0.104	0.019	0.071	0.141
Public sector employment	0.108	0.019	0.071	0.144
Industry fixed effects	0.102	0.018	0.069	0.138
Alternative priors 1	0.105	0.018	0.069	0.140
Alternative priors 2	0.104	0.018	0.069	0.140

even strengthened). Even when including a set of industry fixed effects, so that we only use within-industry changes in concentration levels, we find our results largely confirmed.[26] Finally, To guard ourselves against unduly influence of prior choices, we re-estimate our models under different prior specifications (see Table A.3 for details), and find our results unchanged.

8. Conclusion

Do voters join unions or do unions cause higher turnout among its members? In this paper, we have used a unique survey data set and applied the appropriate statistical method to provide credible estimates of the causal effect of union membership on turnout. To deal with the well-known selection problem, our approach jointly models the decision to join a union and the choice to vote. In doing so, it exploits three distinct sources of causal identification in the presence of unobserved confounders. We find that, on average, the union effect on turnout is robust and substantively large (about 9-10 percentage points). Methodologically, we have drawn on the growing literature in statistics and econometrics that extends the potential outcomes framework to non-linear models that can deal with more complex data structures and allow researcher to exploit additional information, like cognitive tests, to impose less restrictive identification assumptions. This complements the more traditional approach (which we also follow) of trying to find plausible instrumental variables for an endogenous variable like union membership.

One limitation of this study is that it only considers one election (the 2006 mid-term) for the one cohort that makes up the NLSY. This reflects data constraints. While the panel

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

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 (or other aspects of politics). 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 (like the American National Election Study) frequently used to study voting. But there are similar surveys in other countries. So the approach can be applied to different data sets to assess the external validity of the findings.

With this caveat in mind, the results have important implications for group-based theories of political participation and the larger debate about the questions about the causes and consequences of declining turnout. The finding is consistent with mobilization models of turnout, which provide a coherent explanation for how group influence on individual members can generate apparently strategic behavior at the macro level. It also provides micro foundations for macro studies that have found a link between aggregate union strength and turnout. More speculatively, these findings provide further impetus to carefully study the possibility that union decline has an indirect effect, through electoral mobilization, on election results and policy outcomes.

A. Appendices

A.1. Union wage differential

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

Note Calculated from US Bureau of Labor Statistics, series LU. 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.

[a] Difference in median weekly earnings [in contemporary $].

[b] Not calculated since base size less then 50,000

A.2. Model equations

Table A.2: 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

A.3. Prior sensitivity analysis

We conduct a range of prior sensitivity analyses (see Gill (2008: 199f.) for an overview). Table A.3 list hyper-parameter values used in the model (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.

Table A.3: Prior parameters specifications

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

References

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

Abbring, J. H. and Heckman, J. J. (2007). Econometric evaluation of social programs, part III: Distributional treatment effects, dynamic treatment effects, dynamic discrete choice, and general equilibrium policy evaluation. In Heckman, J. J. and Leamer, E. E. (Eds.), *Handbook of Econometrics*, Elsevier, volume 6. pp. 5145–5303.

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

Aldrich, J. H. (1993). Rational choice and turnout. *American Journal of Political Science*, 246–278.

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

Anderson, T. W. and Rubin, H. (1956). Statistical inference in factor analysis. In Neyman, J. (Ed.), *Proceedings of the third Berkeley symposium on mathematical statistics and probability*. Berkeley: University of California Press, volume 5, pp. 111–150.

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**, 101–107.

Belman, D. (1988). Concentration, unionism, and labor earnings: A sample selection approach. *Review of Economics and Statistics*, 391–397.

Budd, J. W. and Na, I.-G. (2000). The union membership wage premium for employees covered by collective bargaining agreements. *Journal of Labor Economics*, **18**, 783–807.

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

Carneiro, P., Hansen, K. T. and Heckman, J. J. (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**, 361–422.

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

Coate, S. and Conlin, M. (2004). A Group Rule: Utilitarian Approach to Voter Turnout: Theory and Evidence. *The American Economic Review*, **94**, 1476–1504.

