Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach

Brendan Nyhan† Christopher Skovron‡ Rocío Titiunik§

August 22, 2015

Abstract

The widespread availability of voter files has changed the study of participation in American politics. However, most studies calculate voter turnout among registrants, which may lead to invalid causal inferences if registration rates differ between treatment and control groups. We introduce a sensitivity analysis approach that can be used to assess the potential bias induced by this problem, which we call differential registration bias, and illustrate its use with two empirical studies of the effects of voting eligibility on subsequent turnout among young voters. In both cases, we find that eligibility decreases turnout as a proportion of total registration. However, these negative effects are highly sensitive to differential registration bias and, in one study, change signs when we instead calculate turnout as a proportion of the population. These results suggest the need for scholars to provide rigorous sensitivity analysis as part of any study of voter file data.

†Assistant Professor, Department of Government, Dartmouth College, nyhan@dartmouth.edu
‡Ph.D. Candidate, Department of Political Science, University of Michigan, cskovron@umich.edu.
§Assistant Professor, Department of Political Science, University of Michigan, titiunik@umich.edu.
The widespread availability of digital voter files from state and local governments has changed the study of voting behavior in American politics. It is now easier than ever for scholars to examine voter behavior using extremely large datasets, enabling new studies of how factors as disparate as majority-minority districts (Barreto, Segura and Woods 2004), minority candidates (Barreto 2007; Fraga 2015), genetic similarity (Fowler, Baker and Dawes 2008), or early voting registration (Holbein and Hillygus 2015) affect turnout at the individual or precinct level. These data have been found to be of high quality, especially when cleaned and aggregated by third party firms like Catalist (Ansolabehere and Hersh 2012; Hersh 2015).

However, studying voter file data necessarily restricts the analysis to registered citizens, which poses significant challenges for causal inference. For many causal factors of interest (e.g., incumbents’ ethnicity or individual voting eligibility), the intervention or “treatment” is likely to affect the likelihood of registration as well as turnout, raising concerns about endogenous selection bias (see, e.g. Elwert and Winship 2014), a form of post-treatment bias (Rosenbaum 1984). Ideally, we would simply analyze the relationship between the treatment of interest and turnout rates without conditioning on the intermediate variable of registration, but analysts of voter files almost always lack comprehensive individual-level data on non-registrants.1

We formalize a definition of bias from differential registration, which demonstrates the threat it poses to valid causal inferences. When possible, scholars should instead adjust turnout rates to the population using data from alternative sources. If these data are not available, however, an alternative approach is needed. We therefore develop a novel sensitivity analysis that makes it possible for scholars to characterize the robustness of their treatment effect estimates to (unobserved) potential differences in registration rates between treatment and control groups.

We illustrate this problem with two studies of political socialization, a research area which has recently begun to use quasi-experimental techniques to estimate the effect of initial election eligibility on subsequent voter turnout and other behavior (e.g., Meredith 2009; Mullainathan and Washington 2009; Dinas 2012, 2014; Coppock and Green 2015; Holbein and Hillygus 2015). These studies typically compare voter turnout and political attitudes among individuals whose 18th birthday fell close to a previous general election. Within a window around the eligibility

---

1This concern is not directly relevant to field experiments that use voter files to identify a sample of interest and conduct a randomized study of the effect of treatment among a group of registrants.
cutoff (turning 18 by election day), individuals who could have taken part in that election and those who could not are assumed to be otherwise identical, differing only in the exact timing of their birth. We use a similar research design that compares turnout rates between young citizens who were narrowly eligible or ineligible to vote in a prior election.

Our first study is an original analysis of a comprehensive sample of registrants from 42 U.S. states and the District of Columbia who were born within four days of the election eligibility cutoff—a far narrower window than previous studies (and thus one that is more likely to satisfy an as-if random assumption). Our empirical analysis of the effects of voting eligibility on subsequent turnout based on this data demonstrates the need to consider the potential bias induced by conditioning on registration. When we naïvely use registered voters in the voter file as the denominator for our turnout analysis, we find results that largely contradict previous studies—initial election eligibility appears to sometimes reduce subsequent turnout or have no effect. However, we show that this result is highly sensitive to differing registration rates between the two groups: a pattern where the registration rate in the treatment group were just five percent higher than in the control group would be sufficient to reverse the negative effect. Consistent with this sensitivity analysis, when we instead use birth totals as the denominator, we find that initial eligibility increases subsequent turnout. This empirical illustration shows that differential patterns of registration between treatment and control groups can severely affect studies of voting behavior and lead to misleading conclusions when turnout rates are calculated as a proportion of total registration.

Our second study is a re-analysis of the Florida analysis in Holbein and Hillygus (2015), where we again examine the effect of voting eligibility on subsequent turnout. Like the first study, when turnout rates are calculated among registered voters, eligibility is found to negatively affect turnout in later elections among young voters but a sensitivity analysis of the data also indicates that even a slightly higher registration rate in the treatment group would reverse these results. However, unlike Study 1, we find that the registration rate is higher in the control group than in the treatment group when we use birth counts to approximate the total population, which suggests that the negative effects Holbein and Hillygus report are not the result of differential registration bias.

These empirical studies illustrate how even small differences in registration rates between treatment and control groups can reverse the conclusions of turnout studies based on registration files.
Our results suggest that scholars who study turnout in voter files should always present a rigorous sensitivity analysis to address concerns about differential registration bias and, whenever possible, estimate registration rates using total population counts from alternative sources.

**Challenges to the study of turnout from voter files**

Voter files are everywhere in the study of contemporary American politics. Especially after being cleaned and aggregated by firms like Catalist, they offer high-quality data on a vast population of U.S. citizens while avoiding the social desirability bias and low statistical power that plague survey studies of self-reported turnout (Ansolabehere and Hersh 2012; Hersh 2015). In particular, these data make it possible to consider research questions that would be difficult or impossible to address in conventional survey samples such as sub-national variation in turnout rates among specific subgroups of interest (see, e.g. Barreto, Segura and Woods 2004; Barreto 2007; Fraga 2015; Holbein and Hillygus 2015).

However, citizen must register to be included in a voter file—an action that is plausibly affected by many independent variables that are also theorized to affect voter turnout. In estimating the causal effect of a treatment only among registered voters, analysts may unwittingly induce a form of post-treatment bias (Rosenbaum 1984) sometimes known as “endogenous selection bias” (Elwert and Winship 2014).² Our particular concern here is the possibility of differential registration between the treatment and control groups. If a treatment affects registration rates as well as voter turnout (the outcome of interest), a simple comparison of turnout rates between treated and control groups among registered voters could lead to mistaken inferences.³

Table 1 illustrates the problem of differential registration bias with a hypothetical example of an election in which 10,000 citizens are in the treatment group and 10,000 citizens are in the control group.⁴ When registration rates differ by treatment status, we can find spurious results. In this case, treatment is associated with higher rates of registration—5,000 individuals in the treatment

---

²We assume that analysts wish to make inferences about the population as a whole.
³For ease of exposition, we assume a binary treatment corresponding to our application below, but our argument also applies straightforwardly to discrete multi-valued treatment variables.
⁴As noted above, we focus on observational studies in which analysts seek to estimate the causal effect of a given treatment from voter file data. Providing convincing evidence for such an inference will of course require addressing other concerns besides differential registration bias that we do not consider here (see, e.g., Imbens and Rubin 2015).
group register compared with 4,000 in the control group (column 2)—but has no effect on turnout (columns 3 and 5). As a result, if we simply compare turnout rates among registered voters using voter file data (column 4), we would conclude that turnout rates are lower among treated voters (2500/5000=50% versus 2500/4000=62.5% among controls) even though the true turnout rate of 25% is identical in both groups (column 5).

