Voter Identification and Nonvoting in Wisconsin, Election 2016*

Kenneth R. Mayer and Michael G. DeCrescenzo
University of Wisconsin–Madison

Updated November 24, 2017

How much did Wisconsin's voter identification requirement matter in 2016? We constructed a survey of registered Wisconsin nonvoters in the counties surrounding Milwaukee and Madison to estimate the number of registered voters who were kept from voting in 2016 due to the voter ID requirement. We find that roughly 11% of nonvoters in these urban areas claimed that voter ID was at least a partial reason why they did not vote in the 2016 presidential election, amounting to approximately 17,000 citizens. Further, we find that 6% of nonvoters (approximately 10,000 registrants) either lack a qualifying ID or list voter ID as their primary reason for not voting in 2016. With a Bayesian analysis that combines our survey with administrative data from the Wisconsin voter file, we produce a range of estimates suggesting that voter turnout in Milwaukee and Dane Counties was reduced by 0.5 to 2 percentage points.

Note for APW: This paper is in its early stages. We would appreciate feedback especially on how credible and robust the findings are, as well as how our results fit into the debate over the various estimation methods—the CCES large-scale survey seems to be the main alternative method. Reactions to the Bayesian approach would be useful as well (having said that, it isn't yet utilized to its full potential!). The key question is how well we have extracted the most credible information that the data can provide. Thanks!

*[*] **Working paper. Please do not cite or circulate without author permission.** Financial support for data collection was provided by Office of the Dane County Clerk. The survey and respondent tracking were implemented with the assistance of the University of Wisconsin Survey Center. The authors thank Barry Burden and David Canon for helpful feedback.

How much did Wisconsin's new voter ID requirement affect voter turnout in the 2016 presidential election? In this paper, we analyze the results of an original survey fielded in the aftermath of the election to estimate the law's impact on turnout.

By surveying voters who did not vote in 2016 nonvoters in Wisconsin's two largest counties—Milwaukee and Dane—we estimate the extent to which Wisconsin's strict photo ID requirement impeded voting among registrants. We then outline and implement a Bayesian strategy for estimating the overall impact on turnout in these two counties. Posterior estimates suggest that voter ID reduced turnout in these two counties 0.5 percentage points (under conservative assumptions) to 2 percentage points (using less conservative assumptions). Because the survey was fielded in these two counties only, we cannot generalize these findings to the overall state of Wisconsin.

The paper proceeds as follows. First, we review existing strategies used to study voter ID laws. Next, we discuss our survey data and estimation approach. Having outlined the approach, we present a small set of posterior estimates from the procedure in its current form. Finally, we discuss potential criticisms about the approach and the robustness of the findings.

Approaches to Studying Voter ID Requirements

Estimating the effect that voter ID laws have on turnout is a difficult empirical problem. Both the individual decision to vote and aggregate turnout are affected by a large number of variables, of which the lack of ID is only one consideration. Moreover, voters may be confused about the law and *believe* that they lack a qualifying ID when their ID in fact does qualify.

Here, we classify the most common methods that have been used to study voter ID requirements, summarize their conclusions, and identify potential problems in the inferences drawn.

Provisional Ballots:

The most direct method for observing the consequences of voter ID examines the number of provisional ballots that voters cast.[1] Under the Help America Vote Act (HAVA), voters who arrive at polling place but are not on the voter rolls (or whose eligibility is challenged)

[1] See, for example, Government Accountability Office (2014).

must be given the chance to cast a "provisional ballot," with standards for counting (or "curing") defined by state law. In Wisconsin, a provisional voter who does not show a qualifying ID on election day has until 4 p.m. on the Friday after an election to show ID at the municipal clerk's office.

Wisconsin typically has issued very small numbers of provisional ballots, primary because of election day registration is allowed under state law. A voter who was not on the list of registered voters could immediately register with the required documentation. In 2012, Wisconsin issues 132 provisional ballots, all of which were related to registration problems (a lack of driver's license or no proof of residence).[2] In the 2016 general election, the first general in which the state's voter IDlaw was in effect, 821 provisional ballots were cast by voters who did not show a qualifying form of ID (0.023% of all votes), 585 of which were rejected (0.011%).[3] This increase, while small in absolute terms, represents a relative 522% increase in the number of provisional ballots and a 540% increase in the provisional balloting rate. Other analyses of ID-related provisional ballots have found similar issuance rates (Pitts 2013).

Provisional ballots may be overinclusive or underinclusive as an estimate of the effect of voter ID laws. Voters might simply forget to bring their IDs with them when they present (a probably situation with the 173 provisional voters who cured their ballots), or possess an ID and not take the time to cure their ballots. More seriously, estimating the effect of ID using provisional ballots will miss any individual who does not attempt to vote or register to vote because they lack ID.[4] The weight of evidence suggests that underestimation is the more serious problem, as the estimates of ID effects using provisional ballots are considerably smaller than estimates produced by other methods (Stewart 2013, 27–38).

ID Possession Rates

A second method—which has been common in litigation over ID laws—estimates the number and percentage of registered voters who possess either a driver's license or a state-issued

[2] Wisconsin State Elections Commission, GAB-190 Summary Statistics 2012 (http://elections.wi.gov/sites/default/files/publication/65/20121106_gab190_summarystats_pdf_14962.pdf). Ohio, by comparison, issued 208,084 provisional ballots and rejected 34,299 in 2012.

[3] Wisconsin State Elections Commission, GAB-190 Summary Statistics 2016 (http://elections.wi.gov/sites/default/files/publication/2016_general_election_summary_statistics_pdf_15354.pdf).

[4] Under Wisconsin law, registering requires less documentation than voting. Individuals who do not have an ID can use the last four digits of their social security numbers to register (allowed under HAVA).

Department of Transportation (DOT) ID. These are by far the most common forms of ID, and they are universally accepted as a qualifying IDin states with voter ID laws.

These estimates can be derived from surveys that ask individuals about the forms of ID that they possess, from record linkage methods that match voter registration files with statewide files of driver's licenses and DOT IDs, and occasionally from other statewide databases such as concealed-carry weapons permits. Because federal privacy laws protect driver's license information, the linkage method "has been carried out almost exclusively by expert witnesses in the context of litigation" (Stewart 2013, 24).

Survey Data: According to the 2012 Survey of the Performance of the American Electorate (SPAE), 93% of registered voters possessed a driver's license, and 41% had a U.S. passport (Stewart 2013, 36). The rate of valid identification possession rate falls to about 84% when individuals are asked follow-up questions about expiration dates and current addresses. Possession rates vary considerably by race, with African American and Hispanic respondents considerably less likely to hold a valid license or ID (63% and 73%, respectively).

The most extensive surveys of ID possession rates have been conducted by Matt Barreto and Gabriel Sanchez, either as part of an ongoing research project or in expert testimony in voter ID litigation. A 2007 survey of registered voters in Indiana estimated that 83.9% of respondents had a current license of state-issued voter ID (Barreto, Nuño, and Sanchez 2009, 113).

In an expert report submitted on behalf of plaintiffs who were challenging Pennsylvania's voter ID law, Barreto and Sanchez (2012b) conducted a survey that found an estimated 87.2% of registered Pennsylvania voters possessed a valid form of ID (unexpired with conforming name), with statistically significant differences in the possession rates between women and men (92.8% and 88.5%, respectively), between Latinos and non-Hispanic whites (81.7% and 86%), and across income levels (78% of respondents with annual household income below $20,000, vs. 81.8% among respondents with household income above $80,000). They did not find a significant difference in possession rates between whites and African Americans.

Similarly, Barreto and Sanchez (2012a) estimated that among registered voters in Milwaukee County, WI, 9.5% reported in 2011 that they did not possess a qualifying form of ID that would allow them to vote, with higher rates of non-possession among Latinos (14.9 percent) and African Americans (13.2%). A similar study in Texas by (Barreto and

Sanchez 2014) found that 4.7% of eligible White voters in Texas did not possess a qualifying form of ID that would allow them to vote, again with higher non-possession rates in minority populations (11.4% among Latinos and 8.4% among African Americans 2014).

