Voting Among Siblings

By MICHAEL D. BLOEM, JOHN B. HOLBEIN, SAMUEL J. IMLAY, AND JONATHAN SMITH*

Draft: June 12, 2025

Using millions of siblings in the U.S., we detail three findings that quantify whether siblings influence one another to vote in national elections. First, and descriptively, younger siblings are 10 percentage points (50 percent) more likely to vote in their first eligible election when their older sibling votes in a prior election. Second, roughly one-third of this is caused by the older sibling voting, as determined by age-of-voting-eligibility thresholds in a regression discontinuity design. Third, the causal impact of a sibling voting runs in the other direction as well—younger siblings increase the probability of their older siblings voting in their early 20's by 14 percent. These results demonstrate the influence and importance of family and peers in creating an engaged citizenry and underscore that across a wide array of policy domains, conventional impact evaluations do not fully account for all of policies' impacts.

Keywords: Voting; Civic Engagement; Political Socialization; Family; Siblings; Spillovers; Peer Effects

I. Introduction

In the United States, only two-thirds of the voting-age population cast a ballot in 2020—an election with record turnout. Among 18-24-year-olds, only 51.4 percent voted, which was, again, a four-decade high.[1] Why do some young people become active in politics, while others fail to do so? There is no shortage of research and explanations as to why people do (or do not) develop into active voters (Holbein and Hillygus, 2020). Scholars across several disciplines have long looked to the family as a potential source of voting behaviors. Much of the family-focused literature examines the parental attributes that lead children to become voters (e.g., Levin, 1961; Clarke, 1973; Jennings and Niemi, 1968; Tedin, 1974; Jennings and Niemi, 2014; Sears and Funk, 1999; Plutzer, 2002; Jennings, Stoker and Bowers, 2009; Torney-Purta, 2017; Easton and Hess, 1962; Niemi and Sobieszek, 1977; Lane, 1959; Achen, 2002; Sapiro, 2004). For example, parents' level of education and marital status are strongly correlated with children's voting behavior (e.g., Weisberg, 1987; Verba, Schlozman and Burns, 2005).

In this paper, we examine whether an individual's decision to vote affects their siblings' decisions to vote. This departs from the earlier literature on family and voting in two distinct ways. First, we focus on the role of siblings

* Bloem: College Board. Holbein: University of Virginia. Imlay: College Board. Smith: Georgia State University and IZA. This paper represents the views of the authors and not their organizations or that of the data providers.

[1] According to Census data compiled by the US Elections Project, since the 1980s youth voter turnout in Presidential Elections has typically hovered around 30-40%. In Midterm Elections, youth voter turnout is lower: usually around 20%.

on one another, not parents on children. In other contexts, siblings have been described as the "ultimate peer" (Black et al., 2021) that give us a window into how close companions can influence one another. However, this possibility has not been examined in political behaviors. Second, we focus on individuals' civic actions on their siblings' voting behavior instead of a family or individual's attributes or circumstances. Focusing on one sibling's act of voting provides a direct and clear connection to the outcome of interest—voting.

Siblings share more in common than most, which creates a challenge in estimating whether siblings impact each other's probability of voting. For example, siblings often attend the same schools, have the same neighbors, watch the same television shows, and are exposed to similar civic and political acts and discussions. Each of these shared characteristics could influence an individual's likelihood of voting, suggesting that observed similarities in voting patterns might arise from these common factors rather than sibling relationships *per se*. These similarities in voting patterns among siblings are borne out in our data.

Using data on 3.9 million sibling pairs from the College Board's educational assessment data linked with administrative voting records, we show that younger siblings are 10 percentage points (50 percent) more likely to vote in their first eligible election when their older sibling votes in a prior election. This relationship survives a host of observable controls, including high school fixed effects. We are the first to demonstrate this exceptionally strong association between siblings' voting decisions. It is also achieved with the accuracy of administrative voting records, which can be an improvement over survey results biased from social desirability (Karp and Brockington, 2005; Holbrook and Krosnick, 2010) and for a much larger sample than ever before.[2]

To move beyond descriptive associations, we use a regression discontinuity design that exploits age-of-voting-eligibility cutoffs. In the U.S., individuals must be 18 years old on or before election day to vote. For example, someone born on November 4th, 1990 was exactly 18 years old on election day in 2008 and therefore eligible to vote. Conversely, someone born one day later—November 5th, 1990—was one day shy of being eligible to vote in the 2008 election. On average, the people born on the 4th versus the 5th should be the same on all other dimensions, thereby setting up our natural experiment. We then compare the voting decisions of the younger (and older) siblings of the people born on the 4th versus the 5th, for example, and test whether they vote at different rates from one another.

We first quantify the above thought experiment that serves as our first stage—exogenous variation in voting using age-of-voting eligibility birth dates. About 20 percent of people, including older and younger siblings, vote just after turning 18 in an election year (30 percent in Presidential Elections). Voting around age 18 then translates into an increased likelihood of voting around age 20 and 22 compared to those who did not vote around age 18.

[2]Our administrative education records includes both registered and unregistered individuals, which is another advantage over prior studies. Studies using only administrative voting records have the potential for differential registration bias (Nyhan, Skovron and Titiunik, 2017). In the past, most scholars have gotten around this potential form of post-treatment bias by leveraging supplemental survey data (Holbein and Rangel, 2020; Holbein et al., 2023).

We are not the first to ask the question of whether voting is persistent or habit-forming; indeed, many economists and political scientists have deemed this a vitally important question to explore in better understanding the roots of political participation (e.g., Coppock and Green, 2016; Fujiwara, Meng and Vogl, 2016; Meredith et al., 2009; Plutzer, 2002; Gerber, Green and Shachar, 2003; Green and Shachar, 2000). That said, our analysis improves upon this prior work. Most similar to our examination, Coppock and Green (2016) use discontinuities in voting by birth date relative to the age of eligibility. However, they only have access to administrative voting data, not a base sample that includes non-registrants, so they rely on density jumps at the aggregate birth date level. We expand on their work with our education dataset that serves as a base dataset with non-registrants, along with data from all states, not just the 17 they use. Our broader analysis confirms that voting is, indeed, persistent or habitual.

Next, we test for spillovers from whether an older sibling votes to whether the younger sibling votes. This is only one path in which older siblings may impact younger siblings' propensity to vote and the one we focus on. Spillovers may flow in this direction if the older sibling has access to a car, serves as a role model, or is a source of information for the younger sibling. These are just a few examples, and we discuss the related literatures and reasons we may see spillovers in the next section.

We find that younger siblings are 3.3 percentage points (14.4 percent) more likely to vote in their first eligible election if their older sibling voted around age 18. This effect is substantively meaningful compared to the impacts of commonly tested and meta-analyzed get-out-the-vote interventions.[3] For example, Green, McGrath and Aronow (2013) show in a meta-analysis that a canvassing contact from a campaign increases voter participation by approximately 2.5 percentage points, a phone conversation increases voter participation by approximately 1.9 percentage points, and mailers increase voter participation by 0.2 percentage points. By another standard, our estimated effect is also roughly one-third of the association previously mentioned. It is primarily driven by older siblings who voted around age 18 during a Presidential Election and for siblings of the same sex. We also find evidence that this effect lasts beyond the younger sibling's first eligible election but not the probability of ever voting.

We also find that younger siblings who vote around age 18 increase the probability of older siblings voting. Older siblings are 4 percentage points (13.7 percent) more likely to vote in the same election when their younger sibling votes around age 18 than older siblings whose younger sibling did not vote. Again, this effect is substantively meaningful. To our knowledge, ours is the only paper to examine the political effects of younger siblings on older siblings and one of the few papers to demonstrate spillovers from younger to older siblings in any context (e.g., Black et al., 2021; Karbownik and Özek, 2023; Humlum, Morthorst and Thingholm, 2024).

[3]We acknowledge that get-out-the-vote interventions are not the same as our treatment; for one, their dosage is likely much lighter. However, get-out-the-vote interventions are the most widely studied approach to increasing voter turnout and the approach most frequently evaluated using methods for causal inference.

Next, we explore several potential mechanisms behind our results. We are unable to find evidence of a single mechanism driving the results but, rather, we find evidence against some mechanisms. For example, our bi-directional spillovers rule out the "role model" explanation as being the sole mechanism—given the plausible assumption that younger siblings are unlikely to serve as role models to older siblings, at least in the traditional sense. Related, heterogeneous effects analyses do not consistently find that "close" siblings are more likely to influence one another. For example, we find that same-sex siblings are driving the results from older to younger siblings, but opposite-sex siblings are driving the results from younger to older siblings. Similarly, the effect sizes are similar for siblings closer in age to one another than those further apart. Finally, siblings in the same state as one another at the time of an election are, if anything, less likely to impact one another than those in different states. This largely rules out a transportation costs mechanism. We do find evidence that the effects are relatively larger for siblings from families whose parents have a bachelor's degree and those with above median PSAT scores, perhaps suggesting that information and discussions are at play.

Our work provides a few important contributions worth highlighting. First, our data are uniquely large in size and scope, making use of millions of siblings across the United States. Second, our identification strategy that exploits voting age eligibility cutoffs makes us the first paper we are aware of that estimates the causal impact of siblings' voting habits on one another. Third, we do not know of any other research that explores or finds bi-directional sibling spillovers, which has important implications for families, policies, and interventions in many contexts. Related, this implies that standard evaluations may miss some of the important benefits (or costs) of these interventions and policies. This general concept of social spillovers being a potentially important component of evaluations is often discussed in other contexts (e.g., Sacerdote, 2011; Hendren and Sprung-Keyser, 2020), but we do so in a political context and with siblings. Finally, Americans' voting rates are far too low, and our results demonstrate the importance of close peers in voting, civic engagement, and beyond.

We next provide a discussion of the existing literature, our relative contribution, and some of the motivation behind our analyses, all in Section II. Then we discuss our data in Section III. Section IV showcases the descriptive relationship between siblings' voting habits. Section V shows how voting is habit forming in that voting around age 18 leads to voting in future elections. This also sets the stage for our regression discontinuity design. Section VI discusses the methods for detecting sibling spillovers before moving to spillovers from older to younger siblings and from younger to older siblings, in Section VII and VIII, respectively. We conclude with Section IX.

II. Prior Work, Background, and Contributions

Our work contributes to several foundational literatures in economics, political science, sociology, and psychology. How young people develop their political behaviors and leanings is a core question across these disciplines.

We first discuss existing literature on families, typically parents, and voting and the absence of empirical work on siblings in this domain. We then discuss how and why siblings might influence each other to vote.

A. *Prior Research on Voting, Family, and Sibling Spillovers*

Parents and Children in Politics and Voting

Within families, scholarly attention has long centered on parental influences, and for good reason: a substantial body of evidence demonstrates that parents' characteristics, beliefs, and behaviors play a fundamental role in shaping their children's political attitudes and behaviors (e.g., Levin, 1961; Clarke, 1973; Sandell and Plutzer, 2005; Jennings and Niemi, 1968; Tedin, 1974; Jennings and Niemi, 2014; Sears and Funk, 1999; Plutzer, 2002; Jennings, Stoker and Bowers, 2009; Torney-Purta, 2017; Easton and Hess, 1962; Niemi and Sobieszek, 1977; Lane, 1959; Achen, 2002; Sapiro, 2004).[4] Indeed, parental influence is crucial to understanding why certain young people become politically engaged while others do not.

Various parental attributes strongly correlate with children's political behavior. For instance, parents' educational attainment strongly shapes political engagement; higher parental education typically fosters an environment that values civic engagement and encourages political participation (Verba, Schlozman and Brady, 1995). Similarly, parents' own political activity is thought to model participatory behavior, directly teaching children about political engagement and indirectly signaling its importance (Jennings and Niemi, 1968). Parental marital status also matters, as stable family structures are positively associated with higher political knowledge and greater feelings of political efficacy among children (Gimpel, 2003). Additionally, these parental attributes frequently overlap with economic factors, particularly parental income. Parents with higher socioeconomic status have greater resources to invest in their children's political and civic education, providing opportunities and networks that facilitate political participation (Akee et al., 2020).

The importance of these familial inputs motivates several aspects of our methodological approach. Specifically, the strong role of family background in shaping voting behaviors informs our inclusion of certain covariates in our condition-on-observables models. Additionally, recognizing how these factors might differentially affect various subgroups, we estimate heterogeneous treatment effects to explore whether sibling influences vary systematically across different educational inputs and family dynamics, potentially highlighting where sibling effects may be strongest or weakest.

[4]There is also a smaller but expanding literature suggesting the reverse effect—children influencing parents' political attitudes and behaviors, referred to as "trickle-up political socialization" (Dahlgaard, 2018; Jennings and Niemi, 2014; Simon and Merrill, 1998; McDevitt and Chaffee, 2000, 2002; Linimon and Joslyn, 2002; Jennings, Stoker and Bowers, 2009; Washington, 2008; Green et al., 2023; Glynn and Sen, 2015; Urbatsch, 2014).

Summarizing the literature on the broader effects of siblings, Urbatsch (2014, 46) states that "the most frequent finding is that siblings tend to conform. In the process of socialization, the younger sibling's tastes and ideas tend to converge toward those of the older." Put simply, "throughout their youth, people tend to look up to and emulate their older siblings... siblings are a potentially important source of political socialization" (Urbatsch, 2011, 693).

Older siblings also often serve as role models for certain behaviors, with siblings significantly impacting the actions of their fellow familial peers. This influence can be positive, such as promoting prosocial behavior, or negative, including the modeling of risky behaviors. "Siblings affect behaviors—such as childbearing—that themselves influence political views. When siblings have such a range of effects, it seems likely that they might also affect outcomes of direct political interest" (Urbatsch, 2011, 693-694). The dynamics of these relationships are complex and can vary based on individual and family factors (Whiteman and Christiansen, 2008). Along these lines, Urbatsch (2014, 47) notes that "siblings affect participation in organizations and extracurricular activities; when these involve groups with particular outlooks toward social relations, this may shape subsequent political beliefs."

The economics literature has recently demonstrated the importance of role models, often in educational settings, but has rarely focused in-depth on bidirectional examinations of siblings (e.g., Bettinger and Long, 2005; Porter and Serra, 2020). The economics literature has shown the influence of older siblings in several contexts, while not being able to pin down the role model mechanism. For example, using data from Chile, Croatia, Sweden, and the United States, Altmejd et al. (2021) study within-family spillovers in college and major choice.[5] They show that "in all four countries a meaningful portion of younger siblings follow their older sibling to the same college or college-major combination" (Altmejd et al., 2021, 183). Qureshi (2018) shows that additional schooling for Pakistani girls increases their younger brothers' attainment. Dustan (2018) shows that students prefer schools their older siblings have previously attended. Goodman et al. (2015) show that in the United States one-fifth of younger siblings enroll in the same college as their older siblings. Bansak, Jiang and Yang (2022) show using data from China that when a younger sibling is enrolled simultaneously in school, an elder child's school performance improves significantly. Nicoletti and Rabe (2019) show that in England there is a significant positive spillover effect from the older to the younger sibling. In sociology, Zang, Tan and Cook (2023, 1529) use school entry birthdate cutoffs in North Carolina to show that "individuals whose older siblings were born shortly after the school entry cutoff date have significantly higher test scores in middle school and that this positive spillover effect is particularly strong in disadvantaged families." In short, this literature demonstrates that there exist substantial

[5] See also Joensen and Nielsen (2018).

spillovers from older to younger siblings, albeit outside a political context.

What's Missing from the Research—Examinations of Sibling Political Spillovers (in any direction)

Generally speaking, the just-discussed literature is missing two pieces of information that we address in this paper. First, although there is research on parents influencing children's political attitudes and behavior, there is a lack of evidence on siblings impacting one another when it comes to politics. Second, there is little evidence on younger siblings impacting older siblings in any context. We believe the dearth of evidence is partially due to data constraints and issues with causal identification, but there is also, simultaneously, a lack of conceptual underpinnings motivating this bidirectional examination.

The expectation of bidirectional sibling influence is theoretically grounded in the reciprocal nature of sibling relationships. Unlike parent-child interactions, sibling relationships are more peer-like and characterized by ongoing, mutual influence. Younger siblings often look to their older siblings as role models, adopting their attitudes and behaviors through observational learning and imitation. Conversely, older siblings may also be influenced by younger siblings through direct interactions, mutual socialization processes, and shared experiences within the family context. For example, younger siblings may introduce new social and political perspectives into the household, thereby influencing older siblings' views and behaviors. Additionally, shared experiences, such as discussing current events or participating together in community activities, may reinforce mutual influences between siblings.

In the social sciences, there is at least a general sense that siblings influence each other when it comes to political development. However, this idea is not fully developed and evidence testing these (somewhat scattered) theories is very sparse. In economic contexts, Black et al. (2021, 101) note that "there is limited evidence regarding the causal role of siblings on children's outcomes." In our context, siblings are rarely—if ever—mentioned in foundational treatises about political development. Indeed, "siblings have received relatively little attention as a source of [political development], despite their conspicuous place in the lives and perceptions of young people" (Urbatsch, 2011, 693). Indeed, "this field of research has, however, overlooked that children within the same family have different experiences" (Bratsberg et al., 2022, 475). As Urbatsch (2014) puts it: "siblings play only a small part in mainstream economics and political science. At best, these fields have acknowledged related influences of birth order and family size: older and younger siblings face divergent family dynamics growing up." Feinberg, Solmeyer and McHale (2012, 43) put it succinctly: "[siblings] are often ignored in research and preventive interventions with youth and families." Urbatsch (2014, 43) further notes that "[siblings] rarely take starring roles in political analyses, perhaps because any independent effects of individuals' siblings can be difficult to observe."[6] Urbatsch

[6] Alwin (2016) observes that "[Urbatsch (2014) shows that] having an older sister is associated with more left-wing views, but he does not seek information on the views of the older sister."

(2014, 43) notes that the roots of this pattern likely arise because "very few studies collect data on both sibling composition and interesting political beliefs" and that social scientists' "reticence concerning siblings may stem from the empirical complexity of the topic" (Urbatsch, 2011, 694).

Datasets containing detailed sibling information alongside political outcomes are rare, and most existing studies rely on conditional-on-observables designs that struggle to clearly identify causal relationships. For instance, existing research tends to examine siblings' influences on political attitudes, such as ideology, rather than behaviors like voting. Ifzaren (2018) provides one of the few examples, showing that Australian women with sisters differ slightly in their party preferences compared to women with brothers. However, such work remains limited in scope and typically correlational rather than causal.

Empirical studies examining such sibling dynamics remain scarce, primarily due to data limitations. One exception to this pattern is Lahtinen, Erola and Wass (2019). Using a register-based dataset of one election in Finland, Lahtinen et al. (2019) show that a quarter of the total variance in voter turnout is shared between siblings. That said, as the authors readily note—"most of the resemblance between siblings was left unexplained" as they only observed parental characteristics in their dataset. Moreover, random (or as-good-as random) assignment of sibling characteristics is incredibly rare, as many features are held constant within sibling clusters. What does exist in this space largely uses conditional-on-observables designs and most of the (limited) work in this space looks at siblings' effects on political beliefs (like political ideology), rather than on their political behaviors (like voting). For example, Ifzaren (2018) shows that women with sisters are more likely to vote for the National or Democrat parties (in Australia) and somewhat less likely to vote for the Labor party than women with brothers.

Our research significantly advances this understudied area. First, our dataset includes millions of sibling pairs from across the U.S., providing unprecedented scale and representativeness. This large and detailed dataset allows us to examine mechanisms and heterogeneous effects rigorously. Second, our data contain precise birth dates, enabling a robust causal identification strategy using age-of-voting-eligibility thresholds. This design moves beyond correlational analyses prevalent in existing literature, allowing us to establish clear causal effects of sibling voting behavior. By addressing these gaps, our study contributes to a deeper theoretical understanding of familial influences in political behavior, especially the under-examined bidirectional dynamics between siblings.