Table 1: Hypothetical example of differential registration bias

<table>
<thead>
<tr>
<th></th>
<th>Total (1)</th>
<th>Registered (2)</th>
<th>Voted (3)</th>
<th>Turnout (% reg.) (4)</th>
<th>Turnout (% pop.) (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated</td>
<td>10,000</td>
<td>5,000</td>
<td>2,500</td>
<td>50%</td>
<td>25%</td>
</tr>
<tr>
<td>Control</td>
<td>10,000</td>
<td>4,000</td>
<td>2,500</td>
<td>62.5%</td>
<td>25%</td>
</tr>
</tbody>
</table>

This problem would of course be avoided if we could calculate turnout rates using the total voting-eligible population in each group rather than the total number of registered voters in each group. When possible, the simplest solution to the problem of differential registration is to obtain the missing population totals from alternative data sources. Often, however, the necessary data are not available or are imperfect. For instance, in the case of the two voter eligibility studies we analyze below, the ideal alternative data source would be a U.S. Census count of the citizen population born on the days of interest. Because these data are not publicly available, however, we instead use birth counts by day, which are a plausible alternative assuming inaccuracies resulting from foreign births, interstate migration, deaths, etc., are not correlated with treatment status.

We formalize a sensitivity analysis approach (Rosenbaum 2005) that can be implemented in cases where the needed quantities—the total population in the treatment and control groups—are unknown or imprecisely estimated. Specifically, our approach identifies the differential in registration rates between the treatment and control groups that would drive the observed difference in turnout rates among registered voters to zero. This method is applicable to the study of any treatment that affects voting behavior, including observational studies, quasi-experimental studies, and randomized experiments, providing a measure of the vulnerability of a treatment effect estimate to differential registration bias when precise population totals are missing. The method can also be

---

5 This approach mirrors, for example, Imai, Keele and Yamamoto (2010), who provide a sensitivity test for mediation that can be used to calculate the level of unobserved correlation in errors between the mediator and outcome equations for which the average causal mediation effect is zero.
used to corroborate findings when population totals from alternative data sources are likely to be estimated with error, which is a pervasive phenomenon since precise counts of the voting-eligible population are not directly available (McDonald and Popkin 2001).

In the example introducing differential registration bias above, we assumed that the true population counts were known. However, we now assume that total population counts for the treatment and control groups are not available and that researchers are working with a voter file that includes only registered voters.\(^6\) This scenario is illustrated in Table 2, where for each group we have the total registration (column 2) and total number of voters (column 3), which allows us to calculate turnout rates as a proportion of total registration in each group (column 4). However, letting the subscripts \(T\) and \(C\) denote the treatment and control groups, respectively, the total population counts in each group, which we denote by \(P_T\) and \(P_C\), are not available. As a consequence, the true turnout rates (i.e., the ratio of voters to the total population in each group) are also unavailable, creating the risk of a mistaken inference if differential registration bias is present.

Table 2: Hypothetical illustration of sensitivity analysis

<table>
<thead>
<tr>
<th>Overall population</th>
<th>Total (1)</th>
<th>Registered (2)</th>
<th>Voted (3)</th>
<th>Turnout (% reg.) (4)</th>
<th>Turnout (% pop.) (5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treated (P_T)</td>
<td>5,000</td>
<td>2,500</td>
<td>50%</td>
<td>2,500/(P_T)</td>
<td></td>
</tr>
<tr>
<td>Control (P_C)</td>
<td>4,000</td>
<td>2,500</td>
<td>62.5%</td>
<td>2,500/(P_C)</td>
<td></td>
</tr>
</tbody>
</table>

We introduce some additional notation. We let \(R_T\) and \(R_C\) denote the total registration counts in the treatment and control groups, respectively, and \(V_T\) and \(V_C\) denote the total numbers of voters who turned out in each group. The desired—but unavailable—turnout rates are thus

\[ T_{\text{Pop}}^T = \frac{V_T}{P_T} \quad \text{and} \quad T_{\text{Pop}}^C = \frac{V_C}{P_C}, \]

where the superscript \(\text{Pop}\) denotes that these turnout rates are calculated as a proportion of the total population. Henceforth, we refer to \(T_{\text{Pop}}^T\) and \(T_{\text{Pop}}^C\) as the *turnout-to-population* rates or simply the *true turnout* rates. However, without observing the total population counts \(P_T\) and \(P_C\), we cannot calculate these turnout-to-population rates directly. We define the registration rates \(r_T = \frac{R_T}{P_T}\)

---

\(^6\)For simplicity, we assume the entire registration file is available. However, our argument applies directly to the case where only a random sample from the voter file is available (as in Study 1 below).
and \( r_c = R_c / P_c \) in the treatment and control groups, respectively. Given these rates, the total registration counts, \( R_T \) and \( R_C \), can be expressed as \( R_T = r_T \times P_T \) and \( R_C = r_C \times P_C \), and we can express the unknown total population as total registration divided by the probability of registration:

\[
P_T = R_T / r_T \quad \text{and} \quad P_C = R_C / r_C.
\]

We then define the turnout-to-registration rate in each group, \( T_{T}^{\text{Reg}} \) and \( T_{C}^{\text{Reg}} \), as follows:

\[
T_{T}^{\text{Reg}} = \frac{V_T}{R_T} \quad \text{and} \quad T_{C}^{\text{Reg}} = \frac{V_C}{R_C}.
\]

Given these definitions, we can express the desired turnout-to-population rates as

\[
T_{T}^{\text{Pop}} = \frac{V_T}{P_T} = \frac{V_T}{R_T / r_T} = T_{T}^{\text{Reg}} \times r_T \quad \text{and} \quad T_{C}^{\text{Pop}} = \frac{V_C}{P_C} = \frac{V_C}{R_C / r_C} = T_{C}^{\text{Reg}} \times r_C.
\]

(1)

In words, the true turnout rate is the turnout-to-registration rate adjusted (multiplied) by the registration rate in each group. In applications where researchers have access to the voter file but not the total population, the registration rates \( r_T \) and \( r_C \) are unknown. Our sensitivity analysis considers how large the difference between \( r_T \) and \( r_C \) would have to be to generate observed difference in turnout-to-registration rates when the true difference in turnout-to-population rates is zero.

Imagine that \( T_{T}^{\text{Reg}} < T_{C}^{\text{Reg}} \), which means that the turnout-to-registration rate is smaller in the treatment group than in the control group (as in the two applications we consider below). Because the registration rates for the two groups \( r_T \) and \( r_C \) are unknown, \( T_{T}^{\text{Reg}} < T_{C}^{\text{Reg}} \) is not enough to conclude that \( T_{T}^{\text{Pop}} < T_{C}^{\text{Pop}} \).\(^7\) In words, a lower (higher) turnout-to-registration rate among the treatment group does not necessarily imply they have a lower (higher) turnout-to-population rate. However, given the observed difference \( T_{T}^{\text{Reg}} - T_{C}^{\text{Reg}} \), we can estimate how different the registration rates would have to be between the treatment and the control groups for this difference to be observed if there is no difference in turnout-to-population rates (i.e., \( T_{T}^{\text{Pop}} - T_{C}^{\text{Pop}} = 0 \)).

To illustrate our approach, we define \( \Delta \) as the treatment-control difference in true turnout rates:

\[
\Delta = T_{T}^{\text{Pop}} - T_{C}^{\text{Pop}} = (T_{T}^{\text{Reg}} \times r_T) - (T_{C}^{\text{Reg}} \times r_C).
\]

\(^7\)Likewise, we can’t conclude that \( T_{T}^{\text{Pop}} > T_{C}^{\text{Pop}} \) or that \( T_{C}^{\text{Pop}} = T_{T}^{\text{Pop}} \) in this case either without knowing \( r_T \) and \( r_C \).
We note two points. First, if \( r_T = r_C = 1 \), the expression simplifies to \( \Delta = T_T^{\text{Reg}} - T_C^{\text{Reg}} \). In other words, if everyone in the treatment and control groups registers, the ratio of voters to registrants is of course identical to the ratio of voters to total population and there are no complications. Second, if turnout rates are identical between groups but not everyone votes (\( r_T = r_C = r \neq 1 \)), the unknown turnout share difference simplifies to \( \Delta = r(T_T^{\text{Reg}} - T_C^{\text{Reg}}) \). Since \( 0 \leq r \leq 1 \), the sign of \( \Delta \) will be equal in this case to the sign of \( T_T^{\text{Reg}} - T_C^{\text{Reg}} \) and \( |\Delta| \leq |T_T^{\text{Reg}} - T_C^{\text{Reg}}| \); moreover, it is straightforward to explore how \( \Delta \) changes as \( r \) varies from 0 to 1.