Record Linkage: Parties to litigation can obtain access to both the full voter registration lists with fields that are otherwise confidential (such as date of birth, social security numbers, an driver's license numbers), as well as DOT files and other databases. This allows for record linkage methods to determine whether a registered voter possesses a license or state-issued ID and thus estimate the aggregate rate of ID possession. In states that do not record race during the voter registration process, the registration data can be supplemented by estimates of registrant race based on name, birth date, and geographic data.

Record linkage is a probabilistic method, as it relies on connecting an individual in one database to the same individual in another, often using databases that are not designed to facilitate such a process.[5] Nonconforming data field, minor differences in name entries, and entry errors can result in both false positives (matching an individual in one database to a different individual in another database) and false negatives (where an individual exists in both databases but is not linked).[6] When data quality is high, combinations of available fields can yield very high probabilities of accurate matches (Ansolabehere and Hersh 2017).

Depending on the state and the qualifying forms of identification, linkage methods generally show that between 5% and 11% of registered voters do not possess a driver's license or state DOT photo ID (Government Accountability Office 2014, 22–25). Stewart (2014) found that 6.1% of registered voters in North Carolina lacked any form of ID necessary to vote (including passports, military IDs, and Tribal IDs), with African Americans more than twice as likely as white registrants to lack ID.

Estimates of ID non-possession rates in Wisconsin range from 4.5% (Hood 2015, 27) to 8.4% (Mayer 2015, 19).

[5] To give one example, the Department of Transportation driver file in Wisconsin includes driver's license holders who are deceased, those with both a driver's license and state photo-ID (one can have one or the other but not both), and uses a different format for name and address fields than the statewide registration database.

[6] A sample conducted by New York City linking the voter registration file to the state motor vehicle database found that nearly 20% of unmatched records were false negatives resulting from typographical errors in the registration data (Levitt, Weiser, and Muñoz 2006, 4).

Overall Turnout and Individual Voting Behavior

A third method analyzes voter turnout and attempts to isolate the specific effect of voter ID on voting levels using aggregate reports of voter turnout, large-scale surveys, or voting histories of individual registrants (typically after being linked to driver's license or ID files).

A GAO investigation compared states with photo ID laws to demographically similar states, estimating that photo ID requirements depressed turnout in Kansas by 1.9 to 2.2 percentage points and turnout in Tennessee by 2.2 to 3.2 percentage points (Government Accountability Office 2014, 48). A Priorities USA analysis asserted that Wisconsin's voter ID law reduced 2016 turnout by 200,000 votes, nearly ten times President Trump's margin of victory (Priorities USA 2017). These estimates are almost certainly far too high, as they attribute nearly all turnout declines to voter ID laws. Turnout is likely to have been affected by other factors not accounted for in either analysis.

Other studies examine validated voting data from voter files. Using voter files that were linked to driver's license databases in Georgia, Hood and Bullock found that Georgia's strict photo ID law lowered aggregate turnout by 0.4 percentage points (Hood and Bullock 2012). Using data from the 2006-2014 Cooperative Congressional Election Study, Hajnal, Lajevardi, and Nielson (2017) find that voter ID laws depress minority turnout among Latinos by 7.1 percentage points, and among African American voters in primary elections by 4.6 percentage points (368). Grimmer et al. (2017) were critical of this finding, arguing that it was likely the result of respondent misreporting and measurement error, and that the data analysis is flawed (concluding that the analysis, when corrected, is inconclusive about the effects). One problem is that as a large scale survey, the "CCES is not designed to be representative of small populations like those lacking photo IDs" (10), a flaw that as be exacerbated when the survey is applied to low frequency events (Ansolabehere, Luks, and Schaffner 2015).

Empirical approach

In this section, we describe our strategy for estimating the turnout effect of Wisconsin's strict photo ID requirement in the 2016 general election. First, we describe the survey data we use to estimate quantities of interest. Because the survey asks registrants about their experience with the ID law in a few different ways, it is possible to construct slightly different indicators for whether a nonvoter was deterred from voting because of the ID law, which we discuss below. We then describe a theoretical decomposition of the set of nonvoting registrants to identify which subset was kept from voting due to voter ID, and we present a framework to estimate this quantity from our survey data.

Survey of Nonvoters in Wisconsin

The goal of this analysis is to estimate the number of nonvoters who were deterred or prevented from voting due to Wisconsin's photo ID law. To construct these estimates, a survey was mailed to 2,400 registered Wisconsinites in Milwaukee and Dane County who *did not vote* in the November 2016 election. These counties contain the two largest metro areas in the state (Milwaukee and Madison) and have the largest low-income and minority populations, which existing research suggest are most likely to be affected by voter ID requirements. Because the sampling frame contains only these two counties, estimates in this study cannot be extrapolated to the state of Wisconsin as a whole.

Nonvoting registrants were identified using voter histories in the Wisconsin registered voter file (also referred to as the "WisVote" file) with the data file generated on February 10, 2017.[7] The file contained 247 individuals who registered to vote after the presidential election on November 8, 2016. These individuals were removed from the voter file before sampling.

Individuals of lower socioeconomic status (SES) are more likely to be affected by the voter ID requirement but often have lower response rates to surveys. For this reason, the sampling design included an oversample of Census tracts with lower aggregate measures of SES. The sample was stratified as follows:

- Dane County: 650 surveys,
- Milwaukee County, high SES: 750 surveys,
- Milwaukee County, low SES: 1,000 surveys,

with all analyses conducted using sample weights generated by the University of Wisconsin Survey Center. Demographic characteristics of the high- and low-SES tracts are available in the Appendix.

The survey asked voters their reasons for not voting, how closely they followed the election campaign, whether they had been contacted by campaign officials during the campaign, their knowledge of qualifying forms of voter identification, and a handful of demographic questions. Because the study was financially supported by the government of Dane

[7] Surveys were collected with assurances of confidentiality and were de-identified before analysis. The project was approved by the Educational and Social/Behavioral Sciences Institutional Review Board (IRB) on February 9, 2017 (protocol number 2017-0056).

County, we included no questions about political party affiliations or vote intentions in the 2016 election season. The full questionnaire can be viewed online.[8]

A total of 293 surveys were returned, with 75 respondents from Dane County, 213 from Milwaukee County, and 5 whose home counties could not be identified. Because weights were constructed using geographic location, the 5 respondents whose counties could not be determined were assigned no weights and thus excluded from the analysis, resulting in 288 surveys returned.

Identifying the "Affected" Group

The survey asked respondents several questions to assess their experiences with voter ID in the 2016 election. First, respondents were asked why they did not vote, with voter ID included among several other reasons.[9] Voters could initially select several *partial* reasons for not voting (which we refer to as "nominal" reasons in the analysis below), and then they were asked to select their *main* reason for not voting. For reasons related to voter ID, respondents could indicate if they believed they lacked a qualifying ID ("You did not have the right photo ID and know you would not be able to vote") or if they attempted to vote but were told that they did not have a qualifying ID ("You tried to vote, but were told at the polling place that you did not have the necessary photo ID"). Later in the survey, respondents were asked about the forms of ID they possess, which we used to determine whether respondents lacked a qualifying voter ID.[10]

Table 1 displays marginal responses to nonvoters' nominal reasons for not voting. Although there is some concern that nonvoters may cite voter ID as a reason for not voting

[8] https://elections.wisc.edu/news/Voter-ID-Study/Voter-ID-Study-Instrument.pdf

[9] The survey included the following potential reasons for not voting: being ill or disabled, being out of town, not having enough time, not being interested in voting, having a transportation problem that prevented them from getting to the polls, not liking the choice of candidates or issues, being unable to obtain an absentee ballot, lacking a qualifying ID, attempting to vote but being told at the polls that their ID was not qualifying, long lines at the polls, encountering a problem with early voting, and believing that one's vote would not matter. These options were derived from a similar question item used in the Census Bureau's Current Population Survey November Voting and Registration Supplement. Other academic surveys (such as MIT's "Survey of the Performance of American Elections") use similar items as well.