B. How Siblings May Impact One Another

Political Socialization

Political socialization—or the process by which children develop their political beliefs, values, attitudes, and behaviors—has garnered widespread attention (e.g., Holbein, 2017; Chyn and Haggag, 2023; Cohodes and Feigenbaum, 2021; Arnzen and Cohodes, 2025).[7] Broadly speaking, scholars across these disciplines have long debated just how much of a role families play in shaping children's political attitudes and behaviors (Sandell and Plutzer, 2005). Early debates in this space focused on the extent to which families or family structures were central (as argued by scholars like Levin, 1961) or secondary to institutions such as schools and the mass media, as argued by scholars (e.g. Clarke, 1973).

Urbatsch (2014, 44) notes that "despite [the] hurdles [to studying sibling effects], examining the role of sisters and brothers may repay the effort." Indeed, there are multiple reasons to suspect that siblings influence each other. Urbatsch (2014, 44) argues that "when siblings have such a range of effects, it seems likely that they might also affect outcomes of direct political interest" thus, "the lack of emphasis on siblings...sits uneasily with evidence from the other social sciences that siblings extensively affect each other's perceptions and behavior." As Urbatsch (2011, 711) puts it, "siblings, especially older siblings, serve as important socializing influences for many people. Most often, siblings provide influences similar on average to those of other family members and peers, since they themselves experience many of the same conditions and frequently travel in the same social networks." Put differently, "siblings are more than mere childhood appendages; they are usually an important presence in people's lives into adulthood, for better or worse" (Urbatsch, 2014, 121). "Brothers and sisters share a bond that is unlike any other relationship" (Black et al., 2021, 101). In short, "in childhood and adolescence, siblings spend considerable time together, and siblings' characteristics and sibling dynamics substantially influence developmental trajectories" (Feinberg, Solmeyer and McHale, 2012, 43).

Economic Reasons

There are several basic economic concepts and literature that can also lead to siblings influencing each other when voting. One example is travel and time costs. Voting often requires going to an official voting center to cast a vote. Travel distance and time are consistently a factor in economic models and research that influence decision-making, and they have even been shown to influence whether people vote (Chen et al., 2022; Dyck and Gimpel, 2005; Schafer and Holbein, 2020; Pettigrew, 2021; Cantoni, 2020). In this context, siblings may be able to share transportation costs or drive one another to the voting center. This is likely only a factor if the siblings live together or close by. This is why, in our analyses, we explore whether there are differential effects for siblings who live in different states.

Economics also has a vast literature on information acquisition across a variety of contexts. The general idea is that acquiring information, such as who to vote for or perhaps how and where to vote, can be time-consuming

[7] Much of this political socialization research is, admittedly, theoretical in nature and somewhat dated (e.g., Dawson and Prewitt, 1969; Searing, Schwartz and Lind, 1973; Greenstein, 1965). Even with some lamenting the "abandonment" of political socialization research in recent decades (Sapiro, 2004, 1) and arguing for a needed "rebirth" of political socialization research (Niemi and Hepburn, 1995, 7), the fact remains: many vitally important political socialization questions remain unanswered.

and costly (Corvalan and Cox, 2018; Blais et al., 2019; Holbein, 2016). People may want to learn about the issues and candidates before they decide to go vote, but doing so is costly (e.g., Martinelli, 2006; Chakravarty, Kaplan and Myles, 2018). In short, siblings may help each other acquire the necessary information to vote. This may be as simple as pointing a sibling to a website to register to vote or the location of polling stations. Or it could be conversations over the phone or dinner table about the issues and candidates of the day. These information spillovers may reinforce voting habits within families.

Finally, the costs of acquiring information are not necessarily evenly distributed, so some populations may find it harder to get the necessary information needed. This is embedded in many of the theoretical economic models. This also motivates some of our heterogeneous analyses.

Psychological Predispositions

One theoretical avenue has to do with psychological predispositions. As (Urbatsch, 2014, 44) notes "having siblings can alter basic, typically lifelong traits such as conscientiousness or extraversion" and work in political science has shown that the Big Five personality traits help shape whether people become active voters or fail to do so (e.g., Weinschenk, 2017; Gerber et al., 2011; Mondak, 2010). Siblings also tend to shape attitudes toward risk (Urbatsch, 2014, 47) and work in political science has linked this attribute with outcomes like voting (Hill, 2020; Schafer, 2021). Siblings also play a crucial role in shaping social skills and perspective-taking abilities. Research indicates that children with siblings demonstrate enhanced social competencies compared to only children. Factors such as age, birth order, and gender can moderate the influence of sibling relationships on social development (Sang and Nelson, 2017). These social skills have been linked to political behaviors such as voting (Holbein, 2017; Holbein et al., 2020).

Social stigma or prosocial behavior may be a factor. As mentioned, siblings are the ultimate peer. Since peers engage with one another, siblings may vote "because others will ask," as documented in DellaVigna et al. (2016). This behavioral and psychological concept of social-image concerns is theorized and affirmed in economics (e.g., Ali and Lin, 2013) and political science (e.g., Gerber, Green and Larimer, 2008; Harbaugh, 1996). Older siblings may value the warm glow of their younger siblings knowing they voted.

Home Environment

An important literature examines the importance of siblings' birth order for various life outcomes (e.g., Black, Devereux and Salvanes, 2005; Jayachandran and Pande, 2017; Black, Grönqvist and Öckert, 2018; Black, 2017; Ejrnæs and Pörtner, 2004; De Haan, 2010). Most pertinent to the present study, Bratsberg et al. (2022) use population-wide individual level register data from Sweden and Norway to show that higher birth order entails lower voter turnout. Generally speaking, research on birth order cannot distinguish whether any established patterns are because of parents' differential treatment of children or because of siblings' influence on one another. But the latter of the two explanations may in part determine our results.

In another line of inquiry, researchers have found that a child's disability status can influence their siblings' outcomes. Using administrative data on children from both Florida and Denmark, Black et al. (2021) show that having a disabled younger sibling affects the other siblings' outcomes. While this example is about younger siblings influencing older siblings, we have no reason to believe the reverse may not have impacts. And again, any estimates may operate through parents' differential treatment of children or because the siblings influence one another.

III. Data, Matching, and Samples

This study primarily uses two datasets: (1) sibling data derived from the universe of PSAT, SAT, and Advanced Placement (AP) takers in the high school graduation cohorts of 2004-2012; and (2) national voting records. We describe each in turn.

A. Sibling Data

The base data for this study are individual-level administrative records for all PSAT, SAT, and AP takers from the 2004-2012 high school graduating cohorts (21.3 million students). The SAT is one of two college entrance exams considered in admissions and program placement by thousands of colleges across the U.S. Approximately 1.3 million students per cohort take the exam. The PSAT is an exam taken prior to the SAT. It qualifies students for scholarships and college outreach and is the qualifying exam for the National Merit Scholarship. It has broader reach than the SAT and is usually taken during the sophomore and/or junior year of high school. Students take AP exams at the end of a school year, typically after taking a corresponding AP course. Performing well on the exams can earn students college credit while still in high school.

These testing data include a sibling identifier, which has been used in several other papers (e.g. Goodman et al., 2015; Gurantz, Hurwitz and Smith, 2020; Altmejd et al., 2021).[8] We exclude twins and multiples because our research question requires older and younger siblings. This leaves over 3 million sets of siblings, which corresponds to over 6 million individuals. This is a uniquely large sibling pool, especially relative to survey-based samples in the U.S. and elsewhere.[9]

[8] Generally speaking, siblings are defined as students with the same last name and home address. See Goodman et al. (2015) for more details and limitations.

[9] The vast majority of these siblings are sets of two. According to our calculations of 2023 ACS data, among households with at least two children residing in the house, 65 percent have exactly two children Ruggles et al. (2025). In our data, 86 percent of the families we can identify with at least two children have exactly two children. This discrepancy is likely partially explained by the limited range of high school cohorts we use. Thus, some families may have additional siblings that we do not observe since they fall outside this cohort range.

The testing data also contains several other pieces of information we rely on. First, it contains individuals' exact date of birth, which we rely on in our regression discontinuity design. It also tells us the birth order within siblings and the disparity in ages. Second, it includes test scores for each exam. Third, it contains basic demographics, such as sex, race/ethnicity, self-reported parental income/education, and the high school attended.

B. Voting Records

In the United States, each state collects and reports its own voting data, but all states publicly detail whether (but not for whom) each registered voter votes. These voting records are increasingly used in academic research as they provide the largest set of validated measures of this key mode of civic participation across the United States (e.g., Hersh, 2015; Fraga and Holbein, 2020; Barber and Holbein, 2022; Chyn and Haggag, 2023; Cohodes and Feigenbaum, 2021; Cantoni and Pons, 2021, 2022). We obtained nationwide voting data from the Data Trust, LLC—one of the many vendors in this space. The Data Trust combines and standardizes data from each state's election governing body on the hundreds of millions of people who are registered to vote and the tens of millions of people who vote in biannual national elections from 2004 to 2016.

The data record each individual's name, date of birth (DOB), state of residence, and voter-turnout status in various primary and general elections. Thus, for someone who was 18 years of age by November 2004, we observe their voting history between 2004 and 2016.[10] We collapse the voting records into a single observation per person—as opposed to a single observation per person-state—based on name and DOB. This process is detailed in Appendix B.

C. Matching

We match the education records to the voting records using students' name and date of birth (DOB). We limit the dataset to individuals who live in the 50 states or D.C., so we remove international test takers, students from U.S. territories, and the very small number of students missing a DOB.

We start by matching unique name and DOB combinations from each dataset and then employ a series of fuzzy matching methods, as described in Appendix B. To indicate potential false matches (and non-matches), we also construct a series of indicators—flags for common names, partially missing DOB in the voting data, and females (who are more likely to change their last name)—which we later use in robustness tests.

Importantly, an individual in the testing data that does not match to the voting data is considered a non-voter. Because of our conservative and imperfect matching process and because the only available datasets to benchmark

[10] Voters can be removed from a state's records. Removals vary over time and by state but do not impact most voter records. Our analyses focus on recent elections where voters have not had a chance to be removed.

our turnout data to that have demographic data are surveys which rely on self-reported voter turnout (which is over-reported due to social desirability; see Holbrook and Krosnick (2010)), our aggregate voting rates in each election are slightly below the true voting rates for people of a similar age. This is to be expected and should not bias our results. In addition, in a series of validation exercises in Appendix B, we show that the patterns of voting rates in our matched dataset follow those of nationally available statistics by age, race/ethnicity, and election.

D. Analytic Samples

The summary statistics for our finalized dataset are in Table 1. The first column is the full sample of test-takers, including over 23 million siblings and non-siblings while the second column is the sample of 6.5 million siblings, which is the focus of the paper.

The top portion of the table shows students' demographics and PSAT scores. The sample is slightly more likely to be female and 57 percent White. The average PSAT score is 91.

The first eligible election is displayed in the second panel. There are fewer people who are first eligible in 2004 because only some of the 2004 high school cohort are eligible for that election and almost all of the 2005 cohort's first eligible election is 2006.

The last panel shows the means of our outcome variables. 16.8 percent of students voted in their first eligible election. This is roughly the average between the next two outcomes—voted in the first Presidential Election and the first Midterm Election. 43.3 percent of students ever vote through 2016, which equates to a mean of 0.924 votes in general elections and only 0.166 votes in primary elections.

The sibling sample follows similar patterns but a few noticeable differences. For example, the sibling sample is more likely to be White and have higher PSAT scores. Their voting rates are also a bit higher than the overall sample.

Many of our analyses focus on individuals who turn 18 close to an election with a sibling relatively close in age. Columns 3-4 of Table 1 also includes summary statistics for those who turn 18 within 30 days of a Presidential Election (i.e., 2004, 2008, 2012). Column 3 includes individuals who are the oldest child in the family who also have a younger sibling at most 4 years younger. Column 4 includes individuals with an older sibling at most 4 years older. The sample sizes drop precipitously but most of the means are similar to the full sibling sample.

IV. Descriptive Statistics and Analyses

We begin by documenting several new facts about the voting habits of siblings in Table 2. First, 67 percent of two-sibling households have neither sibling voting in the first election where both siblings are eligible to vote. This is not surprising given the low voting rates of young adults detailed in Table 1. Only one sibling votes 24.7 percent

TABLE 1—INDIVIDUAL LEVEL SAMPLE SUMMARY STATISTICS.

	Full sample (1)	Siblings sample (2)	Older sibling RD sample (3)	Younger sibling RD sample (4)
Individual characteristics				
Female	0.527	0.518	0.519	0.518
White	0.568	0.671	0.680	0.659
Black	0.139	0.079	0.077	0.079
Asian	0.068	0.073	0.070	0.079
Hispanic	0.157	0.125	0.122	0.132
American Indian	0.008	0.006	0.006	0.006
Other race	0.041	0.036	0.036	0.037
Missing race	0.027	0.016	0.015	0.014
PSAT score	91.411	97.671	98.000	97.902
Missing PSAT score	0.188	0.146	0.143	0.144
Parental educ: no college	0.113	0.102	0.102	0.103
Parental educ: college but no BA degree	0.129	0.136	0.140	0.138
Parental educ: college with BA degree	0.287	0.406	0.412	0.410
Parental educ: missing	0.471	0.356	0.346	0.350
Parental income: less than $50,000	0.129	0.111	0.118	0.107
Parental income: $50,000 - $100,000	0.130	0.168	0.182	0.159
Parental income: greater than $100,000	0.089	0.138	0.143	0.146
Parental income: missing	0.652	0.583	0.558	0.588
First eligible election				
2004	0.106	0.087	0.128	0.023
2006	0.182	0.182	0.212	0.155
2008	0.201	0.234	0.288	0.190
2010	0.211	0.237	0.207	0.276
2012	0.215	0.192	0.159	0.226
2014	0.084	0.068	0.007	0.130
Voting outcomes				
Voted in 1st election	0.168	0.196	0.214	0.193
Voted in 1st presidential election	0.259	0.296	0.296	0.316
Voted in 1st midterm election	0.082	0.098	0.101	0.099
Ever voted	0.433	0.501	0.517	0.501
Total votes	0.924	1.090	1.195	1.012
Number of individuals	23,362,841	6,490,550	225,556	181,994

Notes: The sibling sample includes all individuals in the full sample with a sibling. The older sibling RD sample includes individuals in sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near the 2004, 2008, or 2012 elections. The younger sibling RD sample includes individuals in sibling pairs from families with 2 children with age differences between 9 and 48 months and where the younger sibling turned 18 near the 2004, 2008, and 2012 elections.

of the time and the older sibling is only slightly (one percentage point) more likely to be the voter than the younger sibling. The remaining 8.2 percent of the time both siblings vote. This implies that both siblings vote 25 percent of the time, conditional on at least one sibling voting.

Second, three-sibling households follow similar patterns with 57.5 percent of the trios not voting at all. Conversely, only 3.8 percent of the trios all vote in the youngest sibling's first election.

Third, the correlation coefficient among pairs of siblings voting is about 0.25, which is relatively modest and

TABLE 2—SIBLING VOTING RATES IN YOUNGEST SIBLING'S 1ST ELECTION.

	Sibling pair level	Family level	
	All sibling pairs (1)	Families with 2 children (2)	Families with 3 children (3)
Sibling vote rates			
Younger sibling voted	0.199	0.200	0.203
Middle sibling voted	–	–	0.207
Older sibling voted	0.211	0.211	0.231
Number of siblings voted			
Zero	0.670	0.671	0.575
One	0.249	0.247	0.266
Two	0.081	0.082	0.121
Three	–	–	0.038
Number of sibling pairs	3,886,316	–	–
Number of families	–	2,603,650	356,858

Notes: The sample in column 1 includes all siblings at the sibling pair level. The sample in columns 2 and 3 includes all siblings at the family level for families with 2 children (column 2) or 3 children (column 3).

implies that individuals within families do not typically have the same voting habits. However, there are some correlates that predict whether both siblings vote (Table A1), including relatively higher PSAT scores, parental income and education, and if both siblings are males. These are the same predictors of whether individuals vote in the data.

A. *Correlation Between Siblings' Voting*

We examine whether younger siblings are more likely to vote when their older sibling votes with the following descriptive regression:

$$v_{ie}^{y} = \alpha_0 + \alpha_1 v_{ie'}^{o} + X_i' A + E + \varepsilon_{ie} \tag{1}$$

The binary outcome v_{ie}^{y} is whether the individual i, who is a younger sibling, denoted by the superscript y, votes in election e. We are primarily interested in the coefficient α_1 on whether the older sibling, denoted by the superscript o, votes in election e'. For some analyses and specifications we consider the same elections for the older and younger siblings (i.e., $e = e'$). We include election year-by-state fixed effects E and, in some specifications, we control for the younger sibling's attributes (e.g., sex, test scores, high school), denoted X_i'.

The top panel of Table 3 shows that if older siblings vote in their first eligible election then their younger siblings are 11.4 percentage points more likely to vote in their first eligible election than if older siblings do not vote. Adding in controls for sex (column 2), PSAT scores (column 3), and high school fixed effects (column 4) slightly attenuates that relationship to 10 percentage points. The mean voting rate for younger siblings whose older sibling did not vote is roughly 20 percent, so 10 percentage points represents a 50 percent increase off of the base rate.

TABLE 3—DESCRIPTIVE RELATIONSHIP BETWEEN SIBLING VOTING PATTERNS.

	Outcome = younger sibling voted in 1st election			
	(1)	(2)	(3)	(4)
Older sibling voted in 1st election	0.114***	0.111***	0.108***	0.100***
	(0.001)	(0.001)	(0.001)	(0.001)
Outcome control mean	0.199	0.199	0.199	0.199
Number of sibling pairs	2,575,240	2,575,240	2,575,240	2,573,041
Older sibling voted in younger sibling's 1st election	0.175***	0.171***	0.168***	0.161***
	(0.001)	(0.001)	(0.001)	(0.001)
Outcome control mean	0.199	0.199	0.199	0.199
Number of sibling pairs	2,575,240	2,575,240	2,575,240	2,573,041
Election year by state fixed effects	✓	✓	✓	✓
Sibling demographics		✓	✓	✓
Sibling PSAT scores			✓	✓
Younger sibling high school fixed effects				✓

Notes: The samples includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months. Robust standard errors are reported.

The second panel of Table 3 mimics the top panel but the older siblings' voting variable is changed to whether the older siblings vote in the younger siblings' first eligible election. This election is often a few years after the older sibling's initial election, though sometimes it is the same. The last column shows that when older siblings vote in the younger siblings' first eligible election, the younger siblings are 16.1 percentage points (roughly 80 percent) more likely to vote than those whose older siblings did not vote.

The first two sub-panels of Table A2 display the same analyses but for the subsample of older siblings who are eligible to vote around Presidential Elections. Results are largely unchanged.

V. Voting Begets Voting

This section shows that voting around the age of 18 leads to voting in future elections. This serves three purposes. First, we set up our regression discontinuity design that we use moving forward. The probability of voting, which discontinuously jumps when someone turns 18, serves as our first stage. Second, whether someone votes in downstream elections that take place well after age 18 helps us understand what it means to be treated (i.e., vote around age 18) and some potential mechanisms. Third, whether voting is habit forming is an interesting research question in itself (e.g. Coppock and Green, 2016; Fujiwara, Meng and Vogl, 2016; Meredith et al., 2009; Plutzer, 2002; Gerber, Green and Shachar, 2003; Green and Shachar, 2000).

A. *Method - Regression Discontinuity*

We execute the regression discontinuity design by estimating the following equation:

$$v_{ie} = \beta_0 + \beta_1 Eligible_{ie} + f(age_i) + X_i' B + E + \nu_{ie} \tag{2}$$

As before, the outcome v_{ie} is whether individual i votes in election e. age_i is the running variable, which is the exact date of birth, re-centered at zero relative to the closest election to i's 18th birthday. As an example, November 4th was election day in 2008. Someone born on November 3rd 1990 would have $age_i = -1$ and be eligible to vote in the 2008 election. Someone born on November 5th 1990 would have $age_i = 1$ and be ineligible to vote in the 2008 election. $f(.)$ is the functional form of the running variable, which is most often $f(age_i) = age_i$.