We are interested in the more general case in which both \( r_T \) and \( r_C \) are nonzero and \( r_T \neq r_C \), however. Figure 1 illustrates how a negative “treatment effect” on turnout-to-registration rates can be observed even when the true difference in the turnout-to-population rates is null or even positive. Using the functions defined in Equation (1), the figure plots the turnout-to-population rates \( (T_T^\text{pop}, T_C^\text{pop}) \) against the registration rates \( (r_T, r_C) \) separately for each group holding turnout-to-registration rates \( (T_T^{\text{Reg}}, T_C^{\text{Reg}}) \) fixed. We adopt the scenario in Tables 1 and 2 where \( T_C^{\text{Reg}} = 0.50 \) and \( T_C^{\text{Reg}} = 0.625 \). Thus, we plot the linear functions \( T_T^{\text{pop}} = 0.50 \times r_T \) and \( T_C^{\text{pop}} = 0.625 \times r_C \). Since the slope in the control group (0.625) is higher than the slope in the treatment group (0.50), the dotted line (control) is always above the solid line (treatment). In other words, Figure 1 fixes \( T_C^{\text{Reg}} = 0.50 \) and \( T_C^{\text{Reg}} = 0.625 \), which means that the turnout-to-registration difference is always negative (-0.125).

Imagine first that the probability of registration in the treatment group is 0.8 (\( r_T = 0.8 \)) and the probability of registration in the control group is 0.4 (\( r_C = 0.4 \)). In this case, the true turnout rate in the treatment group \( (T_T^{\text{pop}}) \) of 0.4 is obtained from the y-coordinate of point A on the solid line in Figure 1(a), and the true turnout rate of 0.25 in the control group \( (T_C^{\text{pop}}) \) is obtained from the y-coordinate of point B on the dashed line. Under this scenario, \( \Delta = T_T^{\text{pop}} - T_C^{\text{pop}} = 0.4 - 0.25 = 0.15 \). In other words, the treatment effect on the true turnout-to-population rate is positive despite the turnout-to-registration rate being lower in the treatment than in the control group. Figure 1(b) shows a different scenario in which we assume that \( r_T = 0.8 \) and \( r_C = 0.64 \). In this case, both \( T_T^{\text{pop}} \) and \( T_C^{\text{pop}} \) are equal to 0.4 and thus \( \Delta = 0.4 - 0.4 = 0 \), though the difference in turnout-to-registration rates is still negative (again, treatment slope is 0.5 and control slope is 0.625).

We propose a sensitivity analysis to explore how our inferences are affected by differential registration bias under these circumstances (the most likely scenario in practice). First, define the
differential registration factor, $k$, as the ratio of the registration rate in the treatment group to the registration rate in the control group:

$$k = \frac{r_T}{r_C}$$

We have assumed that both $r_T$ and $r_C$ are nonzero so this ratio is well-defined. Moreover, because $r_T$ and $r_C$ are rates or probabilities, they are both less than (or at most equal to) 1, which means that $k \in (0, \infty)$. Furthermore, since $r_T/k = r_C$ and $r_C$ is a rate, $k$ must satisfy the restriction $0 < r_T/k \leq 1$; that is, the smallest value that $k$ can take is $r_T$.\(^8\) Our sensitivity analysis explores, for a given treatment group registration rate $r_T$, how large the differential registration factor $k$ can be before the implied difference in turnout-to-population rates is zero or has the opposite sign from $T_T^{\text{Reg}} - T_C^{\text{Reg}}$, the observed difference in turnout-to-registration rates.

Given our definition of $k$, we can express $\Delta$, the treatment-control difference in true turnout rates, as a function of the treatment group registration rate $r_T$ and the ratio of treatment/control

\(^8\)In other words, values of $k \in (0, r_T)$ are not allowed because they would imply $r_C \geq 1$.\)
registration rates \( k \):

\[
\Delta(r_T, k) = (T_T^\text{Reg} \times r_T) - (T_c^\text{Reg} \times \frac{r_T}{k}) = r_T(T_T^\text{Reg} - \frac{T_c^\text{Reg}}{k})
\]

where we make the arguments \( k \) and \( r_T \) explicit. Thus, for any nonzero value of \( r_T \), we can calculate the value of \( k \) under which a zero difference in true turnout rates between the treatment and control groups would result in the observed difference in turnout-to-registration rates. We find this value, which we call \( k^* \), as the solution to \((T_T^\text{Reg} - \frac{T_c^\text{Reg}}{k^*}) = 0\), which implies that

\[
k^* = \frac{T_c^\text{Reg}}{T_T^\text{Reg}}.
\]

By definition, \( \Delta(r_T, k^*) = 0 \) for any \( 0 < r_T \leq 1 \).

This sensitivity analysis approach thus allows us to explore how the true turnout difference \( \Delta \) varies with observed turnout-to-registration rates under different assumptions about registration rates for the treatment and control groups (\( r_T \) and \( r_C \)). In particular, we can calculate \( k^* \), the pattern of differential registration that would be required to produce the observed difference in turnout-to-registration rates assuming that there is no difference in the turnout-to-population rates.

We again illustrate the procedure with the hypothetical example presented in Tables 1 and 2. In that example, \( T_T^\text{Reg} = 2500/5000 = 0.5 \) and \( T_c^\text{Reg} = 2500/4000 = 0.625 \), so the difference in turnout rates among registered voters is negative \((T_T^\text{Reg} - T_c^\text{Reg} = -0.125)\). In this case, \( k^* = 0.625/0.50 = 1.25 \), which means that the observed difference in turnout-to-registration rates could occur under a zero (or positive) difference in true turnout rates if the probability of registration were (more than) 25% higher in the treatment than in the control group.

Figure 2 visualizes how the true turnout difference can vary with \( k \) — the ratio of the treatment/control registration rates — for a given difference in turnout-to-registration rates. We plot \( \Delta(r_T, k) \) as a function of \( r_T \) for differing values of \( k \) (which implicitly fixes \( r_C \)). The y-axis is the difference in turnout-to-population rates between the treatment and control groups and the x-axis is the registration rate in the treatment group. As illustrated by the blue curve, when \( k = k^* = 1.25 \), the difference in true turnout rates is zero for every value of \( r_T \). Moreover, when the registration rate is equal in both groups and thus \( k = 1 \), the observed difference in turnout-to-registration rates
$(T_{T}^{\text{Reg}} - T_{C}^{\text{Reg}} = -0.125)$ is equal to the true difference in turnout rates if $r_{T} = 1$.

We can explore sensitivity further in our hypothetical example by calculating the turnout rate difference $\Delta$ for values of the registration ratio $k$ above and below the threshold $k^\star$. For $k > k^\star = 1.25$, the difference in true turnout is positive and the sign of the turnout-to-registration rate difference is reversed. For $1 \leq k < k^\star = 1.25$, the true turnout effect is negative but smaller in absolute value than the difference in turnout-to-registration rates for all $r_{T}$. Finally, for $k < 1$, the true turnout effect is negative and can be larger in absolute value than the difference in turnout-to-registration rates for a high enough $r_{T}$.

In sum, this simple sensitivity analysis allows researchers to fully explore how severe the pattern of differential registration between the treated and control groups would have to be to reverse the sign of the estimated difference in turnout-to-registration rates. Below we illustrate how to apply this approach in two studies of the effects of initial voter eligibility on turnout in subsequent
elections. Though these studies leverage discontinuities in voting eligibility, it is important to note that our approach does not require the use of a regression discontinuity design and can be applied generally to experimental, quasi-experimental, and observational studies of voter turnout.