[10] "Currently, do you have each of the following forms of identification?" Respondents could separately indicate if they possessed several forms of ID, only some of which would satisfy the voter ID requirement. The survey does not indicate to the respondent which forms of ID satisfy the voter ID requirement. The qualifying IDs include a Wisconsin driver's license, Wisconsin Department of Transportation ID, a voting-only ID, a military ID, a Native American tribal ID, a certificate of recent naturalization, and a U.S. passport. The non-qualifying IDs include a credit card, a permit to carry a concealed weapon, a state or federal government ID, and a Social Security card.

TABLE 1: Nominal Reasons for Not Voting

Reason	Yes (%)	No (%)	NA (%)
Unhappy with choice of candidates or issues	50.8	33.5	15.7
Not interested	27.5	49.6	22.9
Not enough time	26.7	51.2	22.2
Vote would not have mattered	26.2	51.2	22.6
Away from home	20.1	62.0	17.9
Ill or disabled	18.4	64.6	16.9
Problem with early voting	12.5	61.5	26.0
Couldn't get absentee ballot	8.1	67.4	24.6
Transportation problems	7.7	69.3	23.0
Did not have adequate photo ID	6.5	69.4	24.0
Lines too long	3.0	71.9	25.1
Told at polling place that ID inadequate	2.9	72.7	24.3

in order to deflect their own responsibility for not voting, there is little evidence of this. Just 6.5% of report that they did not have adequate ID, and 2.9% say that they were turned away at the polls because of voter ID. By contrast, half of all respondents (50.8%) list that they were unhappy with their choice of candidates and issues, and more than a quarter of all respondents said they were not interested and that their votes would note have mattered (27.5% and (26.2%) respectively. The responses shown in Table 1 suggest that voter ID-related nonvoting is rare but not non-existent.

Table 2 shows respondent's main reasons for not voting. When individuals are asked to choose their main reason for not voting, fewer respondents cite ID. Whereas more than 6% cited lacking an ID as a nominal reason, this number drops to 1.7% of individuals who cite lacking ID as their primary reason for not voting. The percentage of individuals who were turned away at the polling place due to ID also falls from about 3% to 1.4%.

Finally, Table 3 shows how many respondents possess each form of ID included in the survey. Because the survey asked respondents about forms of ID that do and do not qualify as valid voter IDs, we include responses to both qualifying and non-qualifying IDs. For subsequent analysis, we code respondents as lacking a qualifying ID if they said they did not possess or did not know if they possess all qualifying forms of ID.[11] With this coding,

[11] We code "don't know" responses as lacking ID because if a respondent is unaware whether they possess a form of ID, presumably they would not be able to use it to vote.

TABLE 2: Main Reasons for Not Voting

Main Reason	Percent
Unhappy with choice of candidates or issues	33.0
Ill or disabled	13.6
Away from home	13.5
Not enough time	9.3
Not interested	8.8
Vote would not have mattered	6.6
NA (None given)	4.9
Problem with early voting	2.9
Transportation problems	2.1
Did not have adequate photo ID	1.7
Told at polling place that ID inadequate	1.4
Couldn't get absentee ballot	1.3
Lines too long	0.9

we find that 3% of the sample lacks a qualifying voter ID.

Using these items about citizens' experiences with voter ID, we construct two ways to define the population of affected citizens. We refer to registrants as *"deterred"* from voting if they lack qualifying ID or mention ID as a reason for not voting. Voter ID could be a nominal reason or the primary reason for not voting. Using a stricter definition, we refer to registrants as *"prevented"* from voting if they lack qualifying ID or list voter ID as their primary reason for not voting. We focus primarily on "deterrence" from voting because we believe it is more consistent with the literature on election laws and voting costs. Electoral reforms can lower an individual's propensity to vote even if they do not make it impossible to cast a ballot. For this reason, we regard nominal reasons for nonvoting as an important manifestation of voter ID's impact on political participation. Furthermore, even if citizens possess a qualifying ID, confusion about the law and which forms of ID are allowed can lead individuals to believe (wrongly) that they cannot vote. A study of nonvoting registrants in Texas presents supporting evidence that confusion about qualifying forms of ID can lower individual propensities to vote (Hobby et al. 2015).

Figure 1 presents our sample-based estimates of the percentage of nonvoters in Dane and Milwaukee Counties deterred and prevented from voting by the ID law. The figure shows point estimates and 95% confidence intervals calculated using the Clopper-Pearson

TABLE 3: ID Possession by Respondents

ID Form	Possess (%)	Lack (%)	DK (%)	NA (%)	Qualifying
WI Voter ID Card	2.4	75.2	1.7	20.7	Yes
WI Driver's License	79.7	14.8	0.8	4.6	Yes
WI DOT ID Card	21.7	59.2	3.1	15.9	Yes
U.S. Passport	42.3	43.2	0.4	14.1	Yes
Naturalization Certificate	3.2	75.7	1.7	19.4	Yes
Native Am. Tribe ID	1.2	78.4	0.4	19.9	Yes
Military ID	5.7	74.3	0.9	19.1	Yes
State/Federal Employee ID	5.0	74.6	0.4	19.9	No
Social Security Card	89.0	3.9	0.7	6.4	No
Non-WI Driver's License	5.6	74.7	0.4	19.3	No
Credit Card	73.8	18.2	0.4	7.6	No
Concealed Carry Permit	6.8	74.4	0.4	18.3	No

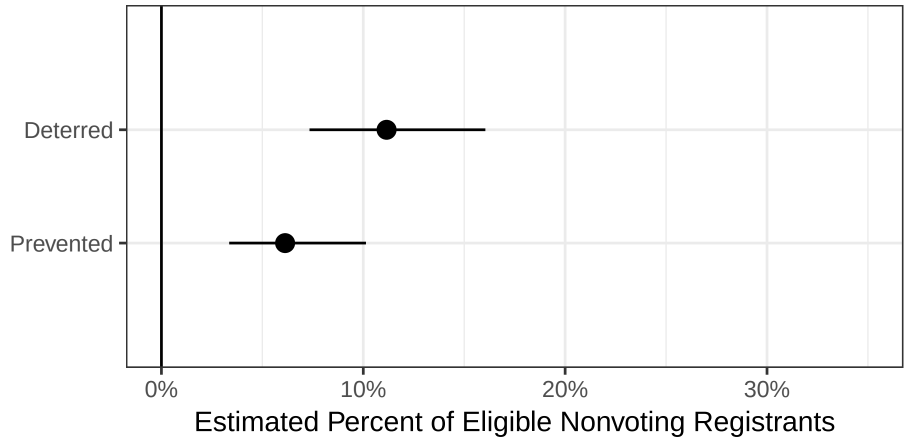
method.[12] We estimate that 11.2% of nonvoting registrants in Dane and Milwaukee counties were "deterred" in some way from voting by the voter ID law, either because they lacked ID, believed they lacked ID, or were told at the polls that their ID did not qualify as valid. $The 95\%$ interval is between 7.3% and 16.0%. The stricter definition of the effect consists of voters who were effectively "prevented" from voting because they lacked an ID or cited ID as the main reason they did not vote. Under this definition, 6.1% of nonvoters were prevented from voting (95% interval: 3.4% to 10.1%).