We are primarily interested in the estimate of β_1, which is the coefficient on $Eligible_{ie}$. This binary variable equals one if individual i is eligible to vote in the election closest to his or her 18th birthday and zero otherwise ($Eligible_{ie} = 1$ if $age_i \leq 0$). β_1 represents the discontinuous jump in the probability of voting relative to the trend in the running variable age_i. We always control for election fixed effects E and some specifications include individual-level controls X_i, but our main specifications do not.

We consider several outcomes, including voting in the election around the age of 18, followed by voting in subsequent elections. After the election around age 18, all the students become eligible to vote, despite using the same instrument $Eligible_{ie}$. Therefore, β_1 is a reduced-form coefficient for outcomes after election e.

Our primary specification uses a bandwidth of 30 days. That implies that we use the small subset of people born in October and November of an election year. The specification only uses the linear version of age, which is graphically justified, and uses robust standard errors. Results are not sensitive to any of these alternative specifications; these are presented in the Appendix.

Support for the Regression Discontinuity Design

Date of birth is unlikely to be manipulated but there are still two lingering concerns. First, date of birth determines whether someone takes the PSAT, SAT, or AP. This may very well be true for age-of-school entry, which is most frequently around August or September, but not around early November.

Second, people in the U.S. are less likely to be born on weekends. This creates lumpiness on certain birth dates, including right around the voting eligibility age/date (see Figure A1 and Figure A2). This can lead to some imbalance in some election years, but these weekend dips are visually obvious (2 out of 7 days) in election and non-election cohorts.

In Table A3, we estimate covariate balance tests that provide support for the research design. Specifically, we regress one covariate at a time (e.g., sex, PSAT scores, etc.) on the running variable and the eligibility threshold. We do not find evidence that the students' characteristics differ from one another on either side of the eligibility threshold.

B. Results

Using all siblings, Table 4 reveals two important results. First, in column 1, those who are barely eligible to vote around age 18 during an election year are 21.5 percentage points more likely to vote than those who are just barely not yet age 18 and ineligible to vote. This implies that 21.5 percent of those who are barely age 18 vote since those who are not yet 18 do not vote at all (visually confirmed in Figure A3). The estimate is 31 percentage points and 9.3 percentage points in Presidential and Midterm Election years, respectively.

The estimates in column 1 serve as the first stage in future analyses. Sometimes the samples change when we focus on an older or younger sibling or a particular election, but the magnitude of the coefficients are roughly the same, as seen across panels in Figure A3. Additionally, the larger coefficient on Presidential Elections provides relatively more statistical power and so many of our main analyses focus on Presidential Elections.

Second, Table 4 shows that voting begets voting, which is visually confirmed in Figure 1. The reduced-form estimate on being eligible to vote from Equation 4 is 1.3 percentage points. This means that students who are just barely 18 before an election are 1.3 percentage points more likely to vote two years later in the next election than those who were just barely not 18. This is a 10 percent increase over the mean voting rate for those who were still 17 at the time of the initial election.

If we scale the 1.3 percentage points by the first column, which serves as a first-stage, our estimate suggests that voting around age 18—as opposed to being ineligible to vote—increases the probability of voting again two years later by 6 percentage points (35 percent).

Being eligible to vote around 18 also increases the probability of voting in four years by 0.9 percentage points, the probability of ever voting by 1.3 percentage points, the number of general election votes cast by 0.046 (see Figure A4 for visual evidence of these alternative outcomes). Overall, these results show that voting around 18 leads to future voting.

The previous results are driven mostly by students who turn 18-years-old around a Presidential Election. This can be seen in the next two panels of Table 4, which are split by students around age 18 during a Presidential Election and Midterm Election, respectively.

TABLE 4—EFFECT OF ELIGIBILITY AT AGE 18 ON OWN FUTURE VOTING, ALL SIBLINGS SAMPLE.

	First stage	Own future voting outcomes			
	Voted at age 18 (1)	Voted in election 2 years later (2)	Voted in election 4 years later (3)	Ever voted in the future (4)	Total future votes (5)
Panel A: 18th birthdays near any election					
Eligible to vote at age 18	0.215***	0.013***	0.009***	0.013***	0.046***
	(0.002)	(0.002)	(0.002)	(0.003)	(0.007)
Outcome control mean	0.007	0.174	0.223	0.480	0.949
Observations	511,470	511,470	511,470	511,470	511,470
Panel B: 18th birthdays near presidential elections					
Eligible to vote at age 18	0.310***	0.021***	0.016***	0.014***	0.057***
	(0.002)	(0.002)	(0.003)	(0.004)	(0.009)
Outcome control mean	0.007	0.085	0.310	0.463	0.893
Observations	286,614	286,614	286,614	286,614	286,614
Panel C: 18th birthdays near midterm elections					
Eligible to vote at age 18	0.093***	0.004	0.002	0.012***	0.032***
	(0.002)	(0.004)	(0.003)	(0.004)	(0.011)
Outcome control mean	0.006	0.289	0.112	0.502	1.021
Observations	224,856	224,856	224,856	224,856	224,856

Notes: The sample includes all siblings with an 18th birthday within 30 days of an election. Regressions include fixed effects for the election year that is closest to their 18th birthday. Bandwidth is 30 days. Robust standard errors are reported.

These results are insensitive to choice of bandwidth, weighting, and polynomial order of the running variable. These robustness tests are shown in Table A4.[11] Table A5 (complemented by Figure A5) and Table A6 (complemented by Figure A6) show qualitatively similar results to Table 4 when using the full sample that includes non-siblings and the subsample of only older siblings, respectively.

VI. Siblings' Spillovers - Methods

We use a regression discontinuity design to assess whether one sibling's voting impacts another sibling's probability of voting. For exposition's sake, the equations below are spillovers from the older sibling to the younger sibling but spillovers in the reverse direction use the same conceptual setup.

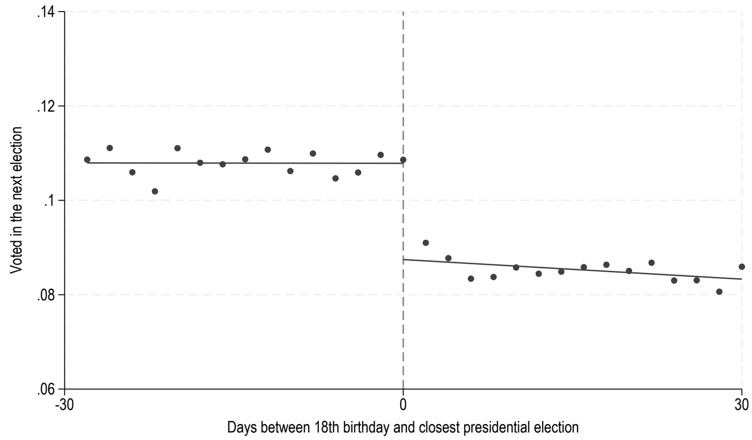
Building on Equation 2, we begin with a first-stage equation that estimates the discontinuous jump in the probability of voting when the older sibling turns 18 around an election with the following:

$$v_{ie}^{o} = \gamma_0 + \gamma_1 Eligible_{ie}^{o} + f(age_i^{o}) + X_i^{o}\Gamma + E + \upsilon_{ie} \tag{3}$$

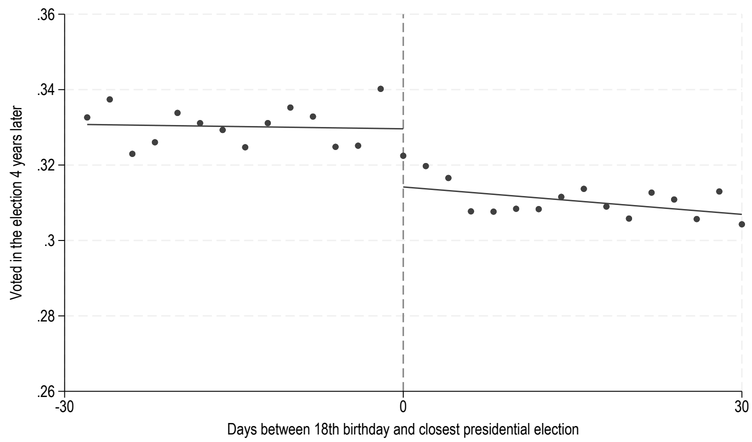
[11] Although not shown, results are also robust to including individual-level controls X_i. These unchanged results further supports the validity of our research design.

FIGURE 1. THE EFFECT OF AGE 18 ELIGIBILITY ON FUTURE VOTING - ALL SIBLINGS SAMPLE

(a) Voted in election 2 years later



(b) Voted in election 4 years later



Note: This figure shows binned scatterplots of average voting outcomes by days between each sibling's 18th birthday and the presidential elections of 2004, 2008, and 2012. Each bin consists of 2 days.

The important distinction with Equation 2 is the superscript o on most of the variables that designates older siblings. The alternative is the superscript y for younger siblings, which appears in later equations.

We are interested in the coefficient γ_1, which measures the increase in the probability of the older sibling voting in election e when he or she turns 18 and becomes eligible to vote. Theoretically, nobody under age 18 should vote so the coefficient can also be interpreted as the voting rate for those who just turned 18.

Next, we consider the impact of the older sibling voting on the younger sibling's probability of voting through the following reduced-form equation:

(4) $$v_{ie}^{y} = \delta_0 + \delta_1 Eligible_{ie'}^{o} + f(age_i^{o}) + X_i^{y}\Omega + E + \omega_{ie}$$

The outcome variable is whether the younger sibling voted in election e. This is the main distinction with the first-stage where the outcome was for the older sibling. But the right-hand-side running and eligibility variables remain about the older sibling. Additionally, the elections that the older and younger siblings need not be the same (i.e., $e \neq e'$).

Lastly, we combine the first-stage and reduced-form equations into a two-stage least squares analysis with the following second stage:

$$v^y_{ie} = \lambda_0 + \lambda_1 \hat{v}^o_{ie'} + f(age^o_i) + X^y_i \Lambda + E + \psi_{ie} \tag{5}$$

$\hat{v}^o_{ie'}$ is the instrumented endogenous variable from Equation 3. As such, λ_1 represents the causal impact of whether the older sibling votes in election e' on whether the younger sibling votes in election e.

A. *Identification and Implementation*

Identification comes from as-good-as random variation in the exact date of birth and the election date. We have no reason to believe that someone born on November 5th versus November 6th are different from one another in any discernible way. We already demonstrated the lack of differences in the covariates Table A3. Columns 2 and 3 of this table repeats this exercise separately for older siblings and younger siblings, respectively, who turn 18 within 30 days of a Presidential Election and finds no concerns about manipulation of the running variable.

We focus our results on siblings who are within four years of age where sibling influences are likely more salient. In practice, we use a bandwidth of 30 days on either side of the eligibility threshold and a linear functional form for the running variable. We later test the sensitivity of these decisions. We report robust standard errors in all results.

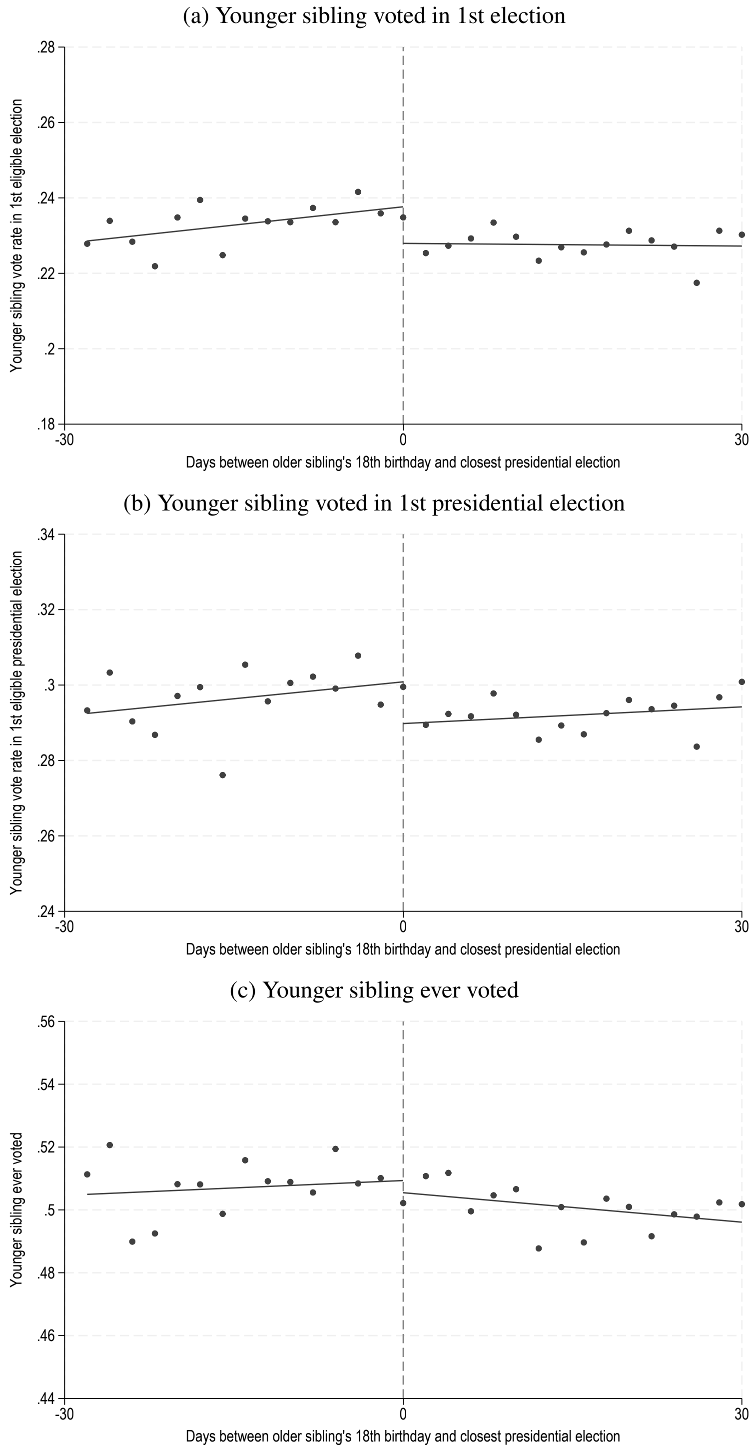
Lastly, we focus on the subsample of older (and younger) siblings who turn 18 around a Presidential Election (i.e., around November of 2004, 2008, 2012). Voter turnout among those just barely 18 years old in Presidential Elections is roughly 30 percent, as seen in Figure A3, compared to 10 percent in Midterm Election years. This larger first-stage helps with statistical power despite dropping siblings who turn 18 around the 2006 and 2010 elections. We show results using the full sample and the subsample of Midterm Elections in the appendix.

VII. Older Siblings' Spillovers

A. *Main Results*

Older siblings who vote around 18 years increase the probability that their younger siblings vote. Table 5 shows the reduced-form and second stage results, which are supported by Figure 2.

FIGURE 2. EFFECT OF OLDER SIBLING'S AGE 18 ELIGIBILITY ON YOUNGER SIBLING'S VOTING



Note: This figure shows binned scatterplots of average younger sibling voting outcomes by days between the older sibling's 18th birthday and the presidential elections of 2004, 2008, and 2012. Voting outcomes are residualized for fixed effects of the younger sibling's first election to match our estimating equation. Each bin consists of 2 days.

TABLE 5—OLDER SIBLING RD, 18TH BIRTHDAYS NEAR PRESIDENTIAL ELECTIONS.

	1st stage	Younger sibling voting outcomes								
	Older sibling voted at age 18 (1)	Voted in 1st election (2)	Voted in 1st pres. election (3)	Voted in 1st midterm election (4)	Ever registered to vote (5)	Ever voted (6)	Voted in 2+ elections (7)	Voted in 3+ elections (8)	Total votes (9)	Total votes after 1st election (10)
Older sibling eligible to vote at age 18 (RF)	0.295***	0.010**	0.011**	-0.001	0.003	0.004	0.014**	0.007*	0.032**	0.022*
	(0.004)	(0.005)	(0.005)	(0.004)	(0.006)	(0.006)	(0.005)	(0.004)	(0.016)	(0.013)
	[5,665]									
Older sibling voted at age 18 (IV)		0.033**	0.038**	-0.004	0.010	0.014	0.046**	0.025*	0.108**	0.075*
		(0.016)	(0.018)	(0.012)	(0.019)	(0.020)	(0.018)	(0.014)	(0.054)	(0.044)
Outcome control mean	0.009	0.229	0.293	0.099	0.667	0.501	0.294	0.152	1.039	0.810
Observations	112,778	112,778	112,778	112,772	112,778	112,778	112,778	112,778	112,778	112,778

Notes: The sample includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year the younger sibling is first eligible for. Bandwidth is 30 days. First stage F-stat reported in brackets in column 1. Robust standard errors are reported.

We find that younger siblings are 1 percentage point more likely to vote in their first election if their older sibling is just barely eligible to vote at age 18 rather than just barely ineligible to vote at age 18. This reduced-form estimate is small in magnitude but precisely measured. When scaled by the first-stage (column 1) in the instrumental variables framework, it implies that older siblings' voting around age 18 increases their younger siblings' probability of voting by 3.3 percentage points. To put that into perspective, 3.3 percentage points is a 14.4 percent increase relative to the 22.9 percent voting rate of the control mean. Putting it all together, younger siblings are 14.4 percent more likely to vote in their first election when their older sibling votes around age 18. This is visually confirmed in Panel A of Figure 2.[12]

The next several columns show different outcomes for the younger sibling. Columns 3 and 4 show that the previous estimate is driven by younger siblings whose first election was a Presidential Election, which is four years after the older sibling turns 18. Column 5 shows that the impacts are not driven by increasing the younger sibling's probability of registering to vote.

The next five columns in Table 5 show how these spillover effects impact younger siblings' cumulative voting record across the multiple elections in our data. In column 5, we estimate no statistically significant effects on whether the younger sibling ever voted in our data (column 6), but we do find evidence of impacts on younger siblings' probability of voting more than once (columns 7-8) and impacts on voting beyond their first eligible election (columns 9-10). Together, these results imply that the observed spillover effects persist across elections but are concentrated among the subset of younger siblings who would have registered and voted at least once even

[12]Additional outcomes that mirror the columns of Table 5 are shown in Figure A7.

absent their older sibling's turnout.

The statistical significance of the main results is somewhat sensitive to choice of bandwidth, kernel weight, and polynomial order, as shown in Table A7. But across specification choices, the coefficients are consistently positive and of a similar magnitude when the outcome is either voting in the focal sibling's first eligible election or first eligible Presidential Election. Moreover, as column 6 of Table A7 shows, the statistical significance holds for nearly all outcomes when using the optimal bandwidth and bias-corrected inferences procedures from Calonico et al. (2017). In a placebo test, we also find no sibling spillovers in the reduced-form estimates among older siblings turning 18 years old in odd years, when there were no national elections (Table A8 and Figure A8).

Finally, we find that the effects are either muted or null when we include older siblings around age 18 in a Midterm Election (Table A9). This is likely due to the weaker first stage for Midterm Elections and perhaps because of the smaller effects on one's own future voting when they are 18 around a Midterm Election.

B. Heterogeneity and Potential Mechanisms

We explore several types of heterogeneous effects to determine where the impacts of siblings' voting are strongest and to probe potential mechanisms. We begin with characteristics of the siblings and family in Table 6.

TABLE 6—OLDER SIBLING RD, HETEROGENEITY RESULTS BY SIBLING AND FAMILY CHARACTERISTICS.