**Application: The effects of election eligibility on subsequent turnout**

We illustrate the problem of differential registration bias and our proposed solutions with two empirical studies of political socialization that focus on the relationship between voting eligibility and subsequent voter turnout. In Study 1, we present an original analysis of voter file data from 42 states. In Study 2, we replicate results in a recent article by Holbein and Hillygus (2015). In both cases, the empirical strategy is based on a comparison of voters who were born close to an election eligibility cutoff, which we describe in more detail below.

Research in political socialization has found long-lasting effects of early experiences and events like parent socialization (Niemi and Jennings 1991; Jennings, Stoker and Bowers 2009) and draft status during the Vietnam War (Erikson and Stoker 2011). The most common and important socializing events for many people as they approach or enter adulthood are elections—the time when politics is most salient in national life and during which people form political identities and establish patterns of behavior that often persist over the life cycle (e.g., Jennings and Markus 1984; Alwin and Krosnick 1991; Sears and Funk 1999; Green, Palmquist and Schickler 2004; Prior 2010; Ghitza and Gelman N.d.). Sears and Valentino (1997), for instance, find that presidential elections appear to be especially potent in forming the political views of adolescents. These findings echo coverage of the putative effects of the 2008 Obama campaign on young people, which have also made the effects of early political experiences particularly salient in recent years.9

These questions have been re-engaged by an emerging literature that studies the effects of initial election eligibility on voter turnout and other political behavior using a quasi-experimental approach (e.g., Meredith 2009; Dinas 2012; Coppock and Green 2015; Holbein and Hillygus 2015). By comparing later turnout and political attitudes among a group of voters whose 18th birthday fell very close to a general election, these studies seek to leverage as-if random variation in birth

---

9One article in *The Atlantic* asked, for instance, “Has Obama Turned a Generation of Voters Into Lifelong Democrats?” (Ball 2013).
timing to compare individuals who had the opportunity to take part in an election and those who did not but are assumed to be otherwise identical. This research strategy, which is based on the discontinuous change in voting eligibility that occurs at election day, is an application of a regression discontinuity (RD) design in which a treatment is given on the basis of whether a score exceeds a known cutoff.

The recent literature on election eligibility effects has implemented this research strategy in different ways. Meredith (2009) uses a difference-in-differences design to estimate the effect of election eligibility on subsequent turnout among registered voters in California. Dinas (2012) instead uses eligibility as an instrument to estimate the downstream effect of self-reported turnout among respondents to the Youth-Parent Socialization Panel Study. Most recently, Coppock and Green (2015) estimate an RD model that finds significant positive effects of initial election eligibility on subsequent vote totals (not turnout rates) in 17 states. By contrast, Holbein and Hillygus (2015) conduct an RD analysis of the Florida voter file and find that just-ineligibles are more likely to vote than just-eligibles in subsequent elections.¹⁰

These RD-style designs are a promising approach to estimating election socialization effects because they can be implemented at a relatively low cost using voter files, but the validity of the results depends on the assumption that the treatment was randomly or as-if randomly assigned—that is, the assumption that the there are no systematic differences between the groups that are being compared except for eligibility status.¹¹ The plausibility of this assumption will of course vary according to the specific details of each implementation. Our focus here, however, is on illustrating the inferential challenges posed by differential registration bias and demonstrating how to address those concerns using our sensitivity analysis approach and alternative population estimates.

¹⁰In addition, two other studies use RD-style designs to estimate the effect of election eligibility on political attitudes. Mullainathan and Washington (2009) find that eligibles are more polarized than in eligibles using a four-year window of 18- to 21-year-olds in American National Election Studies (ANES) data, and Dinas (2014) finds that the act of voting for a party reinforces support for that party in the future using just-eligible or just-ineligible respondents from the Youth-Parent Socialization Panel Study. We focus only on turnout effects in this article.

¹¹Several of the studies described above also assume that election eligibility affects subsequent voting only through its effect on current voting, allowing them to test a habit-formation hypothesis (Meredith 2009; Dinas 2012; Coppock and Green 2015). We do not consider the validity of this additional assumption here and instead focus on the direct effect of election eligibility on subsequent voting.
Studying eligibility effects with a regression discontinuity design

The two examples we analyze below are based on an RD design of the type described above. We therefore discuss this design in more detail before turning to the presentation of each study. The defining feature of an RD design is that subjects are assigned a score and receive treatment if their score exceeds a known cutoff. In the U.S., a discontinuity in voting eligibility occurs when citizens turn eighteen years of age. As a result of the 26th Amendment to the U.S. Constitution (adopted in 1971), people who turn eighteen on or before election day can cast a vote but those who will turn eighteen after election day are ineligible to vote. Thus, date of birth exactly determines voting eligibility and an RD design can be used to study the effects of eligibility on turnout.\textsuperscript{12}

An important feature distinguishing RD designs based on date of birth from most uses of RD is that the score that determines treatment, birthdate, is a discrete variable, which invalidates most identification and estimation results in the RD literature. To address this issue, we adopt the Cattaneo, Frandsen and Titiunik (2015) framework, which interprets the RD design as a local randomized experiment in a fixed window around the cutoff. Importantly, this framework does not require a continuous score for identification or inference. In our context, the RD approach entails assuming that voting eligibility is as-if randomly assigned for people with birthdays near election day.

In order to adopt this local experiment framework, it is crucial to focus on individuals who are born just a few days apart: although factors such as education and civic engagement may differ between 17-year-olds and 19-year-olds, we do not expect major differences between someone who is exactly 18 on election day and someone born just a few days later. Thus, in Study 1, we focus our analysis on individuals who turn eighteen within eight days of election day and assume that eligibility can be considered as-if randomly assigned between those individuals born on election day or the three days earlier (the treatment group) and those born one to four days later (the control group). In Study 2, we use a wider window of one month on either side of election day to ensure comparability with the original analysis in Holbein and Hillygus (2015).

\textsuperscript{12}In conventional RDs, the assigned score is a continuous random variable. If we assume unobserved confounders vary smoothly at the cutoff, the effects of the treatment on the outcome can be estimated by comparing treated and control subjects near the cutoff (for recent reviews, see Imbens and Lemieux 2008 and Lee and Lemieux 2010). More precisely, assuming that potential outcomes are continuous at the cutoff, the average treatment effect at the cutoff can be recovered from the conditional expectation of the observed outcomes given the score (Hahn, Todd and van der Klaauw 2001). Estimation of RD effects based on these continuity assumptions typically relies on local nonparametric estimation (see, e.g., Porter N.d., Imbens and Kalyanaraman 2012, and Calonico, Cattaneo and Titiunik 2014).
Both studies estimate the effects of voting eligibility on subsequent turnout using voter file data, which means they are vulnerable to differential registration bias. Specifically, if just-eligibles are more likely to be registered than just-ineligibles due to the longer period in which they could participate in the political process or be mobilized by campaigns, we could observe a spurious differential in turnout-to-registration rates even if there is no difference in the underlying true turnout rates between the two populations. We therefore use two strategies to assess the robustness of the results. First, we employ the sensitivity analysis approach described above to test how robust our turnout-to-registration results are to potential differences in registration rates between the treatment and control groups. Second, we use estimates of birth counts by date to instead estimate turnout-to-population rates for the treatment and control groups.

**Study 1: Voter eligibility effects in Catalist data**

Our first study is an original application that investigates the effects of voting eligibility on subsequent turnout with an RD design based only on the closest observations to the election day cutoff. Specifically, we examine three cohorts who were narrowly (in)eligible to vote in the 2004, 2006, and 2008 elections, considering only those registrants born within just four days of the election eligibility cutoff—a far narrower window than previous studies, which have used windows measured in months (Meredith 2009; Dinas 2012, 2014; Holbein and Hillygus 2015) or years (Mullainathan and Washington 2009; Coppock and Green 2015).

