

Supporting Information:
Estimating the Effect of Voter ID on Nonvoters in
Wisconsin in the 2016 Presidential Election

Kenneth R. Mayer
Principal Investigator
Professor of Political Science
University of Wisconsin–Madison

Michael G. DeCrescenzo
Ph.D. Candidate, Political Science
University of Wisconsin–Madison

September 25, 2017

More information about the study can be found at
<https://elections.wisc.edu/news/voter-id-study.html>

OVERVIEW

The presidential election of November, 2016, was the first general presidential election held under the state of Wisconsin's new voter identification requirement.

The goal of this analysis is to estimate the number of voting-eligible registered voters in select areas of Wisconsin who did not vote due to the voter ID requirement. This effect of the ID requirement takes multiple forms: citizens may be deterred from voting because they do not possess a qualifying ID, or they may not realize that they possess a qualifying form of ID because they are unsure which forms of ID are qualifying.

SURVEY

To construct these estimates, a survey was mailed to 2,400 nonvoting registrants in Milwaukee County and Dane County, Wisconsin. These counties contain the two largest metro areas in the state (Milwaukee and Madison) and have the largest low-income and minority populations, which research has shown are most likely to be affected by voter ID requirements. For this reason, the estimates cannot be extrapolated to the state of Wisconsin as a whole.

Nonvoters were identified using voter histories in the Wisconsin registered voter file (WisVote) with the data file generated on February 10, 2017.¹ The 247 individuals who registered to vote after the presidential election on November 8, 2016, were removed from the file.

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 survey oversampled registrants residing in Census tracts with lower aggregate SES measures metrics. The survey sample was divided into strata as follows:

- Dane County, 650 surveys
- Milwaukee County (High SES): 750 surveys
- Milwaukee County (Low SES): 1,000 surveys

Demographic characteristics of the high-SES and low-SES tracts are available in the Appendix of this report.

Because the study was funded by the Office of the Dane County Clerk, there were no questions asked about political party affiliations or vote choices.

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. All analysis below is conducted using

1. 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).

sampling weights to account for the stratified sample design. Sampling weights were constructed by the UW Survey Center before any analysis had taken place.²

ESTIMATION

We use a statistical approach to estimate the number of Wisconsinites affected by the law. Specifically, we use the survey to estimate the proportion of nonvoting registrants who did not vote due to the voter ID law. There are a few ways to define what it means for a citizen to be “affected” by the law. Those definitions are discussed in the following section.

A statistical approach presumes that the quantity of interest, the true proportion of ID-related nonvoting registrants, is not directly observable but influences the observable data. And because the data come from a random survey, there is uncertainty due to random sampling in the observable data. A standard statistical approach acknowledges this uncertainty by presenting a point-estimate of the quantity of interest as well as a 95% confidence interval. The range of the confidence interval represents the degree of uncertainty in the data—wider intervals indicate more uncertainty. The formal definition of the 95% confidence interval is an interval that, given the sample size and observed variation in the data, would contain the true parameter in 95 percent of repeated samples from the same data-generating process.

To implement the statistical analysis, we analyze whether voters are affected by the ID law as a binary variable: voters are either affected or not. Further, we analyze the data as a binomial process. For binary outcome data (1s and 0s, “successes” and “failures”, or “affected” and “not affected”), a binomial process describes the expected number of successes for a fixed sample size n and fixed success probability π . More specifically, the binomial distribution describes the probability of finding k many successes given the true success probability π and sample size n .³ The proportion of successes in the sample $\left(\frac{k}{n}\right)$ is an estimate of the true success probability π , but the two will not be exactly equal, so the binomial distribution provides a framework to make inferences about the true value of π from the observed data.

We construct confidence intervals for our estimates of π using the Clopper-Pearson method (Clopper and Pearson 1934). Although it is common to estimate the uncertainty in a binomial process by approximating the binomial distribution using a Gaussian or “normal” distribution, the assumptions underlying these methods 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 binomial distribution rather than from an approximation provided by another distribution. However, they can be “conservative” in the sense that the 95% intervals are sometimes wide enough to obtain more than 95% coverage. In keeping with dominant statistical practices, we opt for conservative uncertainty estimates to avoid overstating the degree of precision in the data.