	Outcome = Younger sibling voted in 1st election							
	Same sex sibling pairs (1)	Opposite sex sibling pairs (2)	Sibling age diff. 2 years or less (3)	Sibling age diff. between 2-4 years (4)	Siblings in same state (5)	Siblings in diff. states (6)	Parent ed: No BA degree (7)	Parent ed: BA degree (8)
Older sibling eligible to vote at age 18 (RF)	0.022***	-0.003	0.008	0.011	0.005	0.020*	0.006	0.018**
	(0.007)	(0.007)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)	(0.008)
Older sibling voted at age 18 (IV)	0.074***	-0.009	0.026	0.037	0.017	0.061*	0.023	0.051**
	(0.023)	(0.024)	(0.021)	(0.023)	(0.020)	(0.031)	(0.038)	(0.023)
Outcome control mean	0.223	0.235	0.107	0.299	0.228	0.237	0.216	0.264
Observations	57,652	55,126	41,339	71,438	82,393	26,869	26,139	45,944

Notes: The sample includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year the younger sibling is first eligible for. Bandwidth is 30 days. Robust standard errors are reported.

First, we explore whether effects are larger among siblings who are similar to one another in various respects. While not a direct test, this is motivated by our discussion of role models and conforming siblings in Section II. In columns 1 and 2, we show that same-sex siblings are driving the results, and there are no spillovers among opposite-sex siblings (we show this visually in Figure A9).[13] Same-sex siblings are 7.4 percentage points (33 percent) more likely to vote in their initial election when their older sibling previously voted around age 18. The difference in the reduced-form estimates between same-sex and opposite-sex siblings is statistically significant.

Second and conversely, we do not find strong evidence that age gaps between siblings are driving the results (columns 3 and 4). This may imply that the "closeness" of siblings or whether an older sibling is viewed as a role model is determined more by sex than age.

Third, we estimate larger spillovers among siblings who live in different states from one another than among siblings who live in the same state (columns 5 and 6). (Since we only know whether the siblings are in different states by the college they attend[14], this also represents a heterogeneity by whether at least one sibling attends college out-of-state.) While we cannot attribute this heterogeneity to geographic distance *per se*, this result is at odds with the economic theory that transportation and time costs are the mechanism behind our results. If these costs were particularly high, we would expect to find larger impacts among siblings living in the same state (or household).

Lastly, results in columns 7 and 8 suggest that voting spillovers are stronger among siblings whose parents have a bachelor's degree than among siblings whose parents do not. There is an underlying positive relationship between parental education and voting; perhaps these estimated spillovers contribute to those turnout disparities.

Next, we examine heterogeneous effects by younger siblings' individual attributes. Following on results in Table 5 which indicated that spillover effects are concentrated among younger siblings who would have voted at least once absent treatment, we first probe for heterogeneities by younger siblings' latent voting propensity. Using the non-sibling sample, we begin by predicting each younger sibling's probability of voting based on his or her attributes, but not on the sibling's attributes. We use basic demographics, high school fixed effects, and PSAT scores. Applying those estimates to our younger sibling sample, we then split the sample into terciles by predicted voting probability in Table 7. We estimate the largest effects for younger siblings in the middle tercile of voting probability, suggesting that the lowest- and highest-propensity voters are the least impacted by their siblings' voting.

Columns 5 and 6 suggest that our estimates are larger when the younger sibling faces an election around age 19, not 18. This can happen for a variety of reasons related to the difference in siblings' ages, given that the elections we consider are every two years. Regardless, at age 19, the younger sibling is more likely to be in college than the 18 year old. Lastly, columns 6 and 7 show a larger spillover for students with higher PSAT scores.

[13] Table A10 splits by all four combinations of siblings sex combinations, which shows that sisters are the largest contributors to the spillover effects.

[14] Test taking data is linked to National Student Clearinghouse data, which is a near census of college enrollment in the U.S.

TABLE 7—OLDER SIBLING RD, HETEROGENEITY RESULTS BY INDIVIDUAL CHARACTERISTICS.

	Outcome = Younger sibling voted in 1st election						
	Lowest tercile predicted voter (1)	Middle tercile predicted voter (2)	Highest tercile predicted voter (3)	Younger sibling age 18 at 1st election (4)	Younger sibling age 19 at 1st election (5)	Below median PSAT score (6)	Above median PSAT score (7)
Older sibling eligible to vote at age 18 (RF)	0.004	0.020**	0.010	0.006	0.016*	0.003	0.016**
	(0.004)	(0.008)	(0.010)	(0.006)	(0.008)	(0.007)	(0.008)
Older sibling voted at age 18 (IV)	0.018	0.072**	0.027	0.019	0.053*	0.010	0.049**
	(0.019)	(0.030)	(0.028)	(0.020)	(0.027)	(0.027)	(0.024)
Outcome control mean	0.046	0.194	0.445	0.201	0.269	0.209	0.258
Observations	36,923	36,924	36,924	66,925	45,611	47,919	47,901

Notes: The sample includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year the younger sibling is first eligible for. Bandwidth is 30 days. Robust standard errors are reported.

VIII. Younger Siblings' Spillovers

A. Main Results

Younger siblings who vote around age 18 increase the probability that their older siblings vote. Table 8 shows these results, which are supported by Figure 3. We find that older siblings whose younger siblings are eligible to vote in a presidential election around age 18 are 1.3 percentage points more likely to vote in that election than if their younger siblings were ineligible to vote around age 18. When this reduced-form effect is scaled by the first-stage coefficient (reported in column 1), we estimate that younger siblings' voting increases their older siblings' probability of voting in the same election by 4 percentage points (13.7 percent). The results in column 2 of Table 8 are insensitive to choice of bandwidth, kernel weight, and polynomial order, as shown in Table A11.

Column 3 shows that this spillover from younger to older siblings does not persist to future elections. We estimate no statistically significant effect on older siblings' voting rates in the election 2 years later. The coefficients in columns 6 and 9 are directionally suggestive of small downstream impacts but are not statistically significant.

Also, unlike the spillover results in the other direction, we find that including sibling pairs with younger siblings turning 18 around Midterm Elections yields similarly large impact estimates. This can be seen in the bottom panel of Table 8.

The fact that we find spillovers from younger siblings to older siblings implies that older siblings acting as role models or younger siblings conforming to older siblings' behavior is not the only mechanism at play, if operative at all. This is consistent with alternative hypotheses in the previous section as well.

TABLE 8—YOUNGER SIBLING RD.

	1st stage	Older sibling voting outcomes							
	Younger sibling voted at age 18 (1)	Voted in election closest to younger sibling's 18th birthday (2)	Voted in election 2 years later (3)	Ever registered to vote (4)	Ever voted (5)	Voted in 2+ elections (6)	Voted in 3+ elections (7)	Total votes (8)	Total votes after election closest to younger sibling's 18th birthday (9)
Panel A: Younger sibling's 18th birthday near presidential election									
Younger sibling eligible to vote at age 18 (RF)	0.332***	0.013**	-0.001	0.004	0.007	0.012**	0.001	0.027	0.014
	(0.004)	(0.006)	(0.004)	(0.006)	(0.007)	(0.006)	(0.005)	(0.019)	(0.015)
	[5,567]								
Younger sibling voted at age 18 (IV)		0.040**	-0.003	0.011	0.020	0.037**	0.003	0.082	0.042
		(0.018)	(0.013)	(0.019)	(0.020)	(0.018)	(0.015)	(0.057)	(0.045)
Outcome control mean	0.006	0.292	0.124	0.674	0.519	0.316	0.177	1.139	0.847
Observations	90,997	90,997	90,997	90,997	90,997	90,997	90,997	90,997	90,997
Panel B: Younger sibling's 18th birthday near any election									
Younger sibling eligible to vote at age 18 (RF)	0.223***	0.009**	-0.000	0.006	0.006	0.006	0.001	0.024	0.014
	(0.003)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.004)	(0.015)	(0.013)
	[6,022]								
Younger sibling voted at age 18 (IV)		0.042**	-0.001	0.027	0.028	0.025	0.005	0.107	0.065
		(0.017)	(0.017)	(0.021)	(0.022)	(0.021)	(0.017)	(0.067)	(0.057)
Outcome control mean	0.006	0.206	0.202	0.666	0.519	0.331	0.199	1.213	1.007
Observations	165,751	165,751	165,751	165,751	165,751	165,751	165,751	165,751	165,751

Notes: The sample includes sibling pairs from families with 2 children with age differences between 9 and 48 months and where the younger sibling turned 18 near the 2004, 2006, 2008, 2010, 2012, or 2014 elections. Regressions include fixed effects for the election year that is closest to the younger sibling's 18th birthday. Bandwidth is 30 days. Robust standard errors are reported.

B. *Heterogeneity and Potential Mechanisms*

Once again, we explore heterogeneous effects to ascertain where sibling spillover effects are strongest and to probe potential mechanisms. The first two columns of Table 9 show that the point estimate for opposite-sex siblings is larger than for same-sex siblings.[15] This is the opposite result from the spillovers from the older to younger sibling, although the difference in the reduced-form estimates between opposite- and same-sex siblings is not statistically significant at conventional levels.[16] Once again, we find no evidence that sibling pairs with smaller age gaps drive the results, and there is evidence that spillover effects are larger among siblings whose parents have a bachelor's degree.

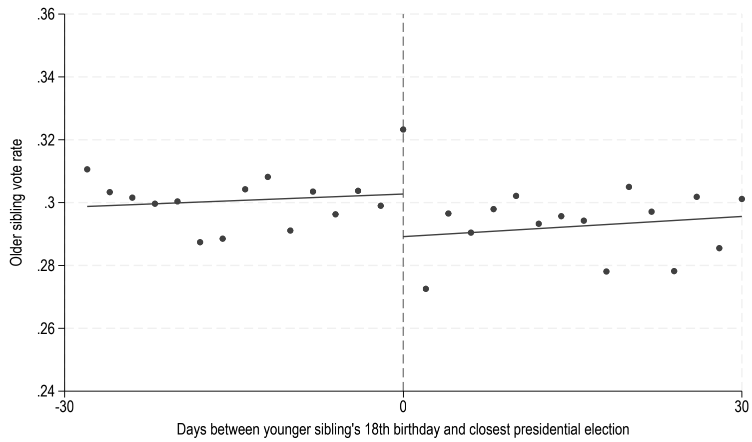
Combining the results on spillovers from older siblings to younger siblings and vice versa, we find inconsistent results on the role of sex composition but consistently null results on differences by age gap. This makes it less likely that "close" siblings are driving the results, and some of the proposed mechanisms (e.g., role model dynamics) should be de-emphasized.

[15] We show this visually in Figure A10.

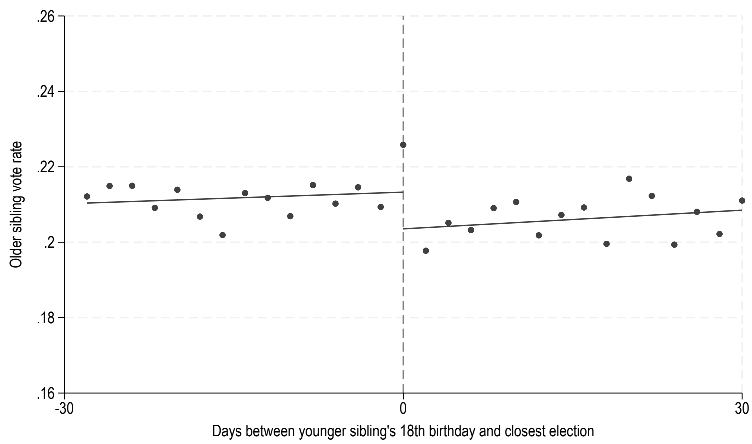
[16] Table A12 splits by all four combinations of siblings sexes and older sisters and younger brothers are the largest contributors to the spillover effects.

FIGURE 3. EFFECT OF YOUNGER SIBLING'S AGE 18 ELIGIBILITY ON OLDER SIBLING'S VOTING

(a) 18th birthdays near presidential elections



(b) 18th birthdays near any election



Note: This figure shows binned scatterplots of average older sibling vote rates in their younger sibling's 1st election by days between the younger sibling's 18th birthday and the election. Panel (a) includes sibling pairs where the younger sibling turned 18 near the presidential elections of 2004, 2008, and 2012. Panel (b) includes midterm elections. Voting outcomes are residualized for fixed effects of the election closest to the younger sibling's 18th birthday to match our estimating equation. Each bin consists of 2 days.

Table 10 shows heterogeneous effects by students' individual characteristics. There is suggestive evidence, at best, that spillover effects are once again greatest for older siblings in the middle tercile of predicted voting probability. In columns 4 and 5, we find that spillover effects are slightly larger when the older sibling was 18 during their first election, compared to 19. This was determined by cohort and date of birth and might imply that older siblings' prior opportunity to vote does not matter much. And again, older siblings with relatively high PSAT scores are most impacted by their younger sibling voting.

TABLE 9—YOUNGER SIBLING RD, HETEROGENEITY RESULTS BY SIBLING AND FAMILY CHARACTERISTICS.

	Outcome = Older sibling voted in election closest to younger sibling's 18th birthday					
	Same sex sibling pairs (1)	Opposite sex sibling pairs (2)	Sibling age diff. 2 years or less (3)	Sibling age diff. 2-4 years (4)	Parent ed: No BA degree (5)	Parent ed: BA degree (6)
Older sibling eligible to vote at age 18 (RF)	0.008	0.019**	0.011	0.015*	0.002	0.016*
	(0.008)	(0.009)	(0.010)	(0.008)	(0.012)	(0.009)
Older sibling voted at age 18 (IV)	0.025	0.056**	0.036	0.042*	0.008	0.039*
	(0.026)	(0.026)	(0.031)	(0.022)	(0.045)	(0.023)
Outcome control mean	0.290	0.295	0.291	0.293	0.284	0.329
Observations	45,961	45,036	35,696	55,301	23,233	38,688

Notes: The sample includes sibling pairs from families with 2 children with age differences between 9 and 48 months and where the younger sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year that is closest to the younger sibling's 18th birthday. Bandwidth is 30 days. Robust standard errors are reported.

TABLE 10—YOUNGER SIBLING RD, HETEROGENEITY RESULTS BY INDIVIDUAL CHARACTERISTICS.

	Outcome = Older sibling voted in election closest to younger sibling's 18th birthday						
	Lowest tercile predicted voter (1)	Middle tercile predicted voter (2)	Highest tercile predicted voter (3)	Older sibling age 18 at 1st election (4)	Older sibling age 19 at 1st election (5)	Below median PSAT score (6)	Above median PSAT score (7)
Older sibling eligible to vote at age 18 (RF)	0.010	0.019*	0.017	0.019**	0.010	0.009	0.023**
	(0.007)	(0.011)	(0.012)	(0.010)	(0.008)	(0.009)	(0.010)
Older sibling voted at age 18 (IV)	0.043	0.056*	0.039	0.056**	0.031	0.033	0.057**
	(0.030)	(0.031)	(0.027)	(0.028)	(0.024)	(0.032)	(0.023)
Outcome control mean	0.098	0.282	0.502	0.295	0.291	0.276	0.323
Observations	28,941	28,941	28,941	36,545	53,917	40,355	38,215

Notes: The sample includes sibling pairs from families with 2 children with age differences between 9 and 48 months and where the younger sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year that is closest to the younger sibling's 18th birthday. Bandwidth is 30 days. Robust standard errors are reported.

IX. Conclusion

This paper documents the impact siblings have on one another's probability of voting. Leveraging data on millions of U.S. siblings and a discontinuity in their age of voting eligibility, we find that older siblings' voting influences younger siblings to vote and vice versa. As far as we know, we are the first to document sibling spillovers in voting. To our knowledge, this is also the only study to show spillovers among siblings that run in both directions—from older to younger siblings and vice versa.

Our findings are important both within and beyond the realm of voting behavior. First, our results establish that siblings influence one another's voting activity in young adulthood, providing some of the most compelling evidence to date that siblings, like parents, play a formative role in young adults' political behavior. At the same time, our findings suggest that sibling influence differs in important ways from parental influence: we estimate bidirectional sibling spillovers of equal magnitude, which suggests that, at least in voting, sibling influence is quite mutual.

Second, the existence of sibling spillovers in voter turnout implies that many turnout-inducing policies, interventions, and campaigns exert aggregate turnout effects larger than previously known, as few prior impact estimates have accounted for sibling spillovers in voting.[17] More generally, our results underscore that impact evaluations that do not account for spillovers among siblings and close peers are liable to understate the benefits and costs of policies and interventions across a wide array of policy domains. Future research should continue to investigate what kinds of policy effects exhibit the largest spillovers and what kinds of behaviors are subject to the greatest peer- and family influence.

Third, voting is an important act of civic engagement in the U.S. and young adults cast their ballots at alarmingly low rates. We show that siblings can play a role in increasing voting rates. Policy and interventions can exploit this knowledge to improve youth turnout, perhaps by targeting particular family structures or specific individuals within families. At the same time, our results suggest that sibling turnout alone is unlikely to improve young adults' low rates of voter registration or reduce the sizeable population of "never voters"; accomplishing these objectives likely requires complementary turnout strategies.

Our results also suggest several promising directions for future research. For one, scholars still have much to learn about the role of the family in voting. We have focused on siblings, but parents and more distant family members are worthy of further exploration as well. We also do not know how siblings impact one another's voting in later adulthood, when voting rates are higher.

Additionally, results from our younger-to-older sibling design show that changes in individuals' voting eligibility (due to aging) can increase turnout among their already eligible family members. Future research might investigate whether similar effects arise when voting eligibility expands within families due to changes in *de jure* voting rights, such as reductions in voting age eligibility or the extension of voting rights from men to women.

Finally, future work can better identify mechanisms and specify under what conditions sibling spillovers are strongest. Two of our findings in particular invite further study: First, it would be helpful to understand why sibling spillovers in voting are largest among those with higher test scores and parental education, disentangling the potential roles of turnout norms, (non)cognitive skills, and voter registration. Second, it would be useful to under-

[17] By the same token, turnout-inducing treatments can be expected to exert smaller aggregate effects in low-fertility and/or sparse-social-network settings, where treated individuals' voting behaviors spill over onto fewer siblings or close peers. Given that many of the world's democracies have experienced low and declining fertility and rising social isolation among youth, this implication has clear political relevance.

stand why sibling voting does not increase registration and instead boosts turnout only among those who would have registered and voted at least once anyway.[18] Answering these questions would equip social scientists with a clearer understanding of how sibling influence operates and a firmer sense of its scope, power, and limitations.

[18] One hypothesis is that sibling turnout influences young adults' interest in voting but does not impart knowledge or skills required to complete the voter registration process. Another possibility is that registration simply requires more effort than voting once registered, and the nudge of a sibling's turnout is sufficient to motivate the latter but not the former.

REFERENCES

Achen, Christopher H. 2002. "Parental socialization and rational party identification." *Political behavior*, 24: 151–170.

Akee, Randall, William Copeland, John B Holbein, and Emilia Simeonova. 2020. "Human capital and voting behavior across generations: evidence from an income intervention." *American Political Science Review*, 114(2): 609–616.

Ali, S Nageeb, and Charles Lin. 2013. "Why people vote: Ethical motives and social incentives." *American economic Journal: microeconomics*, 5(2): 73–98.

Altmejd, Adam, Andrés Barrios-Fernández, Marin Drlje, Joshua Goodman, Michael Hurwitz, Dejan Kovac, Christine Mulhern, Christopher Neilson, and Jonathan Smith. 2021. "O brother, where start thou? Sibling spillovers on college and major choice in four countries." *The Quarterly Journal of Economics*, 136(3): 1831–1886.

Alwin, Duane F. 2016. "Families' Values: How Parents, Siblings, and Children Affect Political Attitudes."

Arnzen, Cameron J, and Sarah R Cohodes. 2025. "Education and the Gender Voting Gap."

Bansak, Cynthia, Xuan Jiang, and Guanyi Yang. 2022. "Sibling spillovers in rural China: A story of sisters." *China Economic Review*, 76: 101873.

Barber, Michael, and John B Holbein. 2022. "400 million voting records show profound racial and geographic disparities in voter turnout in the United States." *Plos one*, 17(6): e0268134.

Bell, D'Wayne, Jing Feng, John Holbein, and Jonathan Smith. 2024. "Do STEM Students Vote?" *American Educational Research Journal*, 61(1): 48–73.

Bettinger, Eric P, and Bridget Terry Long. 2005. "Do faculty serve as role models? The impact of instructor gender on female students." *American Economic Review*, 95(2): 152–157.