Although our sample is considerably smaller than the samples in national surveys, because our estimates are near 0%, they have lower variance than estimates near 50% for an equally-sized sample. As a result, our margins of error are largely comparable to other national surveys (about 3 to 4%).[13]

[12]Although it is common to estimate uncertainty bounds for proportions by approximating the binomial distribution with a normal distribution, the assumptions underlying such a method are less reliable in smaller samples and for success probabilities near 0 or 1. Clopper-Pearson intervals are "exact" in the sense that they are derived directly from the quantiles of the binomial distribution (though they may be "conservative" in the sense that 95% intervals may obtain more than 95% coverage). For k successes in n trials with success probability π, the Clopper-Pearson interval contains all values of π that would produce k successes within the inner 95% of its corresponding distribution. The interval bounds themselves are computed with the related beta distribution: $B\left(\frac{\alpha}{2};k,n-k+1\right) < \pi < B\left(1-\frac{\alpha}{2};k+1,n-k\right)$, where $B(q;y,z)$ represents the q^{th} quantile from a beta distribution with parameters y and z.

[13]Clopper-Pearson intervals allow asymmetric uncertainty regions around point estimates, so statements

FIGURE 1: Sample estimates of the fraction nonvoters deterred and prevented from voting due to voter ID (with Clopper-Pearson confidence intervals)



Importantly, most of the individuals we identify as deterred or prevented from voting report that they possess some form of voter ID. Although 11% of the sample was deterred and 6% was prevented from voting due to ID, just 3% of the sample reports lacking an ID. Although strange at first, these findings are very similar to other surveys of nonvoters arguing that voter ID laws create confusion about qualifying forms of ID and therefore impede voting even among individuals who possess qualifying IDs (Hobby et al. 2015). Although respondents may possess driver's licenses, those licenses may have expired, or those individuals may have recently moved residences or changed their names, which could lead to uncertainty about their ability to vote. We devote more attention to this in the Discussion section below.

Generating a Population Estimate: Theoretical Decomposition

This study is interested in the effect of Wisconsin's voter ID requirement on voter turnout. Appealing to a potential outcomes framework, if we let D be an indicator to represent the presence ($D = 1$) or absence ($D = 0$) of the strict voter ID requirement, then the turnout impact of the ID law (δ) is the difference in turnout across values of D:

$$\delta(D) = Turnout(D = 0) - Turnout(D = 1). \tag{1}$$

about margins of error are not exact.

We do not observe $D = 0$ in Wisconsin for election 2016, so it should be treated as a counterfactual scenario.

Our study narrows our focus to turnout among registered voters. With this denominator, $Turnout(D = 1)$ is straightforward to operationalize as the number of recorded votes divided by the number of eligible registrants:

$$Turnout(D = 1) = \frac{Recorded\,Votes}{Eligible\,Registrants}. \tag{2}$$

To obtain $Turnout(D = 0)$, however, we need to add the number of votes that were suppressed by the ID law to the number of recorded votes:

$$Turnout(D = 0) = \frac{Recorded\,Votes + Suppressed\,Votes}{Eligible\,Registrants}. \tag{3}$$

The number of suppressed votes is unknown, but we present a method to estimate it using survey data and the WisVote registered voter file. We begin by showing, theoretically, how the number of suppressed votes can be decomposed from the set of nonvoting registrants in the WisVote file. This decomposition depends on a small set of unknown parameters. We then discuss how we model and estimate these unknown parameters using the available data.

The WisVote file is a record of voter registrations with one row per registration. Our sampling method began by restricting the file to the counties of Milwaukee and Dane, including only those individuals who do not have a record of casting valid votes for the 2016 election. Let this value, the number of nonvoting records in Milwaukee and Dane Counties, be equal to N.

Routine voter list maintenance had not taken place at the time our sample was taken, so we treat the WisVote file as if only a fraction of it contains citizens who remain eligible to vote in Milwaukee and Dane counties. Citizens in the file would be ineligible if they had died, become disabled, moved counties, or updated their registration due to a change of name. If we let the fraction of eligible records in the WisVote file be ζ, then the number of eligible registrants is $N\zeta$.

Among the set of eligible but nonvoting registrants, let ϕ represent the fraction of the set that is deterred from voting due to the voter ID law either by lacking a qualifying ID, believing they lack an ID, or being turned away at the polls.[14] The number of eligible

[14] Although we present two definitions of the affected population, "deterred" and "prevented" from voting,

nonvoters deterred from voting due to ID is thus represented as $N\zeta\phi$. Although many of the individuals who experienced higher voting costs due to voter ID may eventually vote (by overcoming information costs or by obtaining valid ID), the size of that set does not bear directly on $N\zeta\phi$, since $N\zeta\phi$ contains only nonvoters.

The last unknown deals confronts an issue of counterfactuals. Although $N\zeta\phi$ individuals were deterred from voting in 2016 due to voter ID, it is not appropriate to say that all of those individuals would have voted in 2016 in the absence of the ID requirement; some fraction of those voters would not have voted in either case. To account for this, we let τ be the fraction of deterred nonvoters who would have voted in the absence of the law, with the strong expectation that $\tau < 1$. If these parameters represent the true quantities of interest without error, the number of suppressed votes in 2016 due to voter ID is equivalent to $N\zeta\phi\tau$. Equations 4 through 7 review this theoretical decomposition of the WisVote file.

$$\text{Nonvoters in Voter File} = N \tag{4}$$

$$\text{Eligible Nonvoters} = N\zeta \tag{5}$$

$$\text{Eligible Nonvoters Affected by ID Law} = N\zeta\phi \tag{6}$$

$$\text{Suppressed Votes by Voter ID} = N\zeta\phi\tau \tag{7}$$

Modeling the Effect of Voter ID on Turnout

To estimate the number of suppressed votes as a decomposition of the WisVote file, we must estimate the parameters ζ, ϕ, and τ. Because each parameter is a proportion lying between 0 and 1, it would be intuitive to model the data as binomial random variables to estimate the underlying parameters. Due to certain features of the data, however, our choice of distributions require some additional explanation. We describe our distributional assumptions by presenting the "ideal" distributions, highlighting flaws in the ideal distributions, and presenting alternative distributions.

First, we estimate ϕ and τ using survey data. As described above, suppose that some fraction ϕ of registered nonvoters is deterred from voting due to voter ID. Using the survey data, it is possible to identify a set of respondents whose responses indicate ID-related deterrence. Ideally, the number of respondents in this set could be described by a binomial

methodological discussion simply uses "deterred" for the sake of simplicity.

distribution parameterized by ϕ for a given sample size n:

$$k \sim \text{Binomial}(n, \phi), \tag{8}$$

where k is the in-sample number of deterred respondents. Similarly, we could identify a set of individuals who would have voted absent the strict photo ID requirement. For now, we use whether deterred individuals voted in 2012 as an approximation for the propensity of deterred voters who would have voted without the ID law.[15] One way to model the number of respondents who would have turned out (v) would be to fix k (the set of deterred registrants) and estimate τ using another binomial distribution:

$$v \sim \text{Binomial}(k, \tau). \tag{9}$$

However, because k is itself a random variable (Equation 8), we should not take knowledge of k for granted and treat it as a fixed sample size. Instead, we should model ϕ and τ jointly, acknowledging that our uncertainty about both values depends on the same sample of survey data. We therefore model the intersections of ID-related deterrence and 2012 turnout as a categorically distributed variable with four outcome categories. The probability that a respondent i lies in each category is determined by a vector of four associated probabilities π, which together sum to 1.

$$x_i \sim \text{Categorical}(\pi) \tag{10}$$

Table 4 describes how the four outcome categories are indexed. Importantly, π_1 captures the probability of being both deterred and voting in 2012, while π_2 captures the probability of deterrence and not voting in 2012. The categorical assumption allows us to recover estimates of ϕ and τ while accounting for their dependence in a finite set of survey data. The total deterrence probability (ϕ) is simply $\pi_1 + \pi_2$, and the probability of voting in 2012 conditional on deterrence (τ) is equal to $\frac{\pi_1}{\pi_1+\pi_2}$. Furthermore, the categorical setup implies that $\phi\tau = \pi_1$, allowing a simpler expression of the number of suppressed votes as $N\zeta\pi_1$ instead of $N\zeta\phi\tau$. Because of these equivalences, we estimate π directly rather than ϕ and τ separately (though we can always calculate ϕ and τ using the information in π).