Cox, G. W. (1999). Electoral Rules and the Calculus of Mobilization. *Legislative Studies Quarterly*, **24**, 387–419.

Cox, G. W., Rosenbluth, F. M. and Thies, M. F. (1998). Mobilization, Social Networks, and Turnout: Evidence from Japan. *World Politics*, **50**, 447–474.

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

D'Art, D. and Turner, T. (2007). Trade Unions and Political Participation in the European Union: Still Providing a Democratic Dividend? *British Journal of Industrial Relations*, **45**, 103–126.

Denny, K. and Doyle, O. (2008). Political Interest, Cognitive Ability and Personality: Determinants

of Voter Turnout in Britain. *British Journal of Political Science*, **38**, 291–310.

Duncan, G. J. and Stafford, F. P. (1980). Do union members receive compensating wage differentials? *American Economic Review*, **70**, 355–371.

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

Enos, R. D., Fowler, A. and Vavreck, L. (2014). Increasing Inequality: The Effect of GOTV Mobilization on the Composition of the Electorate. *Journal of Politics*, **76**, 273–288.

Erikson, R. S. (1995). State Turnout and Presidential Voting. *American Politics Research*, **23**, 387–396.

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

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

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

Fowler, A. (2012). Regular Voters, Marginal Voters, and the Electoral Effects of Turnout. University of Chicago.

Fowler, A. (2013). Electoral and Policy Consequences of Voter Turnout: Evidence from Compulsory Voting in Australia. *Quarterly Journal of Political Science*, **8**, 159–182.

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? National Bureau of Economic Research, Working Paper 9992.

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

Gray, M. and Caul, M. (2000). Declining Voter Turnout in Advanced Industrial Democracies, 1950 to 1997. *Comparative Political Studies*, **33**, 1091–1122.

Hansford, T. G. and Gomez, B. T. (2010). Estimating the Electoral Effects of Voter Turnout. *American Political Science Review*, **104**, 268–288.

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

Heckman, J. (1997). Instrumental Variables: A Study of Implicit Behavioral Assumptions Used in Making Program Evaluations. *The Journal of Human Resources*, **32**, 441–462.

Heckman, J. J. (1981). Statistical models for discrete panel data. In Manski, C. and McFadden, D. (Eds.), *Structural Analysis of Discrete Data with Econometric Applications*, Cambridge: MIT Press. pp. 114–178.

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

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

Heckman, J. J. and Robb, R. (1985). Alternative methods for evaluating the impact of interventions: An overview. *Journal of Econometrics*, **30**, 239–267.

Heckman, J. J., Smith, J. and Clements, N. (1997). Making the most out of programme evaluations

and social experiments: Accounting for heterogeneity in programme impacts. *Review of Economic Studies*, **64**, 487–535.

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

Heckman, J. J. and Vytlacil, E. J. (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 Heckman, J. J. and Leamer, E. E. (Eds.), *Handbook of Econometrics*, Elsevier, volume 6, Part B of *Handbook of Econometrics*. pp. 4875–5143.

Hill, K. Q., Leighley, J. E. and Hinton-Andersson, A. (1995). Lower-Class Mobilization and Policy Linkage in the U.S. States. *American Journal of Political Science*, **39**, 75–86.

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

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

Jackman, S. (2008). Measurement. In Box-steffensmeier, J. M., Brady, H. E. and Collier, D. (Eds.), *Oxford Handbook of Political Methodology*, Oxford: Oxford University Press. pp. 119–151.

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

Johnson, G. E. (1975). Economic analysis of trade unionism. *American Economic Review*, **65**, 23–28.

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

Koop, G. and Poirier, D. J. (1997). Learning about the across-regime correlation in switching regression models. *Journal of Econometrics*, **78**, 217–227.

Krosnick, J. A. and Lupia, A. (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.

Lamare, R. J. (2010). Union Influence on Voter Turnout: Results from three Los Angeles Country Elections. *Industrial and Labor Relations Review*, **63**, 545–470.

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

Lee, D. S., Moretti, E. and Butler, M. J. (2004). Do Voters Affect or Elect Policies? Evidence from the U. S. House. *The Quarterly Journal of Economics*, **119**, 807–859.