2. Because weights are determined by geographic location, the 5 respondents whose counties could not be determined were excluded from the analysis.

3. The binomial distribution’s probability mass function is $p(X = k | n, \pi) = \binom{n}{k} \pi^k (1 - \pi)^{n-k}$, where $\binom{n}{k}$ represents a binomial coefficient.

The Clopper-Pearson interval can be represented in terms of the quantiles of the binomial distribution:

$$\left\{ \pi \mid P[\text{Bin}(n; \pi) \leq k] > \frac{\alpha}{2} \right\} \cap \left\{ \pi \mid P[\text{Bin}(n; \pi) \geq k] > \frac{\alpha}{2} \right\}, \quad (1)$$

where $0 \leq k \leq n$ is the number of successes, $\text{Bin}(n; \pi)$ is a binomial random variable with n trials and success probability π , and α represents the significance level (for 95% intervals, α is 0.05). Equation 1 states that the values of π within the confidence interval should only include values where the possibility of k successes given π falls only within the “inner 95%” of the distribution. The interval bounds themselves can be computed with a formula that relies on the related beta distribution:⁴

$$B\left(\frac{\alpha}{2}; k, n - k + 1\right) < \pi < B\left(1 - \frac{\alpha}{2}; k + 1, n - k\right), \quad (2)$$

where $B(q; y, z)$ represents the q^{th} quantile from a beta distribution with parameters y and z .⁵

WHO IS “AFFECTED” BY THE LAW? TWO DEFINITIONS

The survey asked respondents several questions to assess their experiences with voter ID in the 2016 election. Respondents were asked about the forms of ID they possess, which we used to determine whether respondents lacked a qualifying voter ID.⁶ Additionally, respondents were asked why they did not vote, with voter ID included among several other reasons.⁷ Voters could initially select several *partial* reasons for not voting (which we refer to as “nominal” reasons in the analysis

4. Where the binomial distribution describes the distribution of successes k given the values π and n , the beta distribution is often used to describe the distribution of an unknown probability (π). In effect, the Clopper-Pearson interval uses the information from the binomial process (n and k) to perform inference about the unknown value π , which is distributed beta.

5. As mentioned above, the survey data include sampling weights to account for oversampling small populations of citizens. Because analyzing survey data using weights can result in non-integer values, we simply round these non-integer values to the nearest integer in order for statistical software to compute Clopper-Pearson confidence intervals (namely, the `binom.test()` function in the statistical package R). The consequences of this correction are essentially imperceptible and should bias estimates neither upward nor downward overall. For comparisons across demographic categories, we sidestep the issue of rounding by performing direct calculations the Clopper-Pearson interval rather than use the pre-written `binom.test()` function.

6. “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.

7. 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.

below), and then they were asked to select their *main* reason for not voting. Voters 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”). The full questionnaire is available with this release.

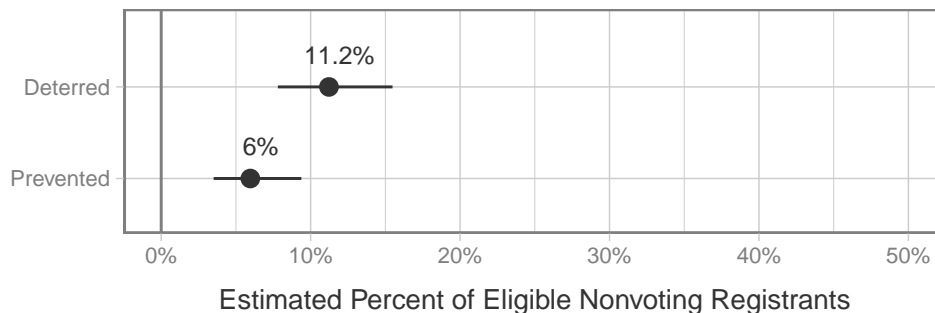
Using these several questions about citizens’ experiences with voter ID, we construct two ways to define the population of affected citizens.