[15] Because presidential campaign activity in Wisconsin was greater in 2012 than in 2016, turnout in 2012 is likely an overestimate of what turnout in 2016 would have been absent the voter ID requirement. We modify this assumption later in the analysis.

TABLE 4: Categorical outcomes for deterrence and past turnout

Variable	Category	Probability
$x_i = 1$	Deterred by ID law and voted in 2012	π_1
$x_i = 2$	Deterred by ID law and did not vote in 2012	π_2
$x_i = 3$	Not deterred by ID law and voted in 2012	π_3
$x_i = 4$	Not deterred by ID law and did not vote in 2012	π_4

The second complicating factor is the presence of survey weights. Although we cannot incorporate weights in a "fully Bayesian" way because we do not have a probability model of the weights, we can incorporate information about weights into the analysis by weighting the loglikelihood of each x_i when computing the posterior distribution.[16] In effect, x's contribution to the overall loglikelihood of the data can be written as follows:

$$\ell(\pi \mid x) = \ell(\pi \mid x_1)\, w_1 + \ell(\pi \mid x_2)\, w_2 \ldots \ell(\pi \mid x_n)\, w_n$$

where each $\ell(\pi \mid x_i)$ and w_i represent the loglikelihood and sample weight, respectively, for each x_i. Intuitively this means that when the data are used to update the posterior distribution, oversampled data provide less information per observation (Lumley 2004).

Finally, to estimate the eligibility rate (ζ) among the registered individuals in the WisVote file, we tracked a sample of 200 survey non-respondents using Lexis/Nexis. Tracking was used to determine the number of non-deceased registrants who still live in their county of registration. The results of the tracking allow us to estimate the rate of eligibility in the WisVote file,

$$z \sim \text{Binomial}(200, \zeta),$$

where z is the number of still-eligible registrants in the sample of 200.

Although we will experiment with different priors as the analysis progresses, we use

[16] Bayesian posterior distributions are proportional to the prior distribution times the likelihood of the data: $p(\theta \mid x) \propto p(x \mid \theta)p(\theta)$. Because weights are introduced during data collection, this method incorporates weighting into the likelihood of the data, $p(x \mid \theta)$.

conjugate priors for π and ζ:

$$\pi \sim \text{Dirichlet}(\gamma), \tag{11}$$

$$\zeta \sim \text{Beta}(2,2). \tag{12}$$

We currently give π a flat prior but intend to introduce more information as the analysis progresses.[17] We give the WisVote eligibility rate a $\text{Beta}(2,2)$ prior to downweight the plausibility of extraordinarily high and extraordinarily low eligibility rates.[18]

We estimate these parameters using Markov chain Monte Carlo as implemented with `Stan` and the `rstan` package for R. We run 50,000 iterations across four chains using a thinning interval of 10, burning the first half of each chain. This gives us 10,000 draws for each parameter (2,500 per chain). With a set of parameter samples, it is straightforward to calculate the posterior estimate of suppressed votes by computing $N\zeta\pi_1$ for each sample draw.

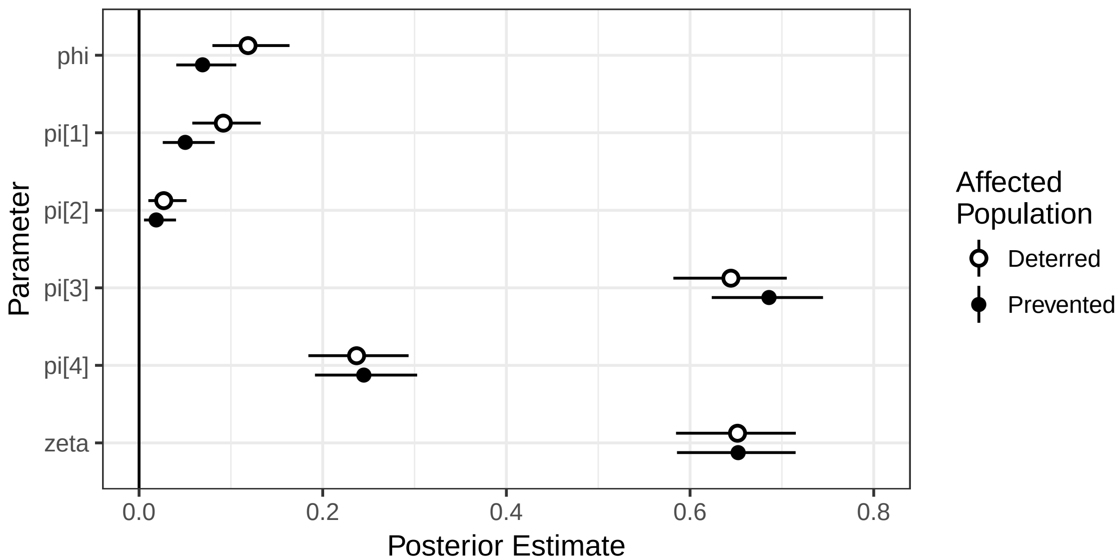
Posterior Estimates

The MCMC returns a sample of parameter values from the joint posterior distribution of all parameters. Figure 2 shows marginal distributions for each estimated parameter from two models. The first model (shown with hollow points) is estimated when the affected population is "deterred" from voting due to voter ID, and the second model codes the affected population as "prevented" from voting. The models therefore return slightly differing values of π, the categorical variable that codes whether a nonvoter was affected by ID and voted in the 2012 general election (as outlined in Table 4). The probability that a nonvoter was both affected and voted in 2012 is represented by π_1, and the total probability that a voter was affected by ID is represented by ϕ, which is the sum of π_1 and π_2. The ζ parameter shows that an estimated 65% of registrants in the WisVote file remain eligible to vote in their same county of registration. This eligibility estimate is not affected by the decision to code

[17] A key benefit of the Bayesian approach is that we can incorporate our prior knowledge that not all categories within x should be equal probability—it is unreasonable that very low and very high rates of deterrence are equally plausible *a priori*. Future versions of the paper will explore this.

[18] The Wisconsin Elections Commission's maintenance process, performed after this sample was taken, removed 51.5% of nonvoters as no longer eligible. We allow our estimate to differ from the the WEC's estimate because their method for determining continued ineligibility—sending mailers to registrants' address—is likely to overstate the rate of ineligibility.

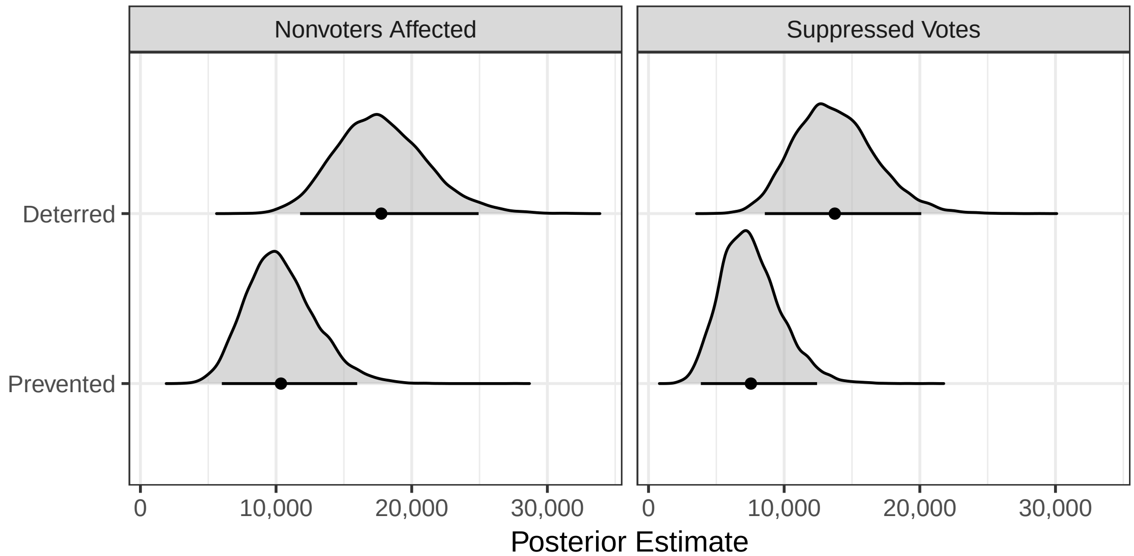
FIGURE 2: Marginal posterior estimates for model parameters (mean and 95% credible interval)