Lee, L.-F. (1978). Unionism and wage rates: A simultaneous equations model with qualitative and limited dependent variables. *International Economic Review*, 415–433.

Leighley, J. (1996). Group Membership and the Mobilization of Political Participation. *The Journal of Politics*, **58**, 447–463.

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

Leighley, J. E. and Nagler, J. (2013). *Who Votes Now? Demographics, Issues, Inequality, and Turnout*

in the United States. Princeton and Oxford: Princeton University Press.

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

Lochner, L. J. and Monge-Naranjo, A. (2011). The nature of credit constraints and human capital. *American Economic Review*, **101**, 2487–2529.

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

Maddala, G. (1986). *Limited-dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.

Mahler, V. A. (2008). Electoral turnout and income redistribution by the state: A cross-national analysis of the developed democracies. *European Journal of Political Research*, **47**, 161–183.

Martinez, M. D. (2010). Why is American Turnout so Low, and Why Should We Care? In Leighley, J. E. (Ed.), *The Oxford Handbook of American Elections and Political Behavior*, Oxford: Oxford University Press.

McCarty, N., Poole, K. T. and Rosenthal, H. (2006). *Polarized America*. Cambridge, MA; London, England: MIT Press.

McFadden, D. L. (1984). Econometric analysis of qualitative response models. In Griliches, Z. and Intrilligator, M. (Eds.), *Handbook of econometrics*, Amsterdam: Elsevier, volume 2. pp. 1395–1457.

Morgan, S. L. and Winship, C. (2007). *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. New York: Cambridge University Press.

Morton, R. B. (1991). Groups in Rational Turnout Models. *American Journal of Political Science*, **35**, 758–776.

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

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

Pencavel, J. H. (1971). The Demand for Union Services: An Exercise. *Industrial and Labor Relations Review*, **24**, 180–190.

Poirier, D. J. (1998). Revising beliefs in nonidentified models. *Econometric Theory*, **14**, 483–509.

Poirier, D. J. and Tobias, J. L. (2003). On the predictive distributions of outcome gains in the presence of an unidentified parameter. *Journal of Business & Economic Statistics*, **21**, 258–268.

Powell, G. B. J. (1986). American Turnout in Comparative Perspective. *American Political Science Review*, **80**, 17–43.

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

Quandt, R. E. (1972). A new approach to estimating switching regressions. *Journal of the American Statistical Association*, **67**, 306–310.

Radcliff, B. (2001). Organized Labor and Electoral Participation in American National Elections. *Journal of Labor Research*, **22**, 405–414.

Radcliff, B. and Davis, P. (2000). Labor Organization and Electoral Participation in Industrial Democracies. *American Journal of Political Science*, **44**, 132–141.

Rosenfeld, J. (2010). Economic Determinants of Voting in an Era of Union Decline. *Social Science Quarterly*, **91**, 379–396.

Rosenstone, S. J. and Hansen, M. J. (1993). *Mobilization, Participation, and Democracy in America*. New York: Macmillan.

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

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

Schnabel, C. (2003). Determinants of trade union membership. In Addison, J. T. and Schnabel, C. (Eds.), *International Handbook of Trade Unions*, Cheltenham: Edward Elgar. pp. 13–43.

Shachar, R. and Nalebuff, B. (1999). Follow the Leader: Theory and Evidence on Political Participation. *The American Economic Review*, **89**, 525–547.

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

Stephens, J. D. and Wallerstein, M. (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*, 390–422.

Verba, S., Schlozman, K. L. and Brady, H. E. (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**, 69–89.

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

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

Western, B. and Rosenfeld, J. (2011). Unions, Norms, and the Rise in U.S. Wage Inequality. *American Sociological Review*, **76**, 513–537.

Zullo, R. (2004). Labor Council Outreach and Union Member Voter Turnout: A Microanalysis from the 2000 Election. *Industrial Relations*, **43**, 324–338.

Zullo, R. (2008). Union Membership and Political Inclusion. *Industrial and Labor Relations Review*, **62**, 22–28.