Black, Sandra E. 2017. "New evidence on the impacts of birth order." *NBER Reporter*, , (4): 15–18.

Black, Sandra E, Erik Grönqvist, and Björn Öckert. 2018. "Born to lead? The effect of birth order on noncognitive abilities." *Review of Economics and Statistics*, 100(2): 274–286.

Black, Sandra E, Paul J Devereux, and Kjell G Salvanes. 2005. "The more the merrier? The effect of family size and birth order on children's education." *The Quarterly Journal of Economics*, 120(2): 669–700.

Black, Sandra E, Sanni Breining, David N Figlio, Jonathan Guryan, Krzysztof Karbownik, Helena Skyt Nielsen, Jeffrey Roth, and Marianne Simonsen. 2021. “Sibling spillovers.” *The Economic Journal*, 131(633): 101–128.

Blais, André, Jean-François Daoust, Ruth Dassonneville, and Gabrielle Péloquin-Skulski. 2019. “What is the cost of voting?” *Electoral Studies*, 59: 145–157.

Bratsberg, Bernt, Christopher T Dawes, Andreas Kotsadam, Karl-Oskar Lindgren, Richard Öhrvall, Sven Oskarsson, and Oddbjørn Raaum. 2022. “Birth order and voter turnout.” *British Journal of Political Science*, 52(1): 475–482.

Calonico, Sebastian, Matias D. Cattaneo, Max H. Farrell, and Rocío Titiunik. 2017. “rdrobust: Software for Regression Discontinuity Designs.” *The Stata Journal*, 17(12): 372–404.

Cantoni, Enrico. 2020. “A precinct too far: Turnout and voting costs.” *American Economic Journal: Applied Economics*, 12(1): 61–85.

Cantoni, Enrico, and Vincent Pons. 2021. “Strict ID laws don’t stop voters: Evidence from a US nationwide panel, 2008–2018.” *The Quarterly Journal of Economics*, 136(4): 2615–2660.

Cantoni, Enrico, and Vincent Pons. 2022. “Does context outweigh individual characteristics in driving voting behavior? Evidence from relocations within the United States.” *American Economic Review*, 112(4): 1226–1272.

Chakravarty, Surajeet, Todd R Kaplan, and Gareth Myles. 2018. “When costly voting is beneficial.” *Journal of Public Economics*, 167: 33–42.

Chen, M Keith, Kareem Haggag, Devin G Pope, and Ryne Rohla. 2022. “Racial disparities in voting wait times: Evidence from smartphone data.” *Review of Economics and Statistics*, 104(6): 1341–1350.

Chyn, Eric, and Kareem Haggag. 2023. “Moved to vote: The long-run effects of neighborhoods on political participation.” *Review of Economics and Statistics*, 105(6): 1596–1605.

Clarke, James W. 1973. “Family structure and political socialization among urban black children.” *American Journal of Political Science*, 302–315.

Cohodes, Sarah, and James J Feigenbaum. 2021. “Why does education increase voting? Evidence from Boston’s charter schools.” National Bureau of Economic Research.

Coppock, Alexander, and Donald P Green. 2016. “Is voting habit forming? New evidence from experiments and regression discontinuities.” *American Journal of Political Science*, 60(4): 1044–1062.

Corvalan, Alejandro, and Paulo Cox. 2018. “The impact of procedural information costs on voting: Evidence from a natural experiment in chile.” *Political Behavior*, 40(1): 3–19.

Dahlgaard, Jens Olav. 2018. “Trickle-up political socialization: The causal effect on turnout of parenting a newly enfranchised voter.” *American Political Science Review*, 112(3): 698–705.

Dawson, Richard E, and Kenneth Prewitt. 1969. “Political socialization: An analytic study.” *(No Title)*.

De Haan, Monique. 2010. “Birth order, family size and educational attainment.” *Economics of Education Review*, 29(4): 576–588.

DellaVigna, Stefano, John A List, Ulrike Malmendier, and Gautam Rao. 2016. “Voting to tell others.” *The Review of Economic Studies*, 84(1): 143–181.

Dustan, Andrew. 2018. “Family networks and school choice.” *Journal of Development Economics*, 134: 372–391.

Dyck, Joshua J, and James G Gimpel. 2005. “Distance, turnout, and the convenience of voting.” *Social Science Quarterly*, 86(3): 531–548.

Easton, David, and Robert D Hess. 1962. “The child’s political world.” *Midwest Journal of Political Science*, 6(3): 229–246.

Ejrnæs, Mette, and Claus C Pörtner. 2004. “Birth order and the intrahousehold allocation of time and education.” *Review of Economics and Statistics*, 86(4): 1008–1019.

Feinberg, Mark E, Anna R Solmeyer, and Susan M McHale. 2012. “The third rail of family systems: Sibling relationships, mental and behavioral health, and preventive intervention in childhood and adolescence.” *Clinical child and family psychology review*, 15: 43–57.

Fraga, Bernard, and John Holbein. 2020. “Measuring youth and college student voter turnout.” *Electoral Studies*, 65: 102086.

Fujiwara, Thomas, Kyle Meng, and Tom Vogl. 2016. “Habit formation in voting: Evidence from rainy elections.” *American Economic Journal: Applied Economics*, 8(4): 160–188.

Gerber, Alan S, Donald P Green, and Christopher W Larimer. 2008. “Social pressure and voter turnout: Evidence from a large-scale field experiment.” *American political Science review*, 102(1): 33–48.

Gerber, Alan S, Donald P Green, and Ron Shachar. 2003. "Voting may be habit-forming: evidence from a randomized field experiment." *American journal of political science*, 47(3): 540–550.

Gerber, Alan S, Gregory A Huber, David Doherty, and Conor M Dowling. 2011. "The big five personality traits in the political arena." *Annual Review of Political Science*, 14(1): 265–287.

Gimpel, James G. 2003. "Cultivating democracy: Civic environments and political socialization in america." *Brookings Institute*.

Glynn, Adam N, and Maya Sen. 2015. "Identifying judicial empathy: Does having daughters cause judges to rule for women's issues?" *American Journal of Political Science*, 59(1): 37–54.

Goodman, Joshua, Michael Hurwitz, Jonathan Smith, and Julia Fox. 2015. "The relationship between siblings' college choices: Evidence from one million SAT-taking families." *Economics of Education Review*, 48: 75–85.

Green, Donald P, and Ron Shachar. 2000. "Habit formation and political behaviour: Evidence of consuetude in voter turnout." *British Journal of Political Science*, 30(4): 561–573.

Green, Donald P, Mary C McGrath, and Peter M Aronow. 2013. "Field experiments and the study of voter turnout." *Journal of Elections, Public Opinion and Parties*, 23(1): 27–48.

Green, Donald P, Oliver Hyman-Metzger, Gaurav Sood, and Michelle A Zee. 2023. "Revisiting a Natural Experiment: Do Legislators with Daughters Vote More Liberally on Women's Issues?" *Journal of Political Economy Microeconomics*, 1(3): 506–516.

Greenstein, Fred. 1965. "Children and Politics." In *Political Socialization*. , ed. Edward Greenburg, 56–63. New Haven, CT: Yale University Press.

Gurantz, Oded, Michael Hurwitz, and Jonathan Smith. 2020. "Sibling effects on high school exam taking and performance." *Journal of Economic Behavior & Organization*, 178: 534–549.

Harbaugh, William T. 1996. "If people vote because they like to, then why do so many of them lie?" *Public choice*, 89(1): 63–76.

Hendren, Nathaniel, and Ben Sprung-Keyser. 2020. "A unified welfare analysis of government policies." *The Quarterly journal of economics*, 135(3): 1209–1318.

Hersh, Eitan D. 2015. *Hacking the electorate: How campaigns perceive voters.* Cambridge University Press.

Hill, Seth J. 2020. "Following through on an intention to vote: Present bias and turnout." *Political Science Research and Methods*, 8(4): 803–810.

Holbein, John. 2016. "Left behind? Citizen responsiveness to government performance information." *American Political Science Review*, 110(2): 353–368.

Holbein, John B. 2017. "Childhood skill development and adult political participation." *American Political Science Review*, 111(3): 572–583.

Holbein, John B., and D. Sunshine Hillygus. 2020. *Making Young Voters: Converting Civic Attitudes Into Civic Action.* Cambridge University Press.

Holbein, John B, and Marcos A Rangel. 2020. "Does voting have upstream and downstream consequences? Regression discontinuity tests of the transformative voting hypothesis." *The Journal of Politics*, 82(4): 1196–1216.

Holbein, John B, D Sunshine Hillygus, Matthew A Lenard, Christina Gibson-Davis, and Darryl V Hill. 2020. "The development of students' engagement in school, community and democracy." *British Journal of Political Science*, 50(4): 1439–1457.

Holbein, John B, Marcos A Rangel, Raeal Moore, and Michelle Croft. 2023. "Is voting transformative? Expanding and meta-analyzing the evidence." *Political Behavior*, 45(3): 1015–1044.

Holbrook, Allyson L, and Jon A Krosnick. 2010. "Social desirability bias in voter turnout reports: Tests using the item count technique." *Public opinion quarterly*, 74(1): 37–67.

Humlum, Maria Knoth, Marius Opstrup Morthorst, and Peter Rønø Thingholm. 2024. "Sibling Spillovers and the Choice to Get Vaccinated: Evidence from a Regression Discontinuity Design." *Journal of Health Economics*, 94(102843).

Ifzaren, Djafra. 2018. "Sibling Gender Effects on Voting Behavior and Attitudes."

Jayachandran, Seema, and Rohini Pande. 2017. "Why are Indian children so short? The role of birth order and son preference." *American Economic Review*, 107(9): 2600–2629.

Jennings, M Kent, and Richard G Niemi. 1968. "The transmission of political values from parent to child." *American political science review*, 62(1): 169–184.

Jennings, M Kent, and Richard G Niemi. 2014. *Generations and politics: A panel study of young adults and their parents.* Vol. 68, Princeton University Press.

Jennings, M Kent, Laura Stoker, and Jake Bowers. 2009. "Politics across generations: Family transmission reexamined." *The Journal of Politics*, 71(3): 782–799.

Joensen, Juanna Schrøter, and Helena Skyt Nielsen. 2018. "Spillovers in education choice." *Journal of Public Economics*, 157: 158–183.

Karbownik, Krzysztof, and Umut Özek. 2023. "Setting a Good Example? Examining Sibling Spillovers in Educational Achievement Using a Regression Discontinuity Design." *Journal of Human Resources*, 58(5): 1567–1607.

Karp, Jeffrey A, and David Brockington. 2005. "Social desirability and response validity: A comparative analysis of overreporting voter turnout in five countries." *The Journal of Politics*, 67(3): 825–840.

Lahtinen, Hannu, Jani Erola, and Hanna Wass. 2019. "Sibling similarities and the importance of parental socioeconomic position in electoral participation." *Social Forces*, 98(2): 702–724.

Lane, Robert E. 1959. "Fathers and sons: Foundations of political belief." *American Sociological Review*, 502–511.

Levin, Martin L. 1961. "Social climates and political socialization." *The Public Opinion Quarterly*, 25(4): 596–606.

Linimon, Amy, and Mark R Joslyn. 2002. "Trickle up political socialization: The impact of kids voting USA on voter turnout in Kansas." *State Politics & Policy Quarterly*, 2(1): 24–36.

Martinelli, César. 2006. "Would rational voters acquire costly information?" *Journal of Economic Theory*, 129(1): 225–251.

McDevitt, Michael, and Steven Chaffee. 2000. "Closing gaps in political communication and knowledge: Effects of a school intervention." *Communication research*, 27(3): 259–292.

McDevitt, Michael, and Steven Chaffee. 2002. "From top-down to trickle-up influence: Revisiting assumptions about the family in political socialization." *Political Communication*, 19(3): 281–301.

Meredith, Marc, et al. 2009. "Persistence in political participation." *Quarterly Journal of Political Science*, 4(3): 187–209.

Mondak, Jeffery J. 2010. *Personality and the foundations of political behavior.* Cambridge University Press.

Nicoletti, Cheti, and Birgitta Rabe. 2019. "Sibling spillover effects in school achievement." *Journal of Applied Econometrics*, 34(4): 482–501.

Niemi, Richard G, and Barbara I Sobieszek. 1977. "Political socialization." *Annual review of sociology*, 3: 209–233.

Niemi, Richard G., and Mary A. Hepburn. 1995. "The rebirth of political socialization." *Perspectives on Political Science*, 24(1): 7–16.

Nyhan, Brendan, Christopher Skovron, and Rocío Titiunik. 2017. "Differential registration bias in voter file data: A sensitivity analysis approach." *American Journal of Political Science*, 61(3): 744–760.

Pettigrew, Stephen. 2021. "The downstream consequences of long waits: How lines at the precinct depress future turnout." *Electoral studies*, 71: 102188.

Plutzer, Eric. 2002. "Becoming a habitual voter: Inertia, resources, and growth in young adulthood." *American political science review*, 96(1): 41–56.

Porter, Catherine, and Danila Serra. 2020. "Gender differences in the choice of major: The importance of female role models." *American Economic Journal: Applied Economics*, 12(3): 226–254.

Qureshi, Javaeria A. 2018. "Siblings, teachers, and spillovers on academic achievement." *Journal of Human Resources*, 53(1): 272–297.

Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari C.W. Williams. 2025. "IPUMS USA: Version 16.0 ACS 2023."

Sacerdote, Bruce. 2011. "Peer effects in education: How might they work, how big are they and how much do we know thus far?" In *Handbook of the Economics of Education*. Vol. 3, 249–277. Elsevier.

Sandell, Julianna, and Eric Plutzer. 2005. "Families, divorce and voter turnout in the US." *Political Behavior*, 27: 133–162.

Sang, Samantha A, and Jackie A Nelson. 2017. "The effect of siblings on children's social skills and perspective taking." *Infant and Child Development*, 26(6): e2023.

Sapiro, Virginia. 2004. "Not Your Parents' Socialization: introduction for a new generation." *Annu. Rev. Polit. Sci.*, 7(1): 1–23.

Schafer, Jerome. 2021. "Delayed gratification in political participation." *American Politics Research*, 49(3): 304–312.

Schafer, Jerome, and John B Holbein. 2020. "When time is of the essence: A natural experiment on how time constraints influence elections." *The Journal of Politics*, 82(2): 418–432.

Searing, Donald D, Joel J Schwartz, and Alden E Lind. 1973. "The structuring principle: Political socialization and belief systems." *American Political Science Review*, 67(2): 415–432.

Sears, David O, and Carolyn L Funk. 1999. "Evidence of the long-term persistence of adults' political predispositions." *The Journal of Politics*, 61(1): 1–28.

Simon, James, and Bruce D Merrill. 1998. "Political socialization in the classroom revisited: The Kids Voting program." *The Social Science Journal*, 35(1): 29–42.

Tedin, Kent L. 1974. "The influence of parents on the political attitudes of adolescents." *American Political Science Review*, 68(4): 1579–1592.

Torney-Purta, Judith V. 2017. *The development of political attitudes in children.* Routledge.

Urbatsch, Robert. 2011. "Sibling ideological influence: A natural experiment." *British Journal of Political Science*, 41(4): 693–712.

Urbatsch, Robert. 2014. *Families' values: How parents, siblings, and children affect political attitudes.* Oxford University Press, USA.

Verba, Sidney, Kay Lehman Schlozman, and Henry E Brady. 1995. *Voice and equality: Civic voluntarism in American politics.* Harvard University Press.

Verba, Sidney, Kay Lehman Schlozman, and Nancy Burns. 2005. "Family Ties." *Social Logic Of Politics: Personal Networs As Contexts*, 95.

Washington, Ebonya L. 2008. "Female socialization: How daughters affect their legislator fathers' voting on women's issues." *American Economic Review*, 98(1): 311–332.

Weinschenk, Aaron C. 2017. "Big five personality traits, political participation, and civic engagement: Evidence from 24 countries." *Social Science Quarterly*, 98(5): 1406–1421.

Weisberg, Herbert F. 1987. "The demographics of a new voting gap marital differences in American voting." *Public Opinion Quarterly*, 51(3): 335–343.

Whiteman, Shawn D, and Abigail Christiansen. 2008. "Processes of sibling influence in adolescence: Individual and family correlates." *Family Relations*, 57(1): 24–34.

Zang, Emma, Poh Lin Tan, and Philip J Cook. 2023. “Sibling spillovers: Having an academically successful older sibling may be more important for children in disadvantaged families.” *American Journal of Sociology*, 128(5): 1529–1571.

Appendix A - Tables and Figures

TABLE A1—DESCRIPTIVE RELATIONSHIP BETWEEN BOTH SIBLINGS VOTING AND FAMILY CHARACTERISTICS.

	Outcome = Both siblings voted in younger sibling's 1st election				
	(1)	(2)	(3)	(4)	(5)
Both siblings female (reference)					
Both siblings male	0.010***			0.009***	0.009***
	(0.000)			(0.000)	(0.000)
Older sibling female, younger sibling male	-0.004***			-0.005***	-0.004***
	(0.000)			(0.000)	(0.000)
Older sibling male, younger sibling female	0.005***			0.004***	0.003***
	(0.000)			(0.000)	(0.000)
Sibling's average PSAT score		0.001***		0.000***	0.000***
		(0.000)		(0.000)	(0.000)
Parental educ: no college (reference)					
Parental educ: college but no BA degree			0.012***	0.011***	0.010***
			(0.001)	(0.001)	(0.001)
Parental educ: college with BA degree			0.024***	0.020***	0.018***
			(0.001)	(0.001)	(0.001)
Parental income: less than $50,000 (reference)					
Parental income: $50,000 - $100,000			0.009***	0.008***	0.008***
			(0.001)	(0.001)	(0.001)
Parental income: greater than $100,000			0.006***	0.004***	0.003***
			(0.001)	(0.001)	(0.001)
High school fixed effects					✓
Number of families	2,626,286	2,626,286	2,626,286	2,626,286	2,626,286

Notes: The sample includes all families with 2 children. Regressions that include siblings' average PSAT scores, parental education, and parental income also include indicators for whether those variables are missing for the family.

TABLE A2—DESCRIPTIVE RELATIONSHIP BETWEEN SIBLING VOTING PATTERNS.

	(1)	(2)	(3)	(4)
Panel A: Outcome = younger sibling voted in 1st election				
Full sample, older sibling's 1st election is presidential				
Older sibling voted in 1st election	0.102***	0.102***	0.100***	0.092***
	(0.001)	(0.001)	(0.001)	(0.001)
Number of sibling pairs	1,284,032	1,284,032	1,284,032	1,281,545
Full sample, older sibling's 1st election is presidential				
Older sibling voted in younger sibling's 1st election	0.167***	0.167***	0.164***	0.158***
	(0.001)	(0.001)	(0.001)	(0.001)
Number of sibling pairs	1,284,032	1,284,032	1,284,032	1,281,545
Older sibling RD sample, 18th birthdays near any election				
Older sibling voted in election closest to 18th birthday	0.101***	0.101***	0.097***	0.089***
	(0.003)	(0.003)	(0.003)	(0.003)
Number of sibling pairs	202,931	202,931	202,931	199,269
Older sibling RD sample, 18th birthdays near presidential election				
Older sibling voted in election closest to 18th birthday	0.097***	0.097***	0.094***	0.086***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of sibling pairs	112,770	112,770	112,770	108,893
Panel B: Outcome = older sibling voted in younger sibling's 1st election				
Full sample				
Younger sibling voted in 1st election	0.189***	0.189***	0.186***	0.178***
	(0.001)	(0.001)	(0.001)	(0.001)
Number of sibling pairs	2,575,240	2,575,240	2,575,240	2,573,041
Full sample, younger sibling's 1st election is presidential				
Younger sibling voted in 1st election	0.183***	0.183***	0.180***	0.168***
	(0.001)	(0.001)	(0.001)	(0.001)
Number of sibling pairs	1,343,926	1,343,926	1,343,926	1,341,584
Younger sibling RD sample, 18th birthdays near any election				
Younger sibling voted in election closest to 18th birthday	0.149***	0.149***	0.144***	0.138***
	(0.004)	(0.004)	(0.004)	(0.004)
Number of sibling pairs	165,734	165,734	165,734	161,876
Younger sibling RD sample, 18th birthdays near presidential election				
Younger sibling voted in election closest to 18th birthday	0.151***	0.151***	0.145***	0.140***
	(0.004)	(0.004)	(0.004)	(0.005)
Number of sibling pairs	90,997	90,997	90,996	86,857
Election year by state fixed effects	✓	✓	✓	✓
Sibling demographics		✓	✓	✓
Sibling PSAT scores			✓	✓
Younger sibling high school fixed effects				✓

Notes: The full sample includes all sibling pairs that include the oldest child in the family with age differences between 9 and 48 months. Robust standard errors are reported.