- 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.

These two definitions of affected registrants are the focus of the analysis in this document. We focus primarily on “deterrence” from voting, since it compasses a fuller range of possible effects. Even if citizens possess qualifying ID, confusion about the law and which forms of ID are allowed can lead them to believe that they cannot vote when in fact they can. A study of nonvoting registrants in Texas finds a similar form of misunderstanding and confusion about the forms of ID that qualify under the law (Hobby et al. 2015). Since many of these individuals presumably would have voted in the absence of the heightened ID requirement, we include these individuals among the group of affected registrants. We outline our method for estimating the total *number* of affected registrants in the following section.

Figure 1 presents our 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 method. 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 7.8% and 15.5%. 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% of nonvoters were prevented from voting (95% interval: 3.5% to 9.4%).

Figure 1: Estimated percent of eligible registered nonvoters in Dane and Milwaukee counties deterred by ID law



CONVERTING PERCENTAGES INTO NUMBERS OF VOTERS

Another quantity of interest is the *number* of eligible registrants who were deterred or prevented from voting. Generating this figure requires multiplying the sample estimates from Figure 1 by the number of eligible registrants in the voter file. This section describes how we compute the number of eligible registrants in the voter file and presents the results.

All voter files naturally contain some ineligible records (Ansolabehere and Hersh 2010). A registrant may be ineligible under an existing record because of death, disability, moving out of state, or an updated registration based on a change of name. Because we are estimating the number of affected voters in Dane and Milwaukee counties, we also consider registrants who move to another address in Wisconsin outside of Dane County or Milwaukee County as “ineligible” to vote in those counties, though they may be eligible to vote at their new address. Although ineligible records are removed over time through the voter list maintenance process, that had not taken place at the time the survey was mailed.⁸

To estimate how many nonvoters were still eligible, The UW Survey Center tracked a random sample of 200 uncompleted surveys—nonresponses as well as mailings returned as “undeliverable”—using Lexis/Nexis to determine whether an individual fell into one of the categories of ineligible voters.⁹ This method found that an estimated 34.7% of nonrespondents in our sample were ineligible. This figure is lower than the percentage of ineligible nonvoters removed through the State Elections Commission list maintenance process (51.5%), a difference attributable to the fact that Wisconsin removes all nonrespondents and undeliverable list maintenance mailings from the list of eligible voters. We did not remove a registrant if Lexis tracking showed that a nonvoting individual was still alive, residing in the same county, and not otherwise ineligible.¹⁰

8. The survey was mailed on March 27, 2017. The Wisconsin Elections Commission completed the 2016 voter list maintenance process on August 1, 2017 (<http://elections.wi.gov/node/5174>). The process consists of a mailing sent to every registrant who had not voted since the 2012 presidential election to ask if the individual wished to remain an active registrant. All nonresponses and nondeliverable mailings are removed from the list of registrants.

9. We are in the process of extending the tracking methodology to all nonrespondents and nondeliverable address.

10. Wisconsin Elections Commission, *Voter Registration Four-Year Record Maintenance*, August 1, 2017. County-

The voter file data showed that there were 229,625 nonvoting registrants in Milwaukee and Dane Counties during the 2016 election. After subtracting the estimated fraction of registrants who were ineligible (34.7%), we estimate that there were 150,010 eligible registrants in Milwaukee and Dane Counties on November 8, 2016. We multiply our estimates in Figure 1 by this number of eligible registrants to project the number of registrants deterred or prevented from voting. We present these results in Table 1. The top two rows of the table contain the estimated percentage of respondents deterred or prevented from voting by the ID law for a given definition (column two) alongside the number of eligible respondents deterred or prevented from voting given each effect size (column four).¹¹ The bottom four rows present the results for the survey questions used to construct the “deterred” and “prevented” categories.