In addition, our sample is the most comprehensive of those considered in this literature to date, which often draw data from a single state voter file (Meredith 2009; Holbein and Hillygus 2015) or small national surveys and panel studies (Mullainathan and Washington 2009; Dinas 2012, 2014, but see Coppock and Green 2015) and also often only consider eligibility in one election (Dinas 2012, 2014). By contrast, our data are drawn from voter files in 42 U.S. states and the District of Columbia and include eligibility variation and turnout data from several national elections. Our data source is voter registration files that were collected, cleaned, and supplemented by the private

---

13 As mentioned above, our focus is not on the details of the RD design per se. In these applications, we use RD because it provides treatment and control groups that are plausibly comparable and has been used in a number of studies to estimate turnout effects from voter files.
company Catalist.\textsuperscript{14} We collect a random sample of voters in the Catalist file born in the eight days around the cutoff date for being eligible to vote (i.e., for being 18 years old on or before election day) in the 2004, 2006, and 2008 elections.\textsuperscript{15}

The three cohorts of individuals in our data were born in 1986, 1988, and 1990, respectively. For example, the 1990 cohort treatment group was born from November 1–4 and were thus eighteen years old and eligible to vote on November 4, 2008, while the control group was born from November 5–8, 1990. Unfortunately, Catalist’s data on unregistered voters are sparse and unreliable, which forces us to focus—like other analysts—on the universe of registrants and thereby introduces the possibility of differential registration bias.\textsuperscript{16} The final dataset includes a total of 49,493 observations in our target windows among the three birth cohorts.

**Effects of eligibility on turnout-to-registration rates**

Before turning to our results, Table 3 explains how we present the findings from our analysis. We compare the behavior of the treatment group of just-eligibles—those who were born just before the election eligibility cutoff—with the control group of just-ineligible voters born just after the cutoff in later elections. The election in the year the cohort turned 18 is denoted E1 and subsequent elections are denoted E2, E3, and E4. For instance, E1 for the 1986 cohort is the 2004 election and the 2006, 2008, and 2010 elections are E2, E3, and E4, respectively, for that cohort.

We analyze the effect of election eligibility in the year voters turned 18 on subsequent turnout-to-registration rates in Table 4, which compares just-eligible and just-ineligible voters who were born in the week surrounding the eligibility cutoff.\textsuperscript{17} These findings initially seem to contradict

\textsuperscript{14}Colorado, Massachusetts, New Jersey, Oklahoma, South Carolina, Vermont, and Washington were excluded due to school entry cutoff dates that overlapped with the election eligibility window, creating a potential confound. Illinois was excluded due to legal restrictions on state voter file use.

\textsuperscript{15}Ansolabehere and Hersh (2010) use Catalist data to analyze the quality of state voter files and find that “Identifying information such as birthdates are generally well collected.” They do identify some problems with missing birth dates and unusual concentrations of voters with birthdates recorded on particular dates but these should not affect the validity of our design. (The only unusual date within our window is November 11, which they find to be unusually prevalent among Texas registrants. Respondents born on November 11, 1988 are part of our sample but we observe no evidence of an excess concentration among respondents in Texas [results available upon request].) Catalist not only cleans and processes data from state voter files but fills in exact birth dates from commercial sources when possible for states that only release month of birth, allowing us to use exact birthdates even in states that do not release them.

\textsuperscript{16}We also drop all observations missing exact birthdates, those with birthdates outside the target range, and those recorded as voting in elections for which they should have been ineligible given their reported birthdate.

\textsuperscript{17}Balance checks reported in the appendix show little imbalance on observables among the registrants in our data.
Table 3: Birth years and election years in 2011 Catalist data

<table>
<thead>
<tr>
<th>Year</th>
<th>E1</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1988</td>
<td>2006</td>
<td>2008</td>
<td>2010</td>
<td>-</td>
</tr>
<tr>
<td>1990</td>
<td>2008</td>
<td>2010</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4: Turnout-to-registration rates by voting eligibility

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>A. 1986 cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eligibility effect</td>
<td>2.12</td>
<td>-2.16</td>
<td>-2.56</td>
<td>(1.07, 3.17)</td>
<td>14.59</td>
</tr>
<tr>
<td>Control group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>52.18</td>
<td>20.37</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>B. 1988 cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eligibility effect</td>
<td>-1.50</td>
<td>-1.91</td>
<td>-</td>
<td>(-2.99, -0.01)</td>
<td>55.51</td>
</tr>
<tr>
<td>Control group</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>20.11</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>C. 1990 cohort</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Eligibility effect</td>
<td>-3.06</td>
<td>-</td>
<td>-</td>
<td>(-4.45, -1.66)</td>
<td>23.10</td>
</tr>
<tr>
<td>Control group</td>
<td></td>
<td></td>
<td>-</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>


findings that eligibility increases subsequent turnout (e.g., Meredith 2009; Dinas 2012; Coppock and Green 2015). While we find a significant positive effect of eligibility on turnout-to-registration rates for the 1986 cohort in the 2006 election, the estimated effect is negative and significant for the 1986 cohort in the 2008 and 2010 elections, the 1988 in the 2008 and 2010 elections, and the 1990 cohort in the 2010 election.

Specifically, registered voters who were born in 1986 and were just eligible to vote in 2004 were significantly more likely to turn out in 2006 than those who were just ineligible. The estimated treatment effect is approximately two percentage points (2.12 percentage points, 95% CI: 1.07, 3.17), which is a substantial increase relative to the low baseline turnout rate for young voters in a midterm election but relatively modest in absolute terms. However, this apparent effect reverses by the second and third subsequent elections—just-eligible voters born in 1986 were significantly less
likely to vote in 2008 and 2010 than their just-ineligible counterparts among the registered voters in our data. We find a similar negative relationship between eligibility and subsequent turnout-to-registration rates for just-eligible registered voters born in 1988 cohort in 2008 and 2010 and the 1990 cohort in 2010. (A full set of RD plots illustrating these estimates is included in the appendix.)

Unfortunately, these findings condition on registration and are therefore vulnerable to differential registration bias. As we show below, previous election eligibility is estimated to have a positive rather than negative effect on subsequent turnout when we condition on total population — a reversal that is the result of differential registration rates in the two groups.

**Sensitivity analysis: Assessing differential registration scenarios**

We now conduct a sensitivity analysis to diagnose how robust the results above are to differential registration bias. Table 5 presents the sensitivity tests for our data. Again, the key term is $k^*$, which is the ratio of registration between the treatment and control groups that could produce the observed difference in turnout-to-registration rates if the true difference in the turnout-to-population rates were zero. When $k^*$ is close to 1, the result is highly sensitive to a small difference in registration rates. By contrast, a $k^*$ that is farther from 1 indicates that the result is more robust.$^{18}$

These results indicate that the positive eligibility effect we observe for turnout-to-registration rates in the 2006 election among the 1986 cohort is robust. The estimated value of $k^*$ is 0.87, which means that just-eligibles would have to register at a much lower rate than just-ineligibles to explain the result if the true effect on turnout-to-population rates was zero. Because theory indicates that eligible voters are more likely to register, we can place some confidence in this result. By contrast, the other estimated values of $k^*$ suggest that the negative estimates of the effect of eligibility on subsequent turnout-to-registration rates are highly vulnerable to differential registration bias. These estimates have corresponding $k^*$ values of 1.03–1.15, which means that only slight registration differentials in the expected direction (i.e., $r_T > r_C$) could produce the observed relationship. Correspondingly, if the registration differentials were larger then $k^*$, the true effect on the turnout-to-population rate would be positive.