TABLE A3—RD SAMPLE COVARIATE BALANCE CHECKS.

	Sample		
	Own-own all sibling RD sample (1)	Older sibling RD sample (2)	Younger sibling RD sample (3)
Own characteristics			
Female	0.006	0.001	0.005
	(0.004)	(0.006)	(0.007)
PSAT score	0.278*	0.163	0.300
	(0.166)	(0.268)	(0.294)
White	-0.002	0.004	-0.008
	(0.003)	(0.006)	(0.006)
Black	-0.002	-0.006*	-0.002
	(0.002)	(0.003)	(0.004)
Asian	0.002	-0.002	0.006*
	(0.002)	(0.003)	(0.004)
Hispanic	0.003	0.007*	0.002
	(0.003)	(0.004)	(0.004)
Sibling characteristics			
Female		0.004	-0.003
		(0.006)	(0.007)
PSAT score		0.053	0.053
		(0.266)	(0.298)
White		0.004	-0.011*
		(0.006)	(0.006)
Black		-0.006*	-0.002
		(0.003)	(0.004)
Asian		-0.001	0.006*
		(0.003)	(0.004)
Hispanic		0.005	0.003
		(0.004)	(0.004)
Sibling age difference (months)		-0.062	0.121
		(0.080)	(0.129)
Age when 1st eligible to vote		0.003	-0.005
		(0.005)	(0.007)

Notes: This table reports covariance balance tests for our three main RD samples. To carry out these tests, we use the same reduced form RD specifications in Equation 2 and Equation 4 while using individual or sibling characteristics as the outcome. Each coefficient comes from a separate regression. For the older sibling RD sample, the sibling characteristics correspond to the younger sibling. Likewise, for the younger sibling RD sample, the sibling characteristics correspond to the older sibling. The sample in column 1 includes all siblings with an 18th birthday within 30 days of a presidential election. The sample in column 2 includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near a presidential election. The sample includes sibling pairs from families with 2 children with age differences between 9 and 48 months and where the younger sibling turned 18 near a presidential election. In all cases, bandwidths are 30 days. Robust standard errors are reported in parentheses.

TABLE A4—ROBUSTNESS CHECKS FOR OWN FUTURE VOTING RD, ALL SIBLING SAMPLE.

	BW = 15 days (1)	BW = 30 days (2)	BW = 30 days; triangle weight (3)	BW = 30 days; quadratic (4)	BW = 60 days (5)	Optimal BW (6)
Outcome = Voted in election 2 years later						
Eligible to vote at age 18	0.020***	0.021***	0.020***	0.020***	0.021***	0.022***
	(0.003)	(0.002)	(0.003)	(0.003)	(0.002)	(0.002)
						{52}
Observations	144,416	286,614	286,614	286,614	584,875	507,346
Outcome = Voted in election 4 years later						
Eligible to vote at age 18	0.015***	0.016***	0.015***	0.015***	0.017***	0.017***
	(0.005)	(0.003)	(0.005)	(0.005)	(0.002)	(0.003)
						{55}
Observations	144,416	286,614	286,614	286,614	584,875	548,759
Outcome = Ever voted in the future						
Eligible to vote at age 18	0.016***	0.014***	0.014***	0.017***	0.016***	0.015***
	(0.005)	(0.004)	(0.005)	(0.006)	(0.003)	(0.003)
						{53}
Observations	144,416	286,614	286,614	286,614	584,875	528,471
Outcome = Number of future votes						
Eligible to vote at age 18	0.059***	0.057***	0.063***	0.065***	0.055***	0.056***
	(0.013)	(0.009)	(0.013)	(0.014)	(0.006)	(0.007)
						{56}
Observations	144,416	286,614	286,614	286,614	584,875	548,759

Notes: The sample includes all siblings with an 18th birthday within 30 days of a presidential election. Regressions include fixed effects for the election year that is closest to their 18th birthday. Column 3 includes a triangular weight. Column 4 includes second order terms of the running variable on each side of the threshold. Column 6 uses the 'rdrobust' package in Stata to compute estimates with the 'optimal' bandwidths, which are reported in curly brackets. Robust standard errors are reported in parentheses.

TABLE A5—EFFECT OF ELIGIBILITY AT AGE 18 ON OWN FUTURE VOTING, FULL SAMPLE.

	First stage	Own future voting outcomes			
	Voted at age 18 (1)	Voted in election 2 years later (2)	Voted in election 4 years later (3)	Ever voted in the future (4)	Total future votes (5)
Panel A: 18th birthdays near any election					
Eligible to vote at age 18	0.178***	0.014***	0.009***	0.014***	0.042***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.003)
Outcome control mean	0.005	0.144	0.201	0.412	0.798
Observations	1,864,816	1,864,816	1,864,816	1,864,816	1,864,816
Panel B: 18th birthdays near presidential elections					
Eligible to vote at age 18	0.253***	0.018***	0.015***	0.018***	0.053***
	(0.001)	(0.001)	(0.002)	(0.002)	(0.004)
Outcome control mean	0.006	0.071	0.276	0.399	0.753
Observations	1,098,813	1,098,813	1,098,813	1,098,813	1,098,813
Panel C: 18th birthdays near midterm elections					
Eligible to vote at age 18	0.069***	0.008***	-0.000	0.008***	0.025***
	(0.001)	(0.002)	(0.001)	(0.002)	(0.006)
Outcome control mean	0.005	0.250	0.092	0.431	0.861
Observations	766,003	766,003	766,003	766,003	766,003

Notes: The sample includes all individuals with an 18th birthday within 30 days of an election. Regressions include fixed effects for the election year that is closest to their 18th birthday. Bandwidth is 30 days. Robust standard errors are reported.

TABLE A6—EFFECT OF ELIGIBILITY AT AGE 18 ON OWN FUTURE VOTING, OLDER SIBLINGS SAMPLE.

	First stage	Own future voting outcomes			
	Voted at age 18 (1)	Voted in election 2 years later (2)	Voted in election 4 years later (3)	Ever voted in the future (4)	Total future votes (5)
Panel A: 18th birthdays near any election					
Eligible to vote at age 18	0.207***	0.015***	0.012***	0.011**	0.054***
	(0.003)	(0.003)	(0.004)	(0.004)	(0.013)
Outcome control mean	0.008	0.183	0.213	0.512	1.134
Observations	202,937	202,937	202,937	202,937	202,937
Panel B: 18th birthdays near presidential elections					
Eligible to vote at age 18	0.295***	0.027***	0.018***	0.014**	0.074***
	(0.004)	(0.004)	(0.005)	(0.006)	(0.018)
Outcome control mean	0.009	0.090	0.288	0.506	1.149
Observations	112,778	112,778	112,778	112,778	112,778
Panel C: 18th birthdays near midterm elections					
Eligible to vote at age 18	0.097***	0.001	0.003	0.007	0.029
	(0.003)	(0.006)	(0.004)	(0.007)	(0.018)
Outcome control mean	0.007	0.300	0.120	0.521	1.114
Observations	90,159	90,159	90,159	90,159	90,159

Notes: The sample includes the older siblings from sibling pairs that include the oldest child in the family with age differences between 9 and 48 months. Regressions include fixed effects for the election year that is closest to the older sibling's 18th birthday. Bandwidth is 30 days. Robust standard errors are reported.

TABLE A7—ROBUSTNESS CHECKS FOR OLDER SIBLING RD.

	BW = 15 days (1)	BW = 30 days (2)	BW = 30 days; triangle weight (3)	BW = 30 days; quadratic (4)	BW = 60 days (5)	Optimal BW (6)
Outcome = Younger sibling voted in 1st election						
Eligible to vote at age 18	0.011	0.010**	0.012*	0.011	0.006*	0.011**
	(0.007)	(0.005)	(0.007)	(0.007)	(0.003)	(0.005)
						{34}
Observations	56,551	112,778	112,778	112,778	230,094	127,004
Outcome = Younger sibling voted in 1st presidential election						
Eligible to vote at age 18	0.010	0.011**	0.014*	0.012	0.005	0.008*
	(0.008)	(0.005)	(0.008)	(0.008)	(0.004)	(0.005)
						{41}
Observations	56,551	112,778	112,778	112,778	230,094	158,745
Outcome = Younger sibling voted in 1st midterm election						
Eligible to vote at age 18	-0.003	-0.001	-0.002	-0.005	-0.001	0.000
	(0.005)	(0.004)	(0.005)	(0.005)	(0.003)	(0.003)
						{39}
Observations	56,548	112,772	112,772	112,772	230,083	146,375
Outcome = Younger sibling ever voted						
Eligible to vote at age 18	-0.002	0.004	-0.000	-0.001	0.005	0.009*
	(0.008)	(0.006)	(0.008)	(0.009)	(0.004)	(0.005)
						{47}
Observations	56,551	112,778	112,778	112,778	230,094	183,072
Outcome = Younger sibling's number of votes cast						
Eligible to vote at age 18	0.017	0.032**	0.025	0.021	0.016	0.031**
	(0.023)	(0.016)	(0.023)	(0.024)	(0.011)	(0.013)
						{45}
Observations	56,551	112,778	112,778	112,778	230,094	174,539
Outcome = Younger sibling voted in 2+ elections						
Eligible to vote at age 18	0.009	0.014**	0.014*	0.010	0.008**	0.009**
	(0.008)	(0.005)	(0.008)	(0.008)	(0.004)	(0.004)
						{63}
Observations	56,551	112,778	112,778	112,778	230,094	239,327
Outcome = Younger sibling total votes after 1st election						
Eligible to vote at age 18	0.005	0.022*	0.013	0.010	0.011	0.019
	(0.019)	(0.013)	(0.018)	(0.020)	(0.009)	(0.012)
						{39}
Observations	56,551	112,778	112,778	112,778	230,094	150,557

Notes: The sample includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year the younger sibling is first eligible for. Column 3 includes a triangular weight. Column 4 includes second order terms of the running variable on each side of the threshold. Column 6 uses the 'rdrobust' package in Stata to compute estimates with the 'optimal' bandwidths, which are reported in curly brackets (Calonico et al., 2017). Robust standard errors are reported in parentheses.

TABLE A8—OLDER SIBLING RD PLACEBO CHECK, 18TH BIRTHDAYS NEAR ODD YEAR SPECIAL ELECTION DATES.

	Younger sibling voting outcomes								
	Voted in 1st election (1)	Voted in 1st pres. election (2)	Voted in 1st midterm election (3)	Ever registered (4)	Ever voted (5)	Voted in 2+ elections (6)	Voted in 3+ elections (7)	Total votes cast (8)	Total votes after 1st election (9)
Older sibling "eligible" at age 18 in odd year	-0.005	-0.002	-0.004	0.001	-0.001	-0.001	0.001	-0.009	-0.003
	(0.004)	(0.004)	(0.003)	(0.004)	(0.005)	(0.004)	(0.003)	(0.014)	(0.009)
Outcome control mean	0.204	0.306	0.096	0.674	0.482	0.247	0.108	1.053	0.684
Observations	173,803	173,803	173,797	173,803	173,803	173,803	173,803	173,803	173,803

Notes: The sample includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months. The running variable in these RD specifications are defined as the number of days between the older sibling's 18th birthday and the nearest odd year election date, using the 'special election' dates in those years (2005, 2007, 2009, 2011). Regressions include fixed effects for the election year the younger sibling is first eligible for. Bandwidth is 30 days. Robust standard errors are reported.

TABLE A9—OLDER SIBLING RD, 18TH BIRTHDAYS NEAR ANY ELECTION.

	1st stage	Younger sibling voting outcomes								
	Older sibling voted at age 18 (1)	Voted in 1st election (2)	Voted in 1st pres. election (3)	Voted in 1st midterm election (4)	Ever registered (5)	Ever voted (6)	Voted in 2+ elections (7)	Voted in 3+ elections (8)	Total votes (9)	Total votes after 1st election (10)
Older sibling eligible to vote at age 18 (RF)	0.206***	0.001	0.005	-0.003	0.001	-0.002	0.006	0.002	0.008	0.008
	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.004)	(0.003)	(0.011)	(0.009)
	[6,262]									
Older sibling voted at age 18 (IV)		0.003	0.022	-0.013	0.005	-0.009	0.031*	0.012	0.040	0.036
		(0.017)	(0.020)	(0.013)	(0.020)	(0.021)	(0.019)	(0.014)	(0.054)	(0.044)
Outcome control mean	0.008	0.209	0.304	0.097	0.671	0.493	0.269	0.131	0.968	0.759
Observations	202,937	202,937	202,937	202,931	202,937	202,937	202,937	202,937	202,937	202,937

Notes: The sample includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near the 2004, 2006, 2008, 2010, or 2012 elections. Regressions include fixed effects for the election year the younger sibling is first eligible for. Bandwidth is 30 days. First stage F-stat reported in brackets in column 1. Robust standard errors are reported in parentheses.

TABLE A10—OLDER SIBLING RD BY SIBLING SEX COMPOSITION.

	1st stage	Younger sibling voting outcomes								
	Older sibling voted at age 18 (1)	Voted in 1st election (2)	Voted in 1st pres. election (3)	Voted in 1st midterm election (4)	Ever registered (5)	Ever voted (6)	Voted in 2+ elections (7)	Voted in 3+ elections (8)	Total votes (9)	Total votes after 1st election (10)
Panel A: Sisters										
Older sibling eligible	0.277***	0.029***	0.039***	0.003	0.016	0.031***	0.039***	0.022***	0.099***	0.070***
to vote at age 18 (RF)	(0.007)	(0.009)	(0.010)	(0.007)	(0.011)	(0.011)	(0.010)	(0.008)	(0.030)	(0.024)
	[1,485]									
Older sibling voted		0.105***	0.141***	0.012	0.058	0.111***	0.141***	0.079***	0.356***	0.252***
at age 18 (IV)		(0.033)	(0.037)	(0.024)	(0.039)	(0.041)	(0.037)	(0.029)	(0.107)	(0.087)
Outcome control mean	0.004	0.220	0.291	0.092	0.635	0.486	0.287	0.145	1.003	0.783
Observations	31,014	31,014	31,014	31,012	31,014	31,014	31,014	31,014	31,014	31,014
Panel B: Brothers										
Older sibling eligible	0.321***	0.014	0.013	-0.003	-0.010	-0.008	0.014	0.005	0.027	0.013
to vote at age 18 (RF)	(0.008)	(0.010)	(0.011)	(0.007)	(0.011)	(0.012)	(0.011)	(0.009)	(0.032)	(0.027)
	[1,457]									
Older sibling voted		0.045	0.041	-0.009	-0.031	-0.024	0.043	0.014	0.084	0.039
at age 18 (IV)		(0.031)	(0.035)	(0.023)	(0.035)	(0.038)	(0.034)	(0.027)	(0.101)	(0.083)
Outcome control mean	0.013	0.227	0.282	0.101	0.701	0.509	0.289	0.150	1.041	0.814
Observations	26,638	26,638	26,638	26,636	26,638	26,638	26,638	26,638	26,638	26,638
Panel C: Older sibing female, younger sibling male										
Older sibling eligible	0.276***	-0.002	0.004	-0.003	0.006	0.000	-0.001	-0.003	0.001	0.002
to vote at age 18 (RF)	(0.008)	(0.010)	(0.011)	(0.008)	(0.011)	(0.012)	(0.011)	(0.009)	(0.033)	(0.027)
	[1,296]									
Older sibling voted		-0.006	0.015	-0.010	0.023	0.001	-0.002	-0.012	0.003	0.008
at age 18 (IV)		(0.036)	(0.040)	(0.027)	(0.040)	(0.044)	(0.040)	(0.032)	(0.119)	(0.098)
Outcome control mean	0.004	0.238	0.294	0.109	0.695	0.515	0.304	0.163	1.086	0.848
Observations	27,566	27,566	27,566	27,566	27,566	27,566	27,566	27,566	27,566	27,566
Panel D: Older sibing male, younger sibling female										
Older sibling eligible	0.309***	-0.003	-0.014	-0.003	0.000	-0.009	0.001	0.005	-0.003	0.001
to vote at age 18 (RF)	(0.008)	(0.010)	(0.011)	(0.007)	(0.011)	(0.012)	(0.011)	(0.008)	(0.032)	(0.026)
	[1,445]									
Older sibling voted		-0.011	-0.046	-0.008	0.002	-0.028	0.002	0.017	-0.009	0.002
at age 18 (IV)		(0.032)	(0.036)	(0.023)	(0.037)	(0.039)	(0.035)	(0.027)	(0.103)	(0.085)
Outcome control mean	0.015	0.232	0.303	0.096	0.642	0.494	0.297	0.150	1.031	0.799
Observations	27,558	27,558	27,558	27,558	27,558	27,558	27,558	27,558	27,558	27,558

Notes: The sample includes sibling pairs that include the oldest child in the family with age differences between 9 and 48 months and where the older sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year the younger sibling is first eligible for. Bandwidth is 30 days. First stage F-stat reported in brackets in column 1. Robust standard errors are reported in parentheses.

TABLE A11—ROBUSTNESS CHECKS FOR YOUNGER SIBLING RD.

	BW = 15 days (1)	BW = 30 days (2)	BW = 30 days; triangle weight (3)	BW = 30 days; quadratic (4)	BW = 30 days; donut (5)	BW = 60 days (6)	Optimal BW (7)
Younger sibling's 18th birthday near presidential election							
Eligible to vote at age 18	0.021**	0.013**	0.027***	0.030***	0.011*	0.012***	0.012***
	(0.009)	(0.006)	(0.009)	(0.009)	(0.006)	(0.004)	(0.005)
							{52}
Observations	45,930	90,997	90,997	90,997	89,901	185,525	162,233
Younger sibling's 18th birthday near any election							
Eligible to vote at age 18	0.015***	0.009**	0.018***	0.018***	0.007*	0.008***	0.008**
	(0.005)	(0.004)	(0.006)	(0.005)	(0.004)	(0.003)	(0.003)
							{46}
Observations	84,069	165,751	165,751	165,751	163,276	336,657	255,744

Notes: The sample includes sibling pairs from families with 2 children with age differences between 9 and 48 months and where the younger sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year the younger sibling is first eligible for. Column 3 includes a triangular weight. Column 4 includes second order terms of the running variable on each side of the threshold. Column 6 uses the 'rdrobust' package in Stata to compute estimates with the 'optimal' bandwidths, which are reported in curly brackets (Calonico et al., 2017). Robust standard errors are reported in parentheses.

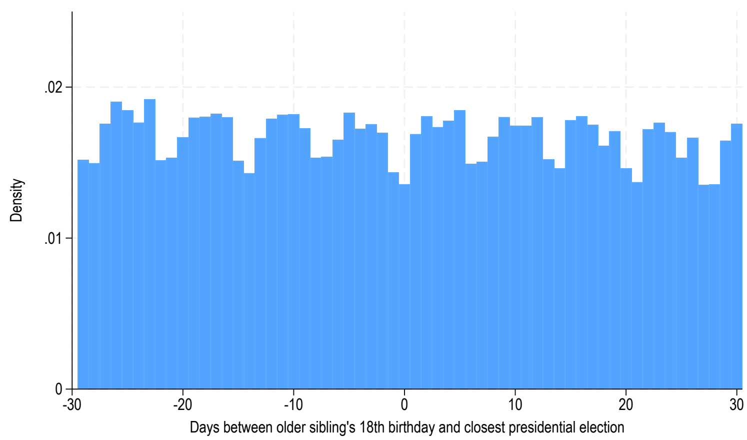
TABLE A12—YOUNGER SIBLING RD BY SIBLING SEX COMPOSITION.