Because the model is currently estimated with flat priors on all parameters (which will change in the future), estimates of deterrence and prevention do not vary greater from the estimates presented in Figure 1. The Bayesian model does provide us an advantage when we calculate other quantities of interest, however, but the posterior samples allow uncertainty in each parameter to proliferate through the calculation. Figure 3 plots posterior estimates for two such quantities. In the left panel, we show the estimated number (rather than the rate) of nonvoters in Milwaukee and Dane Counties who were deterred or prevented from voting due to voter ID ($N\zeta\phi$), and the estimated number of suppressed votes in 2016 due to voter ID ($N\zeta\pi_i$). Each panel shows separate sets of estimates for each definition of the affected group (deterred or prevented). Points and error bars indicate posterior means and 95% credible intervals for each estimate (using the quantile method), and the density curves indicate the entire distribution of posterior samples (10,000 draws for each parameter).

The model estimates that a mean of 17,755 nonvoters were deterred from voting due to voter ID (95% interval from 11,773 to 24,930), resulting in 13,731 suppressed votes (8,573 to 20,104). If we limit the affected group only to those prevented from voting due to ID, we estimate a mean of 10,361 nonvoters prevented from voting (6,003 to 15,980) resulting in 7,548 suppressed votes (3,857 to 12,425). If we consider the full distribution of estimates under both definitions, we approximate that at least 6,000 nonvoters and as many as 25,000

FIGURE 3: Posterior sample distributions, means, and credible intervals for the number of affected nonvoters (left) and number of suppressed votes (right)



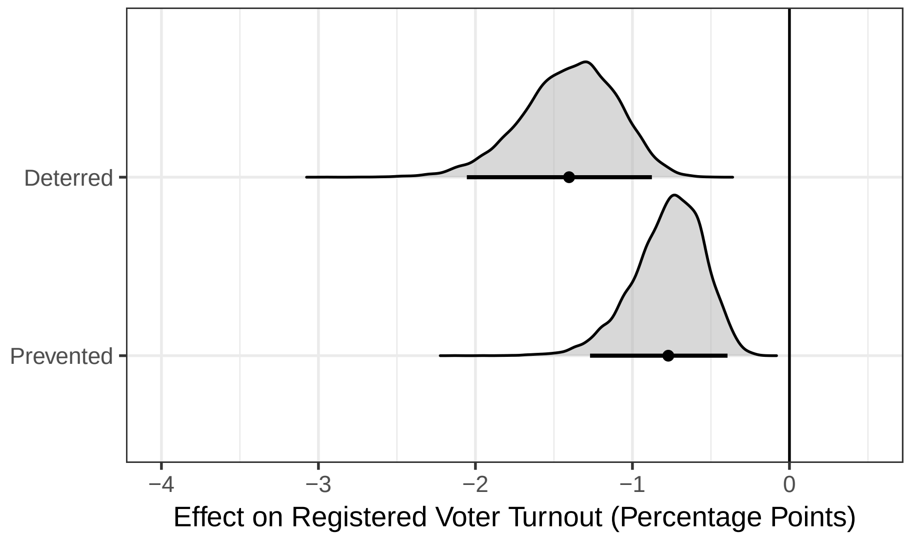
nonvoters had their voting costs raised by voter ID, resulting in 4,000 to 20,000 additional votes that would have been cast without the law.

By adding the number of suppressed votes to the number of validated votes in the WisVote file, we can estimate how much the voter ID law reduced voter turnout in the 2016 election. Using the total number of registrants and votes cast in Milwaukee and Dane County, the turnout rate among registered voters in 2016 was 76.5%. Adding the number of suppressed votes to the number of votes cast in 2016, our mean posterior estimates of turnout without voter ID are 77.9% using the "deterred" definition and 77.3% using the "prevented" definition. These amount to 1.4 and 0.77 percentage-point reductions in registered voter turnout, on average. Figure 4 plots the entire posterior distribution of percentage point reductions in turnout with posterior means and 95% credible intervals.

ROBUSTNESS OF THE FINDINGS

There are a number of potential concerns we wish to address with the data and findings.

FIGURE 4: Percentage point reduction in turnout due to voter ID using posterior samples



Counterfactual turnout

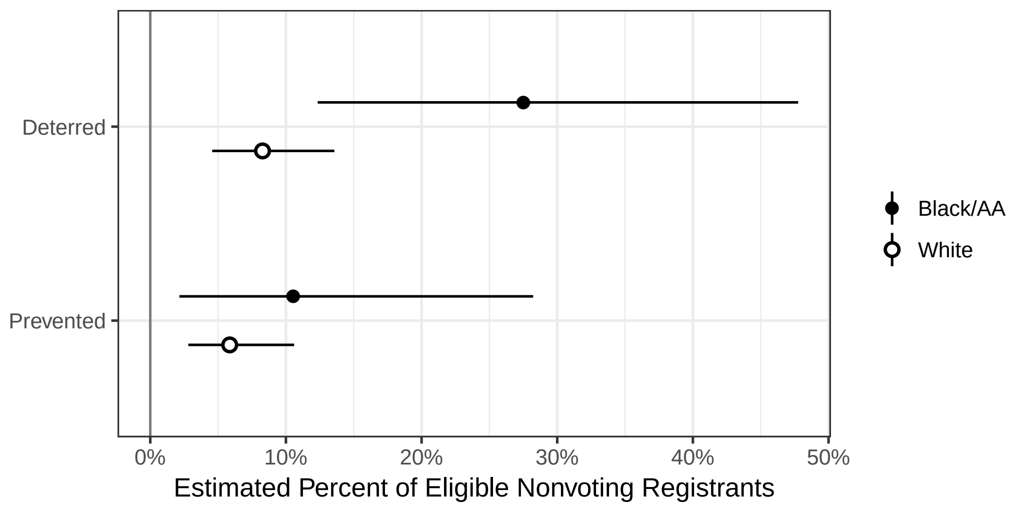
Although we use the sample data to build straightforward estimates of the share of nonvoters affected by voter ID, it is less straightforward to estimate the latent turnout propensity of the affected group. We use 2012 turnout as a rough proxy for the group's turnout rate, but it is likely that 2012 turnout would overestimate a counterfactual 2016 turnout due to differences in local campaign intensity in both elections. We would be interested in suggestions for how best to estimate this turnout rate. We have considered simply simulating the effect size for different turnout rates to demonstrate the sensitivity to this quantity.

Response Error

There is some concern that with a self-reported survey measure of ID-related deterrence, our estimates of the fraction of nonvoters deterred and prevented from voting are largely driven by response error or misreporting. Because voting is a socially desirable action, nonvoters may deflect responsibility for their decision not to vote by blaming external factors such as the voter ID requirement. Although this is a possibility, we have some evidence to suggest that it is unlikely. Most importantly, we observe subgroup variation in survey responses that is consistent with scholarly expectations about voter ID and inconsistent with a hypothesis of survey misreporting.