$^{18}$Because we observe registration only in 2011, we cannot measure changes in registration rates over time. As a result, $k^*$ is constant within each birth cohort.
Table 5: Sensitivity analysis

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$T_{Reg} - T_{Creg}$</td>
<td>$T_{Reg} - T_{Creg}$</td>
<td>$T_{Reg} - T_{Creg}$</td>
</tr>
<tr>
<td></td>
<td>$k^*$</td>
<td>$k^*$</td>
<td>$k^*$</td>
</tr>
<tr>
<td>E2 (2006 midterm)</td>
<td>2.12</td>
<td>-2.16</td>
<td>-3.06</td>
</tr>
<tr>
<td>E3 (2008 presidential)</td>
<td>0.87</td>
<td>1.04</td>
<td>1.15</td>
</tr>
<tr>
<td>E4 (2010 midterm)</td>
<td>-2.56</td>
<td>1.14</td>
<td>-</td>
</tr>
</tbody>
</table>


Estimating turnout rates using an external data source

These findings indicate that the results above in which we condition on registration are highly sensitive to differential registration bias between the treatment and control groups (individuals who were just-eligible or just-ineligible to vote on election day). It is therefore advisable to adjust our turnout rate estimates using population totals from an alternative data source rather than relying on total registrants in the Catalist data. To do so, we calculate daily birth totals within the eight-day window around election day in the 1986 and 1988 cohorts for our sample of 42 states and the District of Columbia using data from Vital Statistics of the United States (1988; 1990). However, exact birth dates were redacted from these data starting in 1989. We therefore estimate daily birth totals for our sample states by scaling the total number of births in the U.S. for each birthdate in our window from the 1990 edition of Vital Statistics (1994) by the proportion of the population living there at the time. The proportion of the U.S. population living in the states in our sample was stable during the study period (varying only from 83.5%–84.1% from 1986–2008 according to Census data) so we did not further adjust these estimates to account for interstate migration.

Using these data, we calculate registration and turnout-to-population rates, which allow us to

---

19 As described above, we exclude Illinois due to legal restrictions on state voter file use plus Oklahoma, South Carolina, Colorado, Massachusetts, New Jersey, Vermont, and Washington, which had school entry birthdate cutoffs that potentially overlap with our date of birth window, creating a potential educational confound for our RD design.
Table 6: Registration rates as of 2011 by voting eligibility as proportion of births

<table>
<thead>
<tr>
<th>Year</th>
<th>Treated</th>
<th>Control</th>
<th>( \hat{k} )</th>
<th>E2</th>
<th>E3</th>
<th>E4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1986</td>
<td>28.42</td>
<td>25.99</td>
<td>1.09</td>
<td>0.87</td>
<td>1.04</td>
<td>1.14</td>
</tr>
<tr>
<td>1988</td>
<td>24.47</td>
<td>23.41</td>
<td>1.04</td>
<td>1.03</td>
<td>1.05</td>
<td>-</td>
</tr>
<tr>
<td>1990</td>
<td>22.59</td>
<td>14.48</td>
<td>1.56</td>
<td>1.15</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

diagnose differential registration bias. In Table 6, we divide the total number of registrants in the treatment and control groups by birth totals from public data sources, which produces estimates of \( r_T \) and \( r_C \). While these figures are not valid estimates of the true registration rate among people in these cohorts (our data are a random sample from Catalist’s voter file and do not include every voter born on the dates in question in our sample states), the difference in registration rates is a valid estimate of differential registration bias due to the use of random sampling. In each row of the table, we can see that the registration rate is much higher in the treatment than in the control group (all \( p < .01 \)). These registration differences are greatest for the 1990 cohort because the just-ineligible (control) group has had less time to “catch up” to the just-eligible (treatment) group, but persist even among the 1986 cohort despite them being measured seven years after turning 18.

Most notably, Table 6 shows that our estimates of the differential registration factor \( \hat{k} \) are well within the range that the sensitivity analysis in Table 5 suggests could explain our negative turnout-to-registration results. For the 1986 cohort, \( \hat{k} \) is 1.09 and the values of \( k^* \) that could explain the negative turnout-to-registration estimates in E3 and E4 are, respectively, 1.04 and 1.14. Likewise, \( \hat{k} \) is 1.04 for the 1988 cohort and the E2 and E3 values of \( k^* \) are, respectively, 1.03 and 1.05. Finally, the \( \hat{k} \) value of 1.56 for the 1990 cohort greatly exceeds the 1.15 estimate of \( k^* \) for E2.

To validate these findings, Table 7 reports a second RD analysis in which we again compare turnout rates between just-eligible and just-ineligible voters but use birth totals rather than registrants as the denominator in our calculations (corresponding RD plots of the raw data are provided in the appendix). Again, because our data is a sample from Catalist’s voter file and does not contain the full universe of people who voted on the birth dates in question, our point estimates of the turnout-to-population rate are not valid for either population. However, because the sample was drawn without respect to birthdate, we can correctly estimate the turnout-to-population difference.
Table 7: Turnout rates by voting eligibility as a proportion of births

<table>
<thead>
<tr>
<th></th>
<th>A. 1986 cohort (first election for just-eligibles: 2004 presidential)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Eligibility effect</td>
<td>0.96</td>
</tr>
<tr>
<td></td>
<td>(0.65, 1.26)</td>
</tr>
<tr>
<td>Control group</td>
<td>3.79</td>
</tr>
<tr>
<td></td>
<td>B. 1988 cohort (first election for just-eligibles: 2006 midterm)</td>
</tr>
<tr>
<td>Eligibility effect</td>
<td>0.22</td>
</tr>
<tr>
<td></td>
<td>(-0.28, 0.71)</td>
</tr>
<tr>
<td>Control group</td>
<td>12.99</td>
</tr>
<tr>
<td></td>
<td>C. 1990 cohort (first election for just-eligibles: 2008 presidential)</td>
</tr>
<tr>
<td>Eligibility effect</td>
<td>1.18</td>
</tr>
<tr>
<td></td>
<td>(0.90, 1.47)</td>
</tr>
<tr>
<td>Control group</td>
<td>3.34</td>
</tr>
</tbody>
</table>


When we adjust for birth totals in this way, we find results that are largely the opposite of what we found when we conditioned on registration. In some elections we consider, being eligible to vote at age 18 significantly increases downstream turnout-to-population rates (E2 and E3 for the 1986 cohort and E2 for the 1990 cohort), which is consistent with prior findings (Meredith 2009; Dinas 2012; Coppock and Green 2015). In the other elections, the estimated effect of eligibility is null. Most importantly, however, we find no evidence of a negative effect of eligibility on subsequent turnout-to-population rates, which confirms that the findings above were skewed by differential registration bias.

The reversal of the negative effects on turnout-to-registration rates is caused by the differential registration rates between the two groups compared. Figure 3 illustrates one way that such a reversal can be observed using data for the 1986 cohort. Even though the total registration and vote counts are similar between the just-eligible and just-ineligible groups, the total population—as estimated by birth counts—is higher in the control group, considerably reducing the turnout-to-population rates relative to the treatment group.
Study 2: Preregistration effects in Florida

Our second study is based on the recent work by Holbein and Hillygus (2015), who investigate the effects of preregistration on future turnout among young people. Preregistration laws typically allow voting-ineligible 16-year-old or 17-year-old citizens to complete a registration application, so that they are automatically added to the registration rolls once they turn eighteen and become eligible to vote. The authors present two different analyses, the first using data from the Current Population Survey, and the second based on the Florida voter file. In both cases, the authors find evidence that preregistration has a positive effect on young people’s subsequent turnout. We refer the reader to the original article for further details.