Table 1: Estimated nonvoters deterred or prevented from voting by voter ID law

	Percent	Interval	N Affected	N Interval
Deterred (Lack ID or Mention ID Law)	11.2	(7.8; 15.5)	16,801	(11,701; 23,252)
Prevented (Lack ID or Mention ID as Main Reason)	6.0	(3.5; 9.4)	9,001	(5,250; 14,101)
Mention ID as Nominal or Main Reason	8.4	(5.5; 12.3)	12,601	(8,251; 18,451)
Mention ID as Nominal Reason	7.4	(4.6; 11)	11,101	(6,900; 16,501)
Mention ID as Main Reason	3.2	(1.5; 5.9)	4,800	(2,250; 8,851)
Lack Qualifying ID	3.2	(1.5; 5.9)	4,800	(2,250; 8,851)

The Wisconsin voter file indicates that 748,777 votes were cast in Dane and Milwaukee Counties in 2016. If all of the estimated registrants deterred from voting had turned out in 2016, voter ID could have reduced turnout by 2.2 percentage points. If all of the estimated registrants prevented from voting had turned out, the law could have reduced turnout by 1.2 percentage points.

Turnout among those deterred from voting in our sample was 80% for the 2012 presidential election, and turnout among those prevented from voting in our sample was 77% in 2012. If the deterred and prevented registrants had voted at their 2012 turnout rates, voter ID could have lowered turnout in 2016 by 1.8 and 0.9 percentage points, respectively.¹² The effects on turnout would be larger if voter ID laws dissuade voters from registering to vote in the first place (Stein and Tchintian 2017). Our data are limited to registrants only, so we cannot estimate how voter ID affects registration.

level data available at <http://elections.wi.gov/node/4420>.

11. The intervals in the right-most column of Table 1 reflect sampling uncertainty around the value in column two.

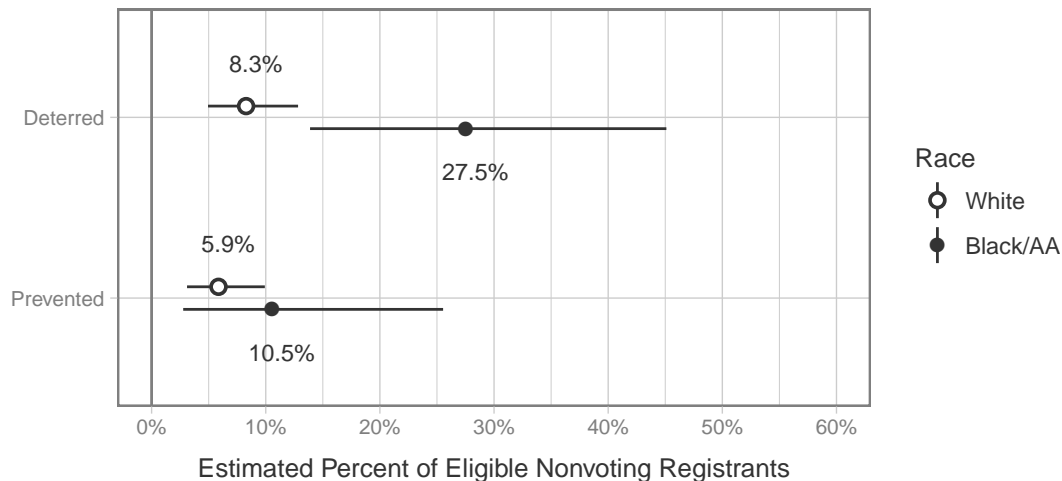
12. Citizens who registered to vote between the 2012 and 2016 elections have no presidential election turnout history to compare to, so we calculate 2012 turnout based on registrants who were registered to vote for the 2012 presidential election or previously. This group of registrants contains all voters with registration dates on or before the date of the 2012 presidential election and—because the voter file lists only the most recent date of registration—any other voter with a validated vote in an election prior to the November 2012 election. The voter file contains elections only as far back as February of 2006, though some registration dates are as early as the 1970s.

DEMOGRAPHIC IMPACT

The results above show the estimated percentage of eligible nonvoting registrants deterred or prevented from voting by the new voter ID requirement. We now turn to an analysis of the demographic impact of these effects.