First, we find that racial variation in those deterred and prevented from voting are consistent with expectations about voter ID's effects on turnout. Figure 5 shows that African Americans are more likely than Whites to report that they were deterred or prevented from voting due to ID. The differences do not appear at conventional levels of significance, but we also have a considerably smaller sample of African Americans who are downweighted due to oversampling (estimates in Figure 5 reflect survey weights). Confidence intervals using unweighted data are narrower.

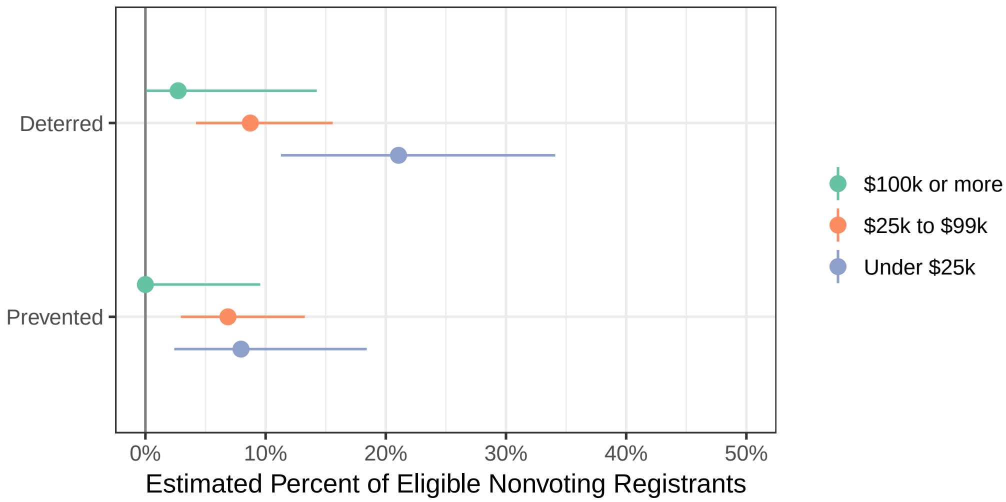
FIGURE 5: Racial differences in ID-related nonvoting



According to recent studies of survey misreporting (Ansolabehere and Hersh 2012), individuals with higher levels of political engagement and knowledge are more likely to misreport socially desirable behaviors. If our findings were driven largely by misreporting, we should expect that indicators and correlates of political engagement should be positively correlated with ID-related nonvoting. By contrast, if responses were largely accurate, we would expect negative correlations between political engagement and ID-related nonvoting, since individuals with lower SES and less knowledge would be more likely to be negatively affected by the ID law.

Our evidence on this front is broadly inconsistent with a misreporting hypothesis. First, we find that individuals with higher incomes are *less*, not more, likely to report that voter ID was related to their decision not to vote. Individuals in the lowest income category, $25,000 or less, were most likely to report that they were deterred or prevented from voting by voter ID. Additionally, we constructed a scale of knowledge about the voter ID law by asking individuals which forms of ID qualified as valid voter IDs. Values on the scale indicate the

FIGURE 6: Income differences in ID-related nonvoting



fraction of IDs each respondent classified correctly.[19] When we use this scale to predict ID-related nonvoting, we find that individuals who are less knowledgeable about the ID law are more likely to be negatively affected by it. Both of these findings run counter to the misreporting hypothesis, which predicts that we should observe higher-SES and more knowledgeable respondents misreporting their experiences with ID.

We do notice that some individuals in the "deterred" and "prevented" categories did vote in the April 2016 presidential primary, in which the voter ID was also enforced. Although there are some valid explanations for finding this pattern, we can make a conservative assumption and drop these individuals from the analysis altogether. Although this does reduce the estimate of ID-related nonvoting, the reduction is not drastic (See Figure 8).

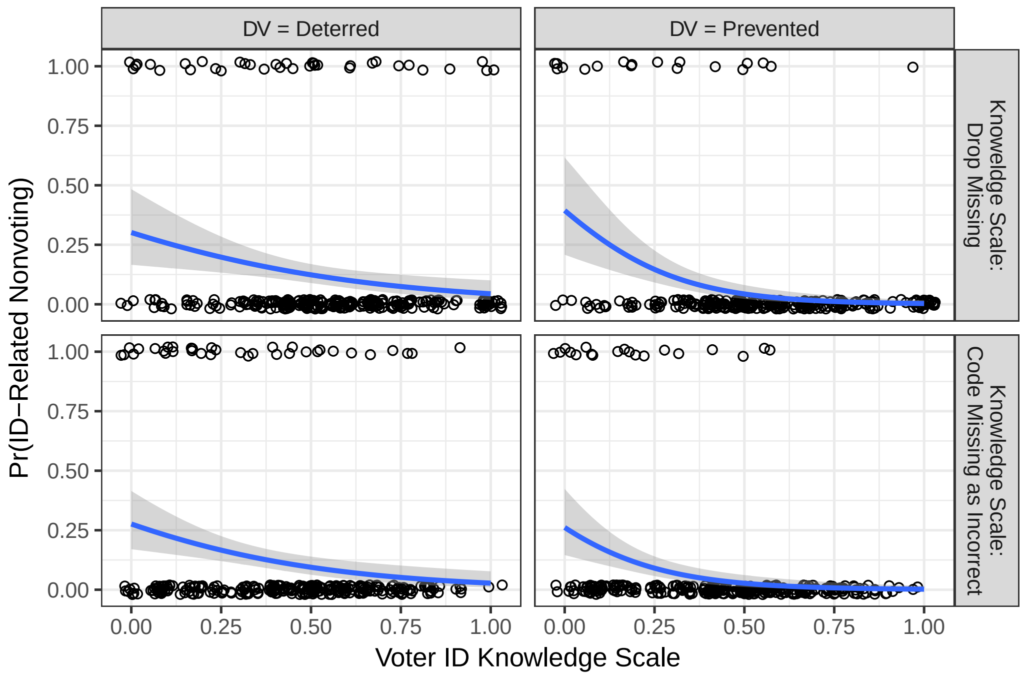
Selecting on the dependent variable

Some critics have alleged that because we cannot make reliable inferences by only studying nonvoters. Because the study does not measure the degree of ID-related costs among individuals who *did* vote in 2016, we cannot be sure about our inferences.

We disagree with these allegations. Supposing that we did ask voters if they felt that their experience voting was made more difficult because of the voter ID law, it is not clear

[19]We code "don't know" responses as incorrect. Because some respondents did not elect to classify all forms of ID, we code two versions of the scale—one that counts missing responses as incorrect, and one that simply omits missing responses from the calculation of the average. Results are not substantively affected by this coding decision.

FIGURE 7: Relationship between ID-related nonvoting and knowledge of ID law (data points jittered)



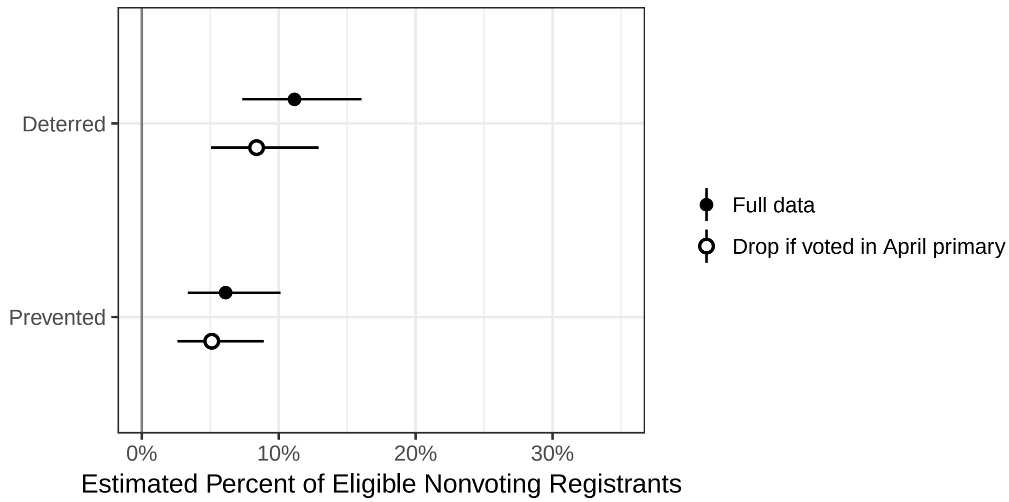
whatsoever how that information would be used to update our estimates. If the fraction of voters who experienced ID-related difficulty voting is greater or less than the equivalent fraction among nonvoters, neither finding would bear on the size of our estimates for nonvoters.