	Younger sibling voted in election closest to 18th birthday	Older sibling voted in election closest to younger sibling's 18th birthday	
	First stage	Reduced Form	Second Stage
	(1)	(2)	(3)
Panel A: Sisters			
RD estimate	0.339***	0.006	0.017
	(0.009)	(0.011)	(0.034)
First stage F-stat	[1,581]		
Outcome control mean	0.003	0.289	0.289
Observations	24,577	24,577	24,577
Panel B: Brothers			
RD estimate	0.317***	0.011	0.033
	(0.009)	(0.012)	(0.039)
First stage F-stat	[1,235]		
Outcome control mean	0.003	0.291	0.291
Observations	21,384	21,384	21,384
Panel C: Older sibling female, younger sibling male			
RD estimate	0.327***	0.033***	0.100***
	(0.009)	(0.012)	(0.037)
First stage F-stat	[1,281]		
Outcome control mean	0.012	0.296	0.296
Observations	22,414	22,414	22,414
Panel D: Older sibling male, younger sibling female			
RD estimate	0.344***	0.005	0.015
	(0.009)	(0.012)	(0.035)
First stage F-stat	[1,477]		
Outcome control mean	0.005	0.293	0.293
Observations	22,622	22,622	22,622

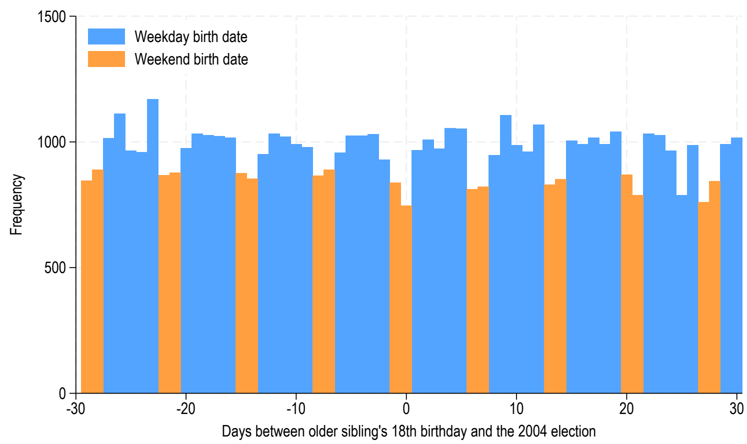
Notes: The sample includes sibling pairs from families with 2 children with age differences between 9 and 48 months and where the younger sibling turned 18 near the 2004, 2008, or 2012 elections. Regressions include fixed effects for the election year that is closest to the younger sibling's 18th birthday. Bandwidth is 30 days. Robust standard errors are reported.

FIGURE A1. DENSITY OF OLDER SIBLING'S 18TH BIRTHDAY RELATIVE TO PRESIDENTIAL ELECTIONS

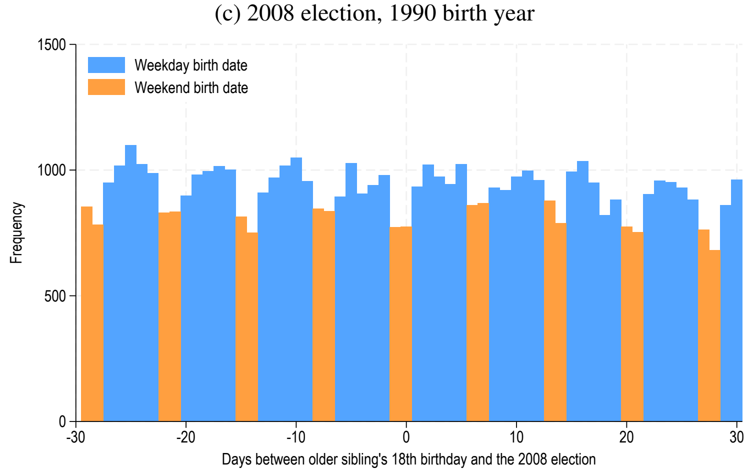
(a) All presidential elections



(b) 2004 election, 1986 birth year



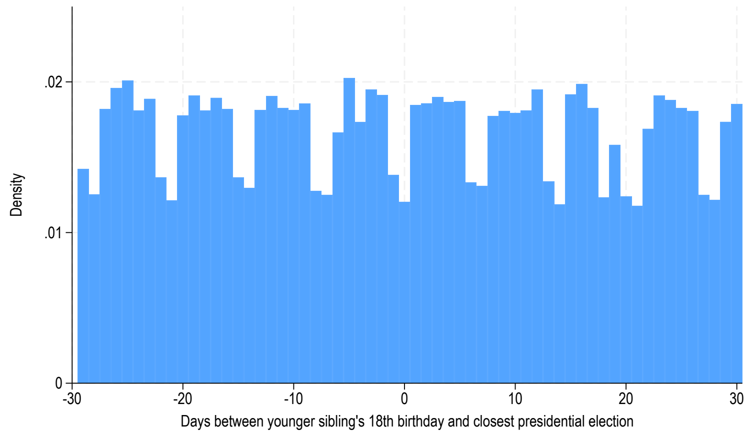
(c) 2008 election, 1990 birth year



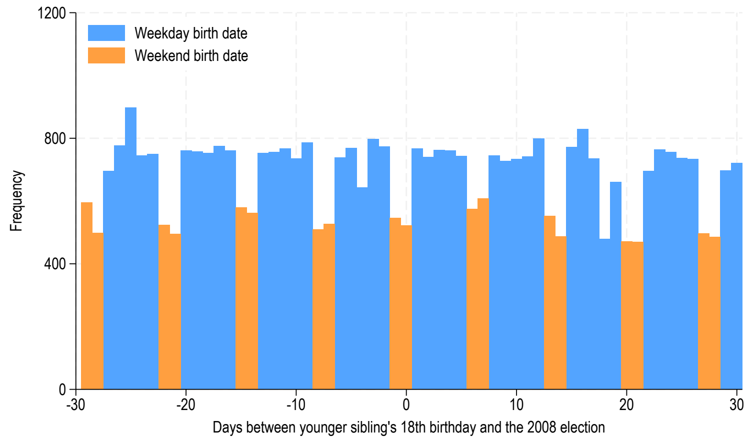
Note: This figure shows histograms of the running variable for the older sibling RD design. Panel (a) shows the histogram for all observations. Panels (b) and (c) show the histograms specifically for the 2004 and 2008 elections to show how the lumpiness of the densities are explained by fewer births on weekends.

FIGURE A2. DENSITY OF YOUNGER SIBLING'S 18TH BIRTHDAY RELATIVE TO PRESIDENTIAL ELECTIONS

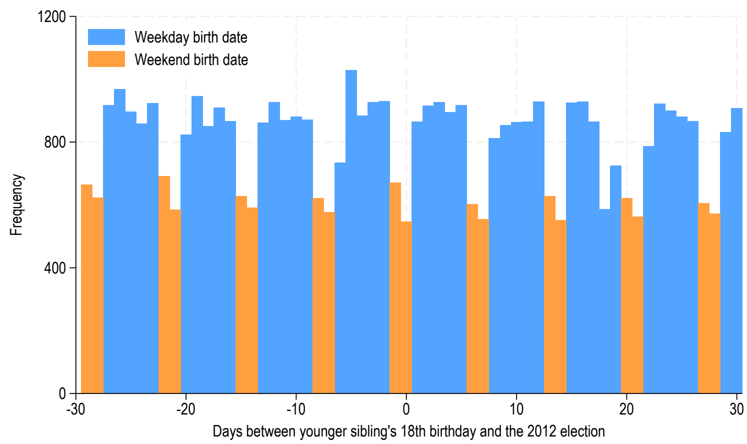
(a) All presidential elections



(b) 2008 election, 1990 birth year



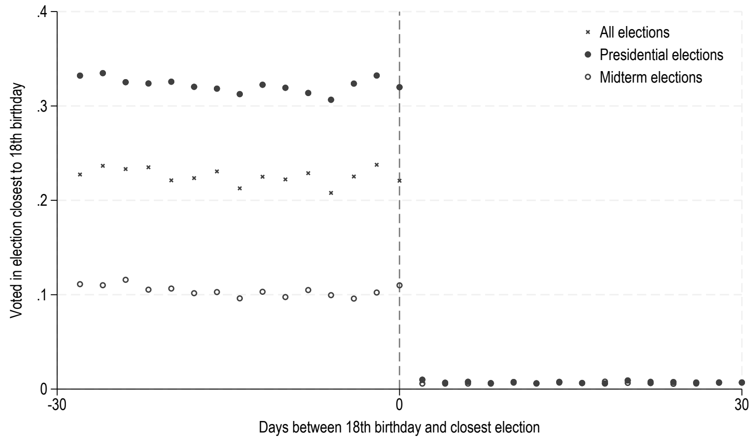
(c) 2012 election, 1994 birth year



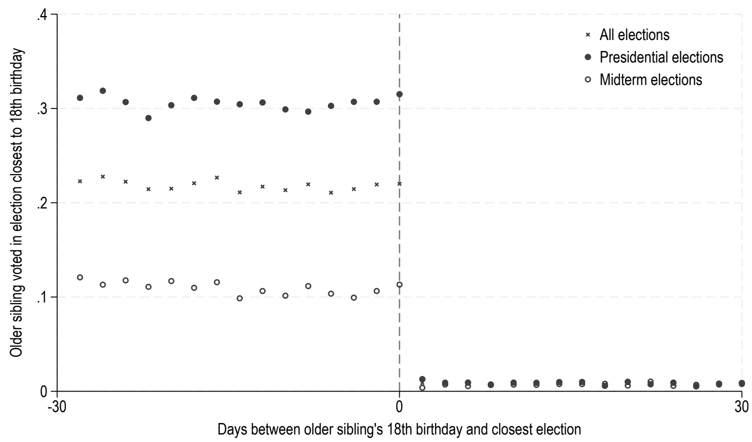
Note: This figure shows histograms of the running variable for the younger sibling RD design. Panel (a) shows the histogram for all observations. Panels (b) and (c) show the histograms specifically for the 2008 and 2012 elections to show how the lumpiness of the densities are explained by fewer births on weekends.

FIGURE A3. FIRST STAGE

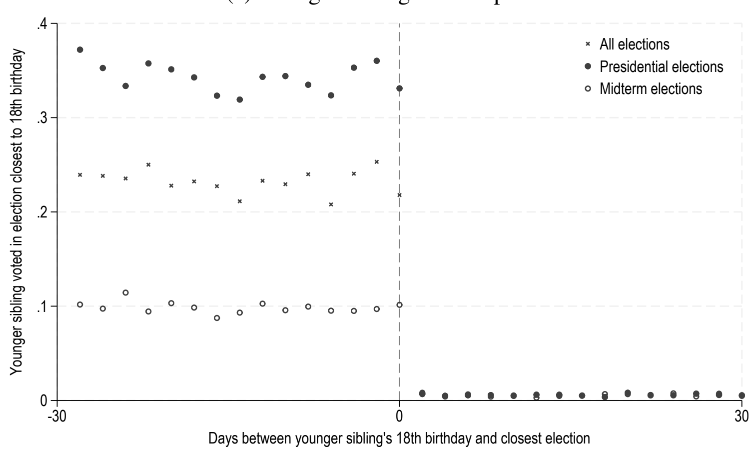
(a) All siblings RD sample



(b) Older sibling RD sample



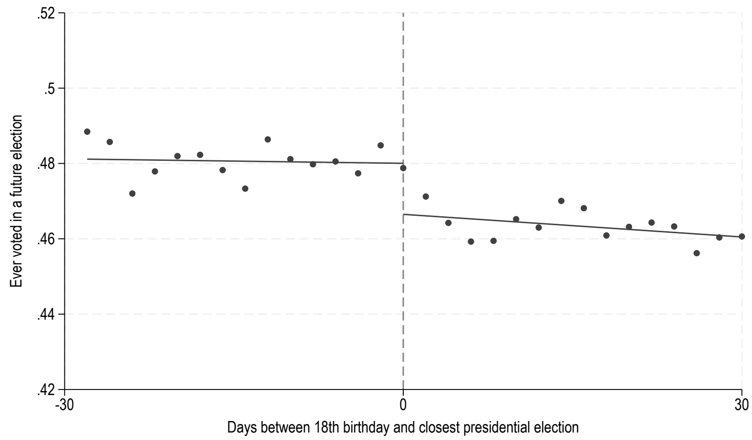
(c) Younger sibling RD sample



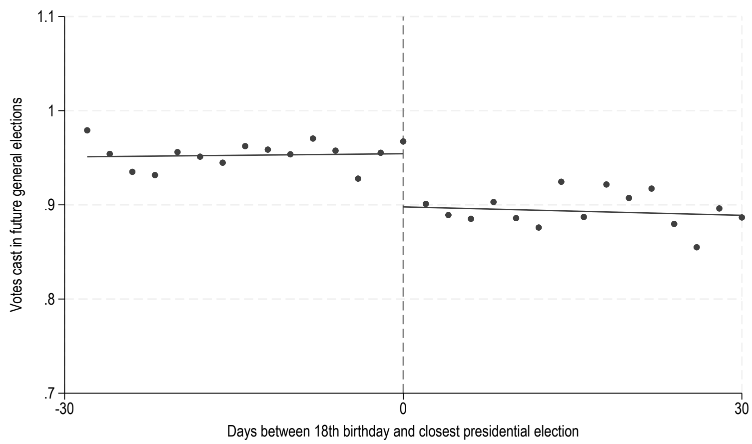
Note: This figure shows binned scatterplots of voting rates in the election closest to 18th birthdays by days between that birthday and the presidential elections of 2004, 2008, and 2012 for our main RD samples. Each bin consists of 2 days.

FIGURE A4. THE EFFECT OF AGE 18 ELIGIBILITY ON FUTURE VOTING - ADDITIONAL OUTCOMES

(a) Ever voted in future election



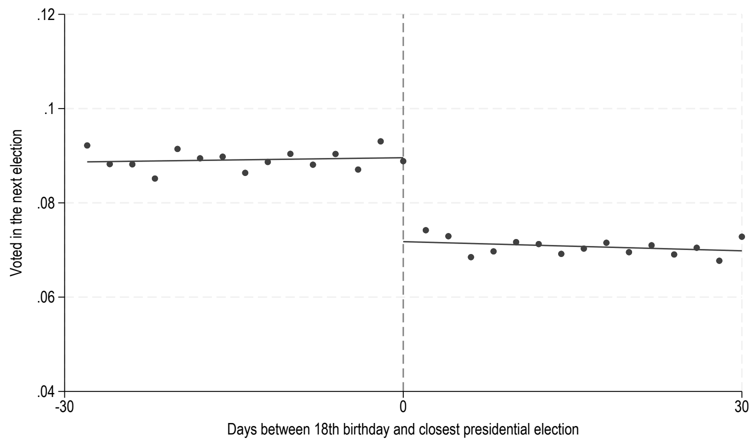
(b) Votes cast in future elections



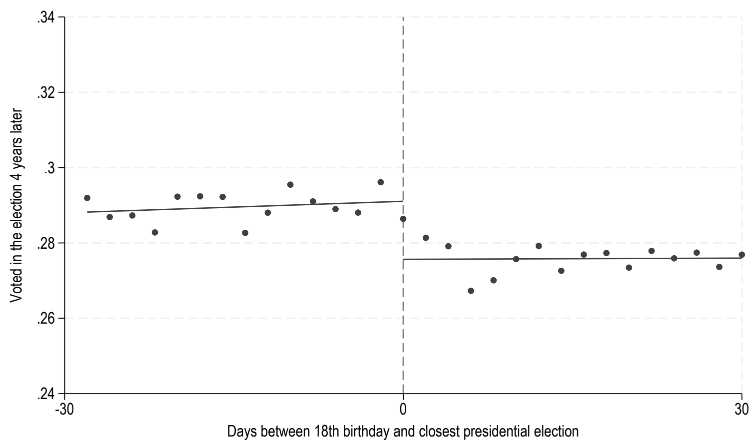
Note: This figure shows binned scatterplots of average voting outcomes by days between each sibling's 18th birthday and the presidential elections of 2004, 2008, and 2012. Each bin consists of 2 days.

FIGURE A5. THE EFFECT OF AGE 18 ELIGIBILITY ON FUTURE VOTING - FULL SAMPLE

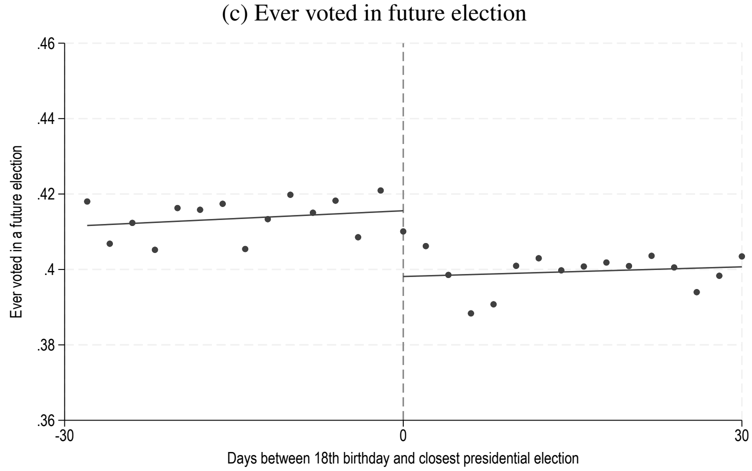
(a) Voted in election 2 years later



(b) Voted in election 4 years later



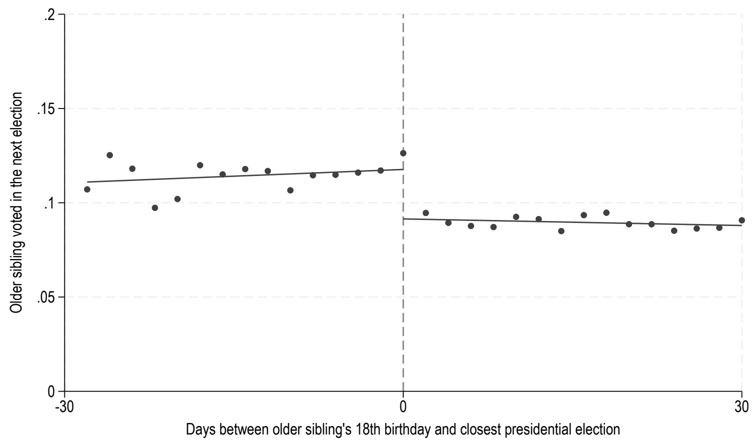
(c) Ever voted in future election



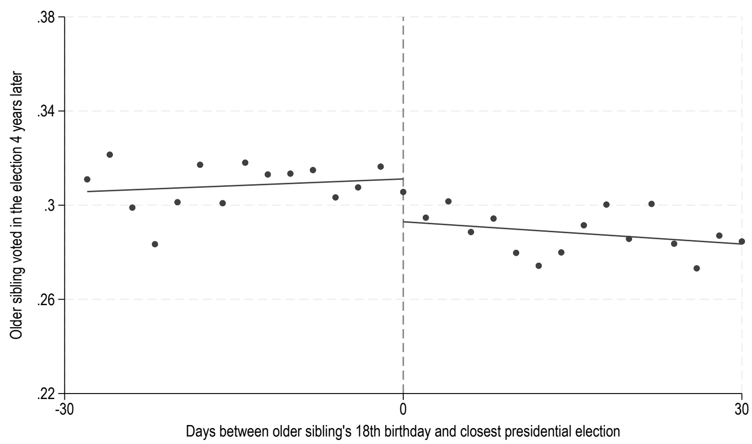
Note: This figure shows binned scatterplots of average voting outcomes by days between an individual's 18th birthday and the presidential elections of 2004, 2008, and 2012. Each bin consists of 2 days.