We focus exclusively on their second analysis, which compares voter turnout among narrowly

Note: Our data are a random sample from Catalist’s voter file and therefore underestimate the turnout-to-population rates for both the treatment and control groups. However, because the data were drawn randomly, we can accurately estimate the difference in turnout-to-population rates between groups. Voting is measured in the 2008 election, and registration is measured in 2011.
eligible and narrowly ineligible Florida voters who were born in 1990 close to the voting-eligibility cutoff for the 2008 presidential election. Holbein and Hillygus (2015) use this design to estimate the effects of preregistration. In Florida, where preregistration is allowed, narrowly ineligible voters are exposed to the opportunity to preregister, while most of those who are narrowly eligible to vote register “regularly” (i.e., when they are already eighteen). They conceptualize narrowly ineligible voters as the treatment group and narrowly eligible voters as the control group, where ineligibility is an instrument for preregistration—the treatment of interest. Their analysis is based on a fuzzy RD design where ineligibility induces preregistration.

Our re-analysis of Holbein and Hillygus’ Florida results, which uses the comprehensive replication materials they generously provided, differs from their original study in important ways. We are primarily interested in illustrating how differential registration patterns between treatment and control groups can affect turnout studies that calculate turnout rates as a proportion of registration. For this reason, we re-analyze Holbein and Hillygus (2015) as a sharp regression discontinuity design where, analogously to our Study 1, the treatment of interest is voting eligibility (as opposed to preregistration), narrowly eligible voters are the treatment group, and narrowly ineligible voters are the control group. Our design is thus analogous to the intent-to-treat (ITT) analysis that they report in the article except that the treatment and control group labels are inverted.

**Effects of eligibility on turnout-to-registration rates**

We first estimate the effect of voting eligibility on future turnout-to-registration rates and then conduct a sensitivity analysis to determine if the results could be driven by differential registration bias. For our analysis, we subset the data to people born from October 4–December 4, 1990 to match the Holbein and Hillygus (2015) window of approximately one month on either side of election day. Within this window, we treat the assignment of voting eligibility in 2008 as locally random and compare the turnout-to-registration rate in 2012 between just-eligibles and just-ineligibles.

Table 8 reports the results. Our estimated treatment effect on turnout-to-registration rates is -2.80 percentage points, meaning that just-ineligible registrants who were exposed to the option to preregister in 2008 voted at a higher rate in 2012 (51.88%) than registrants who in 2008 were just-eligible (49.08%). This estimate is very close to the three percentage-point effect that Holbein
Table 8: Sensitivity analysis of 1990 Florida data

<table>
<thead>
<tr>
<th>$T_{\text{Reg}}^{\text{T}}$ (just-eligible turnout)</th>
<th>$T_{\text{Reg}}^{\text{C}}$ (just-ineligible turnout)</th>
<th>$T_{\text{Reg}}^{\text{T}} - T_{\text{Reg}}^{\text{C}}$</th>
<th>$k^*$</th>
</tr>
</thead>
<tbody>
<tr>
<td>49.08</td>
<td>51.88</td>
<td>-2.80</td>
<td>1.06</td>
</tr>
</tbody>
</table>

Florida voter file data for 1990 births one month before or after November 4th; sample size = 30,979.

and Hillygus (2015) report for their ITT estimate.\(^{20}\)

**Sensitivity analysis: Assessing differential registration scenarios**

Our goal is to establish the robustness of the above finding to potential differential registration patterns. The sensitivity analysis gives a $k^*$ value of 1.06, which means that if the rate of registration were six percent higher in the treatment than in the control group, the negative turnout-to-registration effect we observe would reflect a turnout-to-population effect of 0 (and any difference greater than six percent would switch the effect from negative to positive). In 2008, the nationwide percentage of 18-year-olds who reported being registered to vote was approximately 49% (Herman and Forbes 2010). Assuming that the registration rate among just-eligible (treated) Florida voters within the window we consider was also 49%, a $k^*$ value of 1.06 implies that the rate of registration among just-ineligible (control) voters would only have to be 46.23% to change the sign of the point estimate—a difference of just 2.77 percentage points.

We therefore conclude that, as in Study 1, the negative eligibility effect found in Florida is not very robust to *positive* differential registration—i.e., to a situation where the registration rate is higher in the treatment than in the control group. However, we show below that when we approximate population totals for the two groups in Florida, there is no evidence of differential registration bias. In fact, the difference in registration rates seems to move in the opposite direction (i.e., registration rates are higher in the control group), suggesting that the Florida results reported in Holbein and Hillygus (2015) are indeed robust.\(^{21}\)

\(^{20}\)The difference is likely due to the fact that, unlike Holbein and Hillygus (2015), our analysis reports a simple difference in means and does not include controls.

\(^{21}\)The main focus of Holbein and Hillygus (2015) is not the ITT, but rather the TOT effect—the effect of voting ineligibility on people who preregister. This quantity is essentially the ITT effect divided by the difference in preregistration rates. A change in the sign of the ITT effect, however, would result in a change in sign in the TOT effect (the preregistration rate is, by construction, positive).
Estimating turnout rates using an external data source

We now use births in 1990 to approximate the total population of Florida residents who are just eligible or ineligible to vote in 2008. Ideally, we would adjust the results to birth counts, as we did above with the data from our 42-state sample. However, because of changes in privacy policy, the CDC’s National Vital Statistics do not contain birth rates disaggregated by state and exact date from 1990 onward. We therefore cannot precisely estimate the 2008 population of Floridians in the treatment and control groups.

Because we use a one-month window around the cutoff date of November 4, 1990, though, we can approximate the correct birth totals using the October totals for the treatment group and the November totals for the control group. Because our totals will be inaccurate for four dates in each group, these results should be considered exploratory and not definitive. Using this admittedly incorrect total for the population still gives us some insight into the robustness of the Holbein and Hillygus (2015) estimate, however. In Table 9, we compare the results using both registered voters and births as the denominator. Unlike our analysis of the Catalist data in Study 1, the effect remains negative when we calculate turnout based on births. Indeed, the point estimate is very similar to the effect based on registration: we estimate that 2012 turnout among Floridians who were narrowly eligible in 2008 was -2.67 percentage points lower than their just-ineligible counterparts.

This result can be explained by the fact that the number of births in Florida during November 1990 was significantly lower than the number of births in October. Except for November, the total monthly births in Florida between July and December 1990 were above 17,000. The total number of births in November 1990, however, was only 16,289. Thus, although the number of people who voted in 2012 was very similar in the treatment and the control groups (7,888 and 7,734, respectively), the turnout rate is higher as a proportion of the control group’s (approximate) size. The total number of people registered is thus apparently lower in the control group (which results in the negative effect reported in Table 8) not because of a lower rate of registration but rather because the total control population is smaller than the total treated population. In sum, using births to approximate the actual size of the treated and control population, we find no evidence that the results in Holbein and Hillygus (2015) are caused by differential registration patterns.
Table 9: RD estimates of eligibility effect in 1990 Florida cohort

<table>
<thead>
<tr>
<th>Denominator</th>
<th>$T^\text{Reg}_T$ (just-eligible turnout)</th>
<th>$T^\text{Reg}_C$ (just-ineligible turnout)</th>
<th>Effect size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Registration</td>
<td>49.08</td>
<td>51.88</td>
<td>-2.80</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-3.92, -1.68)</td>
</tr>
<tr>
<td>Births</td>
<td>44.81</td>
<td>47.48</td>
<td>-2.67</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(-3.74, -1.60)</td>
</tr>
</tbody>
</table>

Florida voter file data for 1990 births one month before or after November 4th; sample size = 30,979. Estimated births calculated from Florida vital records for October and November 1990. 95% confidence intervals in parentheses.

Conclusion

Turnout studies using observational voter file data are more common than ever in American politics. However, these data implicitly condition on voter registration, risking what we call differential registration bias, which can distort estimates of how a treatment affects turnout rates. We formalize a definition of this problem and develop a new approach to sensitivity analysis that allows scholars to assess how robust their estimates are to potential differences in registration rates between treatment and control groups. We then show why obtaining population total estimates from alternative data sources is a superior approach when possible. The use of these methods is illustrated with two studies of turnout based on voter file data — an original analysis of voter file data from 42 U.S. states and a reanalysis of Holbein and Hillygus (2015) using data from Florida voter files.