What we *would* like to do, however, is field a placebo study in a state without an ID law as strict as Wisconsin's. If a state without a strict ID law (such as Minnesota) produces a similar fraction of nonvoters who report ID-related difficulty, then this would cast doubt on the validity of our findings for Wisconsin.

A related concern is that, because we only survey nonvoters, any non-zero estimate of ID-related nonvoting will manifest as a negative effect on turnout. This, we admit, is true. However, given the scholarly knowledge of political participation and voter ID laws, it is unreasonable to expect a zero effect from voter ID. It is virtually certain that *some* individuals will be deterred or prevented from voting due to voter ID; the question we aim to answer is *how many* individuals are deterred or prevented.[20]

[20]If there is a positive "counter-mobilization" effect, this is a separate mechanism from the main effect of ID laws. If a study found a net positive effect of voter ID on turnout, we would argue that the study fails to

FIGURE 8: ID-related nonvoting after dropping all affected voters who participated in April 2016 primary



DISCUSSION

Although researchers have used several approaches to study voter identification requirements, most studies approach the subject by assuming that the vulnerable population consists of individuals lacking a qualifying ID. Early research into voter ID identified that racial minority and low-SES populations were less likely than White and higher-income groups to possess qualifying IDs. More recent studies that use sophisticated record linkage methods are also designed to identify a vulnerable population that consists of individuals who are registered to vote but who do not appear in DOT or passport databases. Our work differs from these earlier studies by broadening the pool of individuals who are potentially vulnerable to depressed turnout to include those who actually possess ID (see also Hobby et al. 2015). We contend, and find consistent evidence, that voter ID laws impose informational barriers to voting regardless of ID possession.

For example, although most registered Wisconsin voters possess a Wisconsin driver's license, these licenses expire every eight years. This means that in each four-year presidential election cycle, roughly half of all Wisconsin voters see their voter IDs expire. Voters may also move residences between elections, so the address on their driver's licenses may differ from the address at which they are registered to vote. Many voters also experience name changes predominantly due to marriage and divorce. Although Wisconsin's ID statute

directly quantify the number of voters kept from voting because of the law, which is the focus of our study.

provides some legal allowances for these individuals to vote with their current IDs, these allowances are highly detailed and not widely understood. While our study does not ask detailed follow-up questions about individual IDs, other studies find that initial estimates of ID ownership overestimate the prevalence of invalid IDs. Stewart (2013, 40), for example, finds that the rate of driver's license possession falls from 91% to 80% after accounting for expiration dates (1.6% of licenses), name changes (1.3%), and address changes (9.7%). A similar dropoff occurs for passport ownership rates after accounting for expiration and name changes (41% to 35%).

We believe that these sources of confusion have implications for the study of voter ID that cannot be ignored. Although record-linkage methods provide highly accurate estimates of the share of registrants who lack a qualifying IDs (e.g. Ansolabehere and Hersh 2017), the fact remains that ID laws raise voting costs on a much wider set of respondents than those without ID. Studies that restrict their focus to individuals lacking ID will fail to observe these effects. Despite heightened skepticism among researchers about the validity of survey responses (especially regarding political participation Ansolabehere and Hersh 2012), it is difficult to image how researchers can accurately quantify these sources of confusion without consulting voters and nonvoters directly. We therefore believe that surveys remain an important instrument for understanding the impact of voter ID requirements on political participation.

REFERENCES

Ansolabehere, Stephen, and Eitan Hersh. 2012. "Validation: What big data reveal about survey misreporting and the real electorate." *Political Analysis* 20 (4): 437–459.

———. 2017. "ADGN: An Algorithm for Record Linkage Using Address, Date of Birth, Gender and Name." Forthcoming, *Statistics and Public Policy.*

Ansolabehere, Stephen, Samantha Luks, and Brian F Schaffner. 2015. "The perils of cherry picking low frequency events in large sample surveys." *Electoral Studies* 40:409–410.

Barreto, Matt A, Stephen A Nuño, and Gabriel R Sanchez. 2009. "The disproportionate impact of voter-ID requirements on the electorate—new evidence from Indiana." *PS: Political Science & Politics* 42 (01): 111–116.

Barreto, Matt A., and Gabriel R. Sanchez. 2012a. *Rates of Possession of Accepted Photo Identification, Among Different Subgroups in the Eligible Voter Population, Milwaukee County, Wisconsin.* Expert report submitted on behalf of plaintiffs, *Frank v. Walker*, 11-cv-01128 (LA). April 23.

———. 2012b. *Rates of Possession of Valid Photo Identification, and Public Knowledge of the Voter ID Law in Pennsylvania.* Expert report submitted on behalf of plaintiffs. *Applewhite et al. v. Commonwealth of Pennsylvania, et al.*, No. 33 MD 2012. July 16.

———. 2014. *Accepted Photo Identification And Different Subgroups in the Eligible Voter Population, State of Texas, 2014.* Expert report submitted on behalf of plaintiffs, *Veasy v. Perry*, 13-cv-00193. June 27.

Government Accountability Office. 2014. *Issues Related to Voter Identification Laws.* GAO-14-634. September, revised February 2015.

Grimmer, Justin, Eitan Hersh, Marc Meredith, Jonathan Mummolo, and Clayton Nall. 2017. "Comment on 'Voter Identification Laws and the Suppression of Minority Votes'." Manuscript (March).

Hajnal, Zoltan, Nazita Lajevardi, and Lindsay Nielson. 2017. "Voter Identification laws and the suppression of minority votes." *The Journal of Politics* 79 (2): 363–379.

Hobby, Bill, Mark P. Jones, Jim Granato, and Renée Cross. 2015. "The Texas voter ID law and the 2014 election: A study of Texas's 23rd Congressional District." *Whitepaper: University of Houston Hobby Center for Public Policy and Rice University Baker Institute for Public Policy.*

Hood, M.V. III. 2015. *Expert Report of M.V. Hood.* Expert report for defendants, *One Wisconsin Institute v. Nichol* 3:15-CV-324. January 11.

Hood, MV, and Charles S Bullock. 2012. "Much ado about nothing? An empirical assessment of the Georgia voter identification statute." *State Politics & Policy Quarterly* 12 (4): 394–414.

Levitt, Justin, Wendy R. Weiser, and Ana Muñoz. 2006. *Making the list: Database matching and verification processes for voter registration.* Brennan Center for Justice. March.

Lumley, Thomas. 2004. "Analysis of complex survey samples." *Journal of Statistical Software* 9 (1): 1–19.

Mayer, Kenneth R. 2015. *Expert Report of Kenneth R. Mayer.* Expert report for plaintiffs, *One Wisconsin Institute v. Nichol* 3:15-CV-324. December 10.

Pitts, Michael J. 2013. "Photo ID, Provisional Balloting, and Indiana's 2012 Primary Election." *U. Rich. L. Rev.* 47:939.

Priorities USA. 2017. *Voter Suppression Analysis.* May 3.

Stewart, Charles III. 2013. "Voter ID: Who Has Them; Who Shows Them." *Okla. L. Rev.* 66:21.

———. 2014. *Declaration of Charles Stewart III, Ph.D. United States v. North Carolina* 1:13-CV-861. February 12.