Race: Figure 2 presents the estimated percentage of White and Black/African-American registrants deterred and prevented from voting. Using the broader definition of the effect (“deterred,” as we do above), we find that about 8.3% of Whites lack ID or mention ID as a reason for not voting, compared 27.5% of Blacks. The confidence intervals for these two estimates do not overlap, indicating that Blacks are statistically significantly more affected than Whites in Dane and Milwaukee Counties.¹³

Figure 2: Estimated percentage of White and Black nonvoting registrants affected by ID law

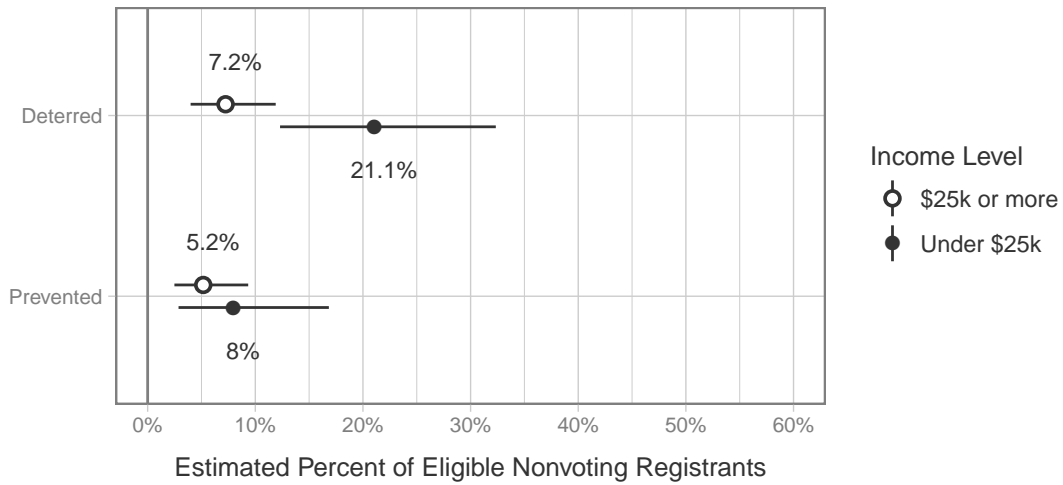


Socioeconomic status: We present an analysis of socioeconomic status in two forms. First, because scholarship on voting and voter turnout suggests that the effect of voter ID laws should be localized to low-SES voters in particular, we compare individuals with incomes lower than \$25,000 to the remainder of the sample. These estimates are presented in Figure 3. An estimated 21.2% of novoting registrants with incomes under \$25,000 are deterred from voting, compared to 7.2% of those with incomes at or above \$25,000—a statistically significant difference. We estimate that 8% of low-income registrants are prevented from voting compared to 5.2% in the remainder of the sample, but this difference is not statistically significant.¹⁴

13. Confidence intervals around the estimates for Black registrants are wider than the intervals for White registrants due to the smaller number of Black registrants in the sample.

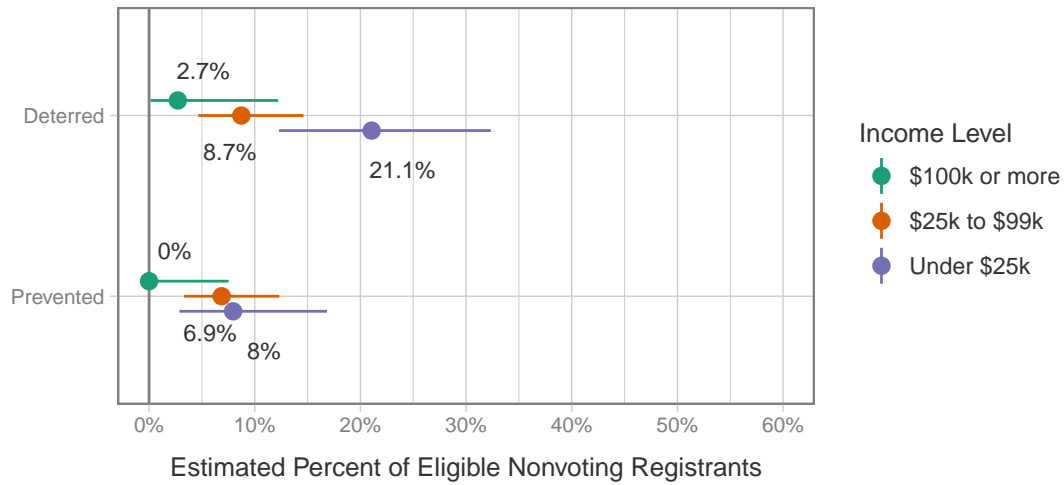
14. Wider confidence intervals are again driven by fewer low-income respondents.