In both cases, we show that comparisons of turnout-to-registration rates are highly sensitive to differential registration bias and would reverse if the registration rate were slightly higher in the treatment group than the control group. However, when we use alternative population estimates to approximate the total population in each group, we find important differences between the two studies. In Study 1, we find that the true effect of voting eligibility on turnout-to-population rates is actually positive, which reverses the sign of the estimated turnout-to-registration effect. In contrast, the alternative population measure we use in Study 2 indicates that the control group’s registration rate is higher than the treatment group, indicating that the negative effect in turnout-to-registration rates reflects a negative effect on true turnout rates rather than differential registration bias.

Substantively, our findings using the Catalist data provide new evidence about the effect of early political experiences on subsequent political participation. The results in Study 1 provide
the most rigorous evidence to date to support the claim that election eligibility can increase subsequent voter turnout. Turnout rates are often significantly higher among just-eligible voters than just-ineligibles in a narrow window of birth dates around the election eligibility cutoff when we condition on total births by date instead of registrations. We also show that the original results in Holbein and Hillygus (2015) based on turnout-to-registration rates in Florida are robust to differential registration bias, enhancing the credibility of original study’s findings about the positive effects of preregistration on turnout.

Of course, this study has limitations that should be acknowledged. While our sensitivity test can provide the exact registration differential that would explain the observed turnout difference between treatment and control groups in voter file data, alternative data sources that can be used to approximate the population totals may provide inexact estimates. In the applications we consider, for instance, it is necessary to approximate the relevant population totals at the state level. For this approach to be valid, it is necessary to assume that interstate migration rates do not differ discontinuously around election day (we know of no evidence or theory suggesting such a possibility). More generally, the turnout effect estimates we report necessarily depend on the assumptions of our RD design—notably, that individuals are as-if randomly assigned to voting eligibility in the narrow window we use around the eligibility cutoff. Future research should continue to assess the validity of our findings on the effects of eligibility and preregistration on turnout and the possibility of differential registration by eligibility status.

Nonetheless, this study illustrates that rigorous study designs like regression discontinuity and even experiments can be vulnerable to post-treatment bias despite other study assumptions being met (e.g., balance on observables). If we condition on or control for the treatment of interest in any way, our estimates of causal effects can be severely biased. This possibility is especially insidious in cases of endogenous selection bias, where we implicitly condition on the treatment in defining the sample. Differential registration bias is only one way in which failing to take this threat into account can lead us to mistaken inferences about how politics works.
**References**


Appendix

Balance checks

To demonstrate that the as-if randomized assumption of our RD design is plausible among the registrants in our data, we compare the demographic characteristics of just-eligible and just-ineligible voters in Table A1 using covariates in the Catalist data, which combines public and commercial records of gender, marital status, race/ethnicity, and religious affiliation. In the pooled data, the differences in means are small and generally not significant despite the very large sample size, suggesting the as-if randomization assumption is likely to be satisfied among registrants.\(^1\)

Table A1: Balance statistics in Catalist data

<table>
<thead>
<tr>
<th></th>
<th>All</th>
<th>Treatment</th>
<th>Control</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.464</td>
<td>0.457</td>
<td>0.11</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.095</td>
<td>0.095</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.160</td>
<td>0.154</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.634</td>
<td>0.644</td>
<td>0.12</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.156</td>
<td>0.159</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Catholic</td>
<td>0.265</td>
<td>0.272</td>
<td>0.06</td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>0.275</td>
<td>0.276</td>
<td>0.93</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1986</th>
<th>Treatment</th>
<th>Control</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.461</td>
<td>0.454</td>
<td>0.44</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.126</td>
<td>0.122</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.156</td>
<td>0.156</td>
<td>0.92</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.643</td>
<td>0.645</td>
<td>0.90</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.156</td>
<td>0.156</td>
<td>0.89</td>
<td></td>
</tr>
<tr>
<td>Catholic</td>
<td>0.262</td>
<td>0.268</td>
<td>0.35</td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>0.283</td>
<td>0.277</td>
<td>0.38</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1988</th>
<th>Treatment</th>
<th>Control</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.472</td>
<td>0.457</td>
<td>0.04</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.085</td>
<td>0.089</td>
<td>0.33</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.161</td>
<td>0.156</td>
<td>0.39</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.637</td>
<td>0.645</td>
<td>0.28</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.153</td>
<td>0.153</td>
<td>0.96</td>
<td></td>
</tr>
<tr>
<td>Catholic</td>
<td>0.265</td>
<td>0.268</td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>0.274</td>
<td>0.68</td>
<td>0.40</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>1990</th>
<th>Treatment</th>
<th>Control</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>0.461</td>
<td>0.462</td>
<td>0.91</td>
<td></td>
</tr>
<tr>
<td>Married</td>
<td>0.070</td>
<td>0.061</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Black</td>
<td>0.164</td>
<td>0.148</td>
<td>0.01</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>0.632</td>
<td>0.644</td>
<td>0.16</td>
<td></td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.159</td>
<td>0.174</td>
<td>0.03</td>
<td></td>
</tr>
<tr>
<td>Catholic</td>
<td>0.267</td>
<td>0.286</td>
<td>0.02</td>
<td></td>
</tr>
<tr>
<td>Protestant</td>
<td>0.269</td>
<td>0.286</td>
<td>0.03</td>
<td></td>
</tr>
</tbody>
</table>

2011 Catalist data; \(n = 49,493\) (1986: 18,472; 1988: 17,185; 1990: 13,836)

\(^1\)There are a few imbalances in the 1990 cohort, which may be the result of the shorter interval between the treatment election for this cohort (2008) and the year of data collection (2011).
RD plots: Registrant turnout rates

As is conventional in RD analyses, we plot raw turnout rates among registrants binned by date of birth. Figure A1 illustrates how turnout varies by eligibility in the election after treatment (which we call E2), while Figure A2 presents corresponding results for the next two elections (E3 and E4).

Figure A1: RD effects of voting eligibility on turnout in subsequent election

![Graph showing RD effects of voting eligibility on turnout in 2006 for 1986 cohort (E2)](a)

![Graph showing RD effects of voting eligibility on turnout in 2008 for 1988 cohort (E2)](b)

![Graph showing RD effects of voting eligibility on turnout in 2010 for 1990 cohort (E2)](c)

2011 Catalist data; \( n = 49,493 \) (1986: 18,326; 1988: 17,153; 1990: 13,792). 95% confidence intervals in parentheses. Lines represent means and 95% confidence intervals for just-eligibles and just-ineligibles.
Figure A2: RD effects of voting eligibility on turnout in second and third subsequent elections

(a) Turnout effects in 2008 for 1986 cohort (E3)

(b) Turnout effects in 2010 for 1988 cohort (E3)

(c) Turnout effects in 2010 for 1986 cohort (E4)

RD plots: Population turnout rates

Figures A3 and A4 plots the raw data for turnout rates by date of birth when adjusted by population totals rather than the number of registrants in the data.

Figure A3: RD estimates of voting eligibility effects on population turnout rates (E2)

(a) Turnout effects in 2006 for 1986 cohort (E2)

(b) Turnout effects in 2008 for 1988 cohort (E2)

(c) Turnout effects in 2010 for 1990 cohort (E2)

2011 Catalist data; \( n = 49,493 \) (1986: 18,326; 1988: 17,153; 1990: 13,792). 95% confidence intervals in parentheses. Lines represent means and 95% confidence intervals for just-eligibles and just-ineligibles.
Figure A4: RD estimates of voting eligibility effects on population turnout rates (E3–E4)

(a) Turnout effects in 2008 for 1986 cohort (E3)

(b) Turnout effects in 2010 for 1988 cohort (E3)

(c) Turnout effects in 2010 for 1986 cohort (E4)

2011 Catalist data; \( n = 49,493 \) (1986: 18,326; 1988: 17,153; 1990: 13,792). 95% confidence intervals in parentheses. Lines represent means and 95% confidence intervals for just-eligibles and just-ineligibles.