Figure 3: Estimated percentage of nonvoting registrants affected by ID law, by income (two-category)



We also divide registrants with incomes at or above \$25,000 into two categories: high-earning individuals with incomes at or above \$100,000, and a middle category with incomes between \$25 and \$99,000. Figure 4 shows the estimates from this three-category classification. Intervals for these estimates overlap at times, but the data suggest a pattern where low-income registrants are more affected than high-income registrants.

Figure 4: Estimated percentage of nonvoting registrants affected by ID law, by income (three-category)



REFERENCES

- Ansolabehere, Stephen, and Eitan Hersh. 2010. “The quality of voter registration records: A state-by-state analysis.” *Report, Caltech/MIT Voting Technology Project*.
- Clopper, Charles J., and Egon S. Pearson. 1934. “The use of confidence or fiducial limits illustrated in the case of the binomial.” *Biometrika*: 404–413.
- 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*.
- Stein, Robert M., and Carolina Tchintian. 2017. “A butterfly effect: Voter ID laws, voter registration and voter turnout.” *Presented at Election Science, Reform, and Administration Conference, 2017*.

APPENDIX A HIGH-SES AND LOW-SES CENSUS TRACTS IN MILWAUKEE COUNTY

Likely higher SES tracts:

- 168 out of the 296 tracts in Milwaukee County are coded as higher SES (and more likely to respond).
- Mean percent African-American: 6.71%
- Mean percent with household incomes under 200% of the poverty line: 30.58%
- Mean number of non-voting registrants per tract: 473.4
- Total number of non-voting registrants in the 168 “high-SES” tracts: 79,531

Likely lower SES tracts:

- 128 out of the 296 tracts in Milwaukee County are coded as lower SES (and less likely to respond).
- Mean percent African-American: 63.45%
- Mean percent with household incomes under 200% of the poverty line: 67.04%
- Mean number of non-voting registrants per tract: 519.8
- Total number of non-voting registrants in the 128 “low-SES” tracts: 66,532

In keeping with sound statistical practice, estimates are adjusted for oversampling using sampling weights.

APPENDIX B SAMPLE DETAILS AND MARGINALS

Respondent Location:

County	N (Weighted)	Percent
Dane County	96.3	33.7%
Milwaukee County	189.1	66.3%
No Data	0.0	0%

Respondent Gender:

Gender	N (Weighted)	Percent
Male	126.5	44.3%
Female	153.1	53.6%
No Data	5.8	2%

Respondent Race:

Race	N (Weighted)	Percent
White	212.2	74.4%
Black/AA	35.3	12.4%
Asian	12.3	4.3%
Native	2.4	0.8%
Multiple	3.7	1.3%
Other	10.8	3.8%
No Data	8.6	3%

Respondent Income:

Income	N (Weighted)	Percent
Under 25k	71.3	25%
25k to 49k	63.1	22.1%
50k to 74k	42.9	15%
75 to 99k	37.0	13%
100k or more	47.2	16.5%
No Data	24.0	8.4%

Exposure to Voter ID Information:

See Info about Voter ID	N (Weighted)	Percent
Yes	170.2	59.6%
No	107.0	37.5%
No Data	8.2	2.9%

Nominal Reasons for not Voting (weighted responses):

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

Main Reason for not Voting (weighted responses):

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
No Reason 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