FIGURE A6. THE EFFECT OF AGE 18 ELIGIBILITY ON FUTURE VOTING - OLDER SIBLINGS

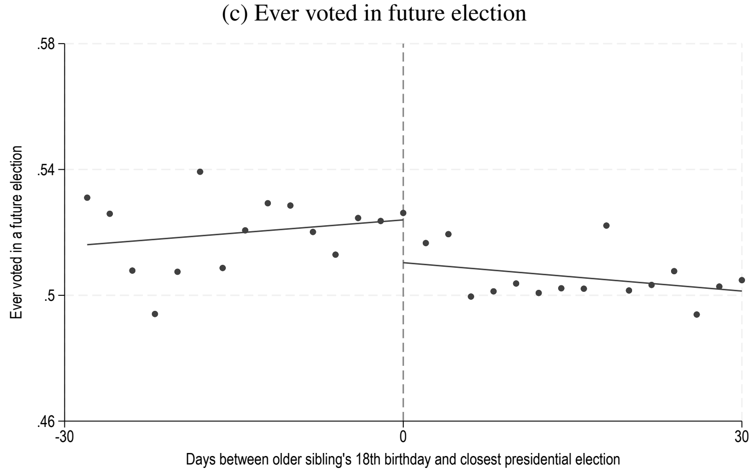
(a) Voted in election 2 years later



(b) Voted in election 4 years later



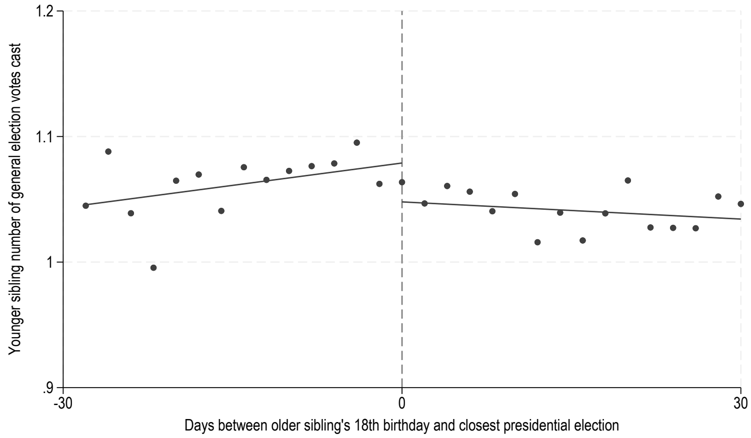
(c) Ever voted in future election



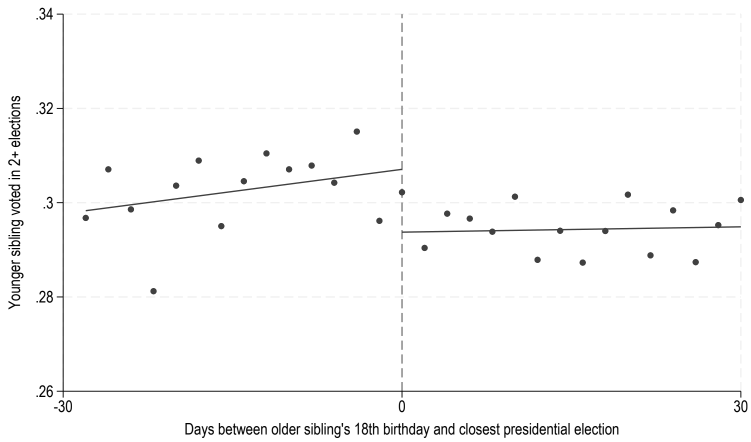
Note: This figure shows binned scatterplots of average older sibling voting outcomes by days between the older sibling's 18th birthday and the presidential elections of 2004, 2008, and 2012. Each bin consists of 2 days.

FIGURE A7. EFFECT OF OLDER SIBLING'S AGE 18 ELIGIBILITY ON YOUNGER SIBLING'S VOTING - ADDITIONAL OUTCOMES

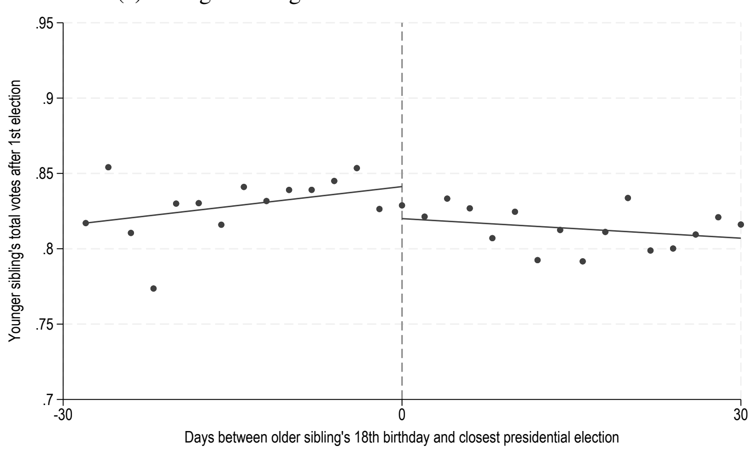
(a) Younger sibling total votes cast



(b) Younger sibling voted 2 or more times



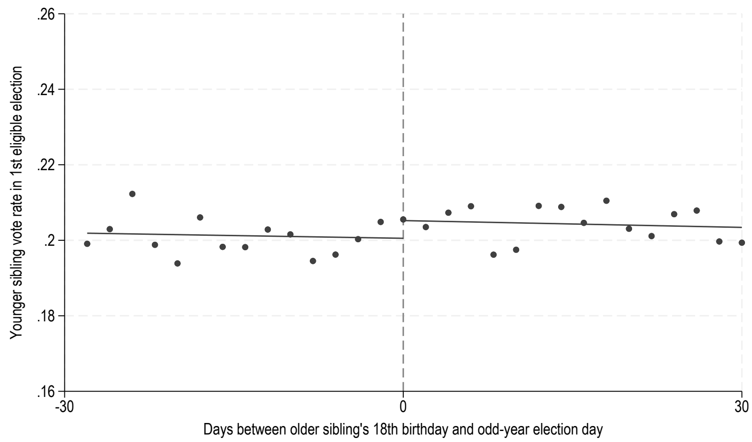
(c) Younger sibling's total votes cast after 1st election



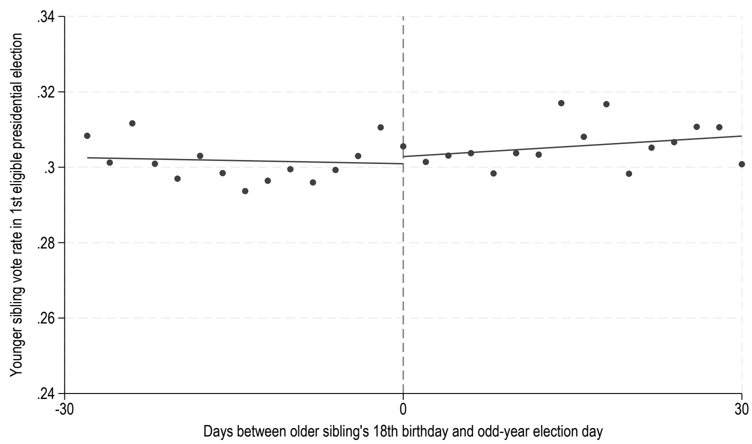
Note: This figure shows binned scatterplots of average younger sibling voting outcomes by days between the older sibling's 18th birthday and the presidential elections of 2004, 2008, and 2012. Voting outcomes are residualized for fixed effects of the younger sibling's first election to match our estimating equation. Each bin consists of 2 days.

FIGURE A8. PLACEBO TEST - EFFECT OF OLDER SIBLING'S AGE 18 ELIGIBILITY FOR 'ODD YEAR' ELECTIONS ON YOUNGER SIBLING'S VOTING

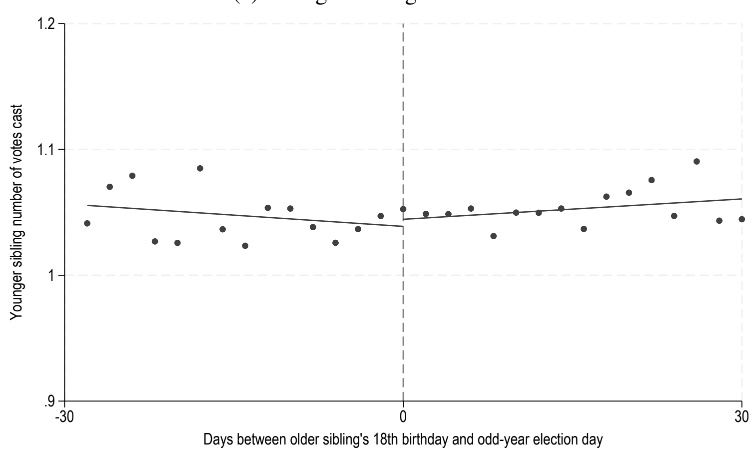
(a) Younger sibling voted in 1st election



(b) Younger sibling voted in 1st presidential election



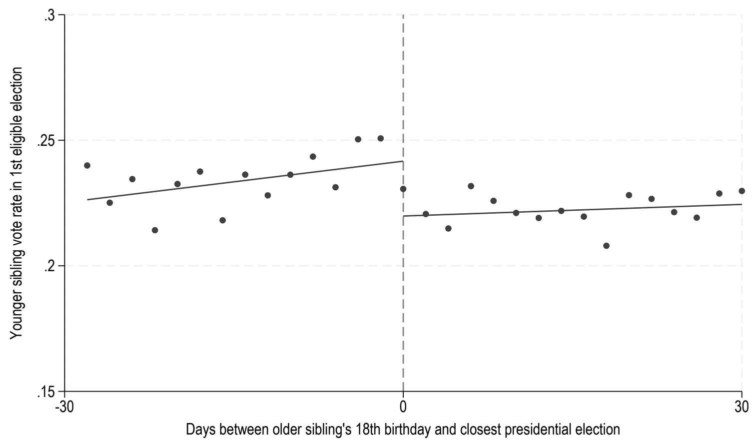
(c) Younger sibling votes cast



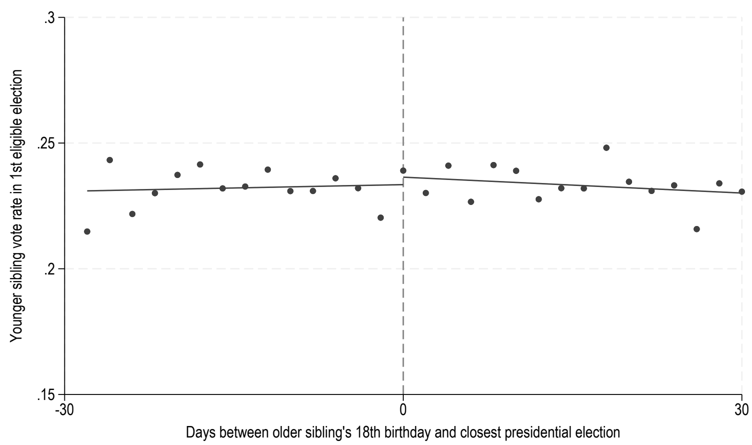
Note: This figure shows binned scatterplots of average younger sibling voting outcomes by days between the older sibling's 18th birthday and the nearest odd year election date, using the 'special election' dates in those years (2005, 2007, 2009, 2011). Each bin consists of 2 days.

FIGURE A9. EFFECT OF OLDER SIBLING'S AGE 18 ELIGIBILITY ON YOUNGER SIBLING'S VOTING - BY SIBLING SEX COMPOSITION

(a) Same sex sibling pairs



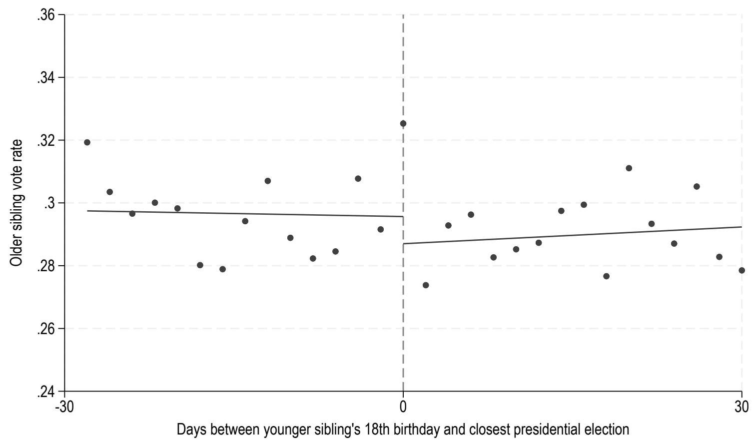
(b) Opposite sex sibling pairs



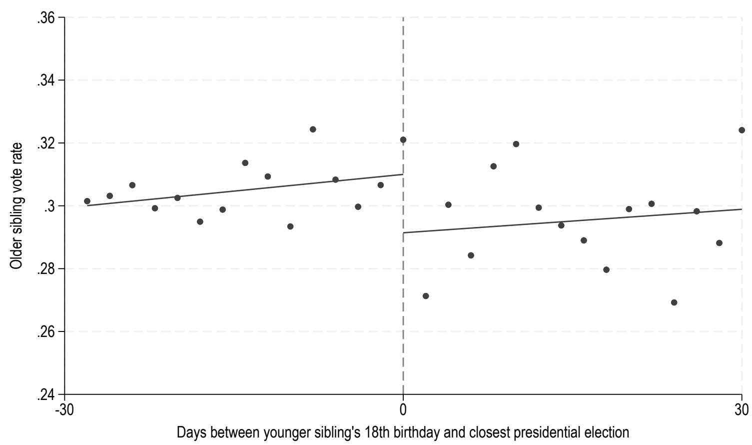
Note: This figure shows binned scatterplots of average younger sibling 1st election vote rates by days between the older sibling's 18th birthday and the presidential elections of 2004, 2008, and 2012. Voting outcomes are residualized for fixed effects of the younger sibling's first election to match our estimating equation. Each bin consists of 2 days. Panel A includes same sex sibling pairs (brother and sister pairs). Panel B includes opposite sex sibling pairs.

FIGURE A10. EFFECT OF YOUNGER SIBLING'S AGE 18 ELIGIBILITY ON OLDER SIBLING'S VOTING - BY SIBLING SEX COMPOSITION

(a) Same sex sibling pairs



(b) Opposite sex sibling pairs



Note: This figure shows binned scatterplots of average older sibling vote rates in their younger sibling's 1st election by days between the younger sibling's 18th birthday and the presidential elections of 2004, 2008, and 2012. Voting outcomes are residualized for fixed effects of the election closest to the younger sibling's 18th birthday to match our estimating equation. Each bin consists of 2 days. Panel A includes same sex sibling pairs (brother and sister pairs). Panel B includes opposite sex sibling pairs.

Appendix B - Matching Process

In this appendix, we discuss the process of matching the education records with the voting records and then we assess match quality by comparing the matched data against a nationally representative dataset on voting rates. Since this study employs the same base data as Bell et al. (2024), material in this appendix is largely reproduced from Appendices A and B of Bell et al. (2024) with the authors' permission.

B1. Preparing the Education Data

The College Board data initially consists of 39.6 million observations across cohorts graduating high school between 2004 and 2018. We then restrict the sample to students who attended high school in one of the 50 states or Washington, DC, and those who have complete information on birth date (i.e. month, day, and year). This leaves over 37.6 million observations that serve as the base sample. Our analyses focus on the 21.3 million observations in the 2004 to 2012 cohorts.

B2. Preparing the Voting Data

We begin with a dataset that keeps voters with a birth year after 1984 to be consistent with the education data. We also keep voters where their year of birth is missing, leaving us with 52,921,628 observations.

With this initial voting dataset, the goal is to get one observation per person. Getting one observation per person presents two challenges relative to the education data. First, the DOB may be partially incomplete and in rare cases, entirely missing.[19] In the education data, we removed those observations from the base sample. We cannot drop those voting records because someone with missing DOB may be a match to the education data. The supposed match in the education data will look as if they never voted if we remove those voting observations. Second, the same person can show up more than once across states if they move (and potentially within the same state if a state's records are inexact and/or missing a piece of the birth date).

To address the above issues, we first go from one observation per person per state to one observation per person across states. This involves finding exact matches and then fuzzy matches using name and DOB, all while making sure there is no overlap in the same election, which is indicative of this being more than one person. We also create a series of flags, which we use in robustness tests, when there are common names or missing information on the DOB.

[19] 4.5 percent had a missing birth year, 14.6 percent had a missing birth month, and 21 percent had a missing birth day.

B3. Identifying Voters That Are Not Duplicates

We first identify all voter observations across all states that do not need to be deduplicated and set them aside to what will be the final dataset of deduplicated voters. To do so, we start with all observations with a unique first and last name combination across all states. At the end of this step, 34.4 percent of the voter data are considered unique observations.

With the set of voters with non-unique first and last name combinations, we identify all observations that were unique by first name, last name, and birth year. After this step an additional 21.8 percent of the voter data are considered unique observations. We then identify all observations that were unique by first name, last name, birth year and birth month. At the end of this step, an additional 20.4 percent of the voter data are considered unique observations. Then we do the same by adding birth day, producing an additional 10.9 percent of the voter data considered unique. We do the same thing again with the addition of middle initial. After this step an additional 3.8 percent of the voter data are considered unique. Along each step, we remove the "unique" voter but generate a flag that describes how the uniqueness was determined.

At the end of this process, 91.3 percent of the voter data are considered unique observations.

B4. Identifying Duplicate Voters

With the remaining 8.7 percent of voting data, we identify the observations with the same DOB information, the same exact first and last name, but differing voting history (e.g. one observation voted in an election in 2010 but did not vote in 2012 and the observation with the same name and DOB information had the opposite voting pattern). The vast majority of these observations were collapsed into a single observation and we retained whether we collapsed two or three (Less than 1.1 percent of these observations were unable to be collapsed into a single observation). Almost all were initially two observations that we collapsed into one.

This deduplication process accounted for roughly 7.3 percent of the original voting observations, leaving just 1.5 percent of original voting observations. We take these observations and place them with the uniquely identified and deduplicated observations and treat them as unique.

From here, we created our final dataset to use in the matching to the College Board data. It contains each dataset that from each round of deduplication to identify unique observations and the remaining observations that were not uniquely identified. In all, we created a dataset with 50,975,728 observations of voting history.

After each de-duplication round, in which observations were uniquely identified, we created a flag to identify all the observations that we determined were uniquely identified for that round to use in a series of robustness tests.

B5. Exact Matching

We start by using exact matching of the education data to the voting data. An exact match required perfect agreement between the first name, last name, middle initial, and date of birth (DOB), none of which could be incomplete in either dataset for this step. We identified 14,175,803 unique matches.[20] This implies that 35.5 percent of the education sample was found in the voting records through exact matching.

B6. Fuzzy Matching

After removing any observation that was exactly matched from both the education and voting datasets, we implement a fuzzy matching algorithm, in three broad steps.

First, we use College Board's 23 step fuzzy matching algorithm - a process they use in other applications. The algorithm starts by very slowly loosening the exact matching criteria. Specifically, it starts with exact matches on first name, last name, and DOB but one or both of the datasets are missing a middle initial, everything matches exactly but for one edit to one name, and everything matches but the first and last name are swapped. The most relaxed criteria - the 23rd step - matches exactly on first name, DOB, and gender but one of the two last names from the two datasets is a suffix of the other and only one of the middle initials is missing. After each step, the matched observations are not replaced for additional matches. We also retain each step in which the observations are matched for robustness tests. This process generates 6,547,716 additional matches, which is an additional 16.4 percent of the education observations.

Second, we used the data linkage method (also known as the editing distance method) employed by Dusetzina, Tyree, Meyer, Meyer, Green, & Carpenter (2014). The method calculates a probability that two strings are a match with the following formula:

$$\sum_{i=1}^{2}[1-(length(name_i)*spedis(name_i,name_{-i}))/2400)]/2 \quad \text{(B1)}$$

The two names from each dataset, indexed by i, are compared in both character length and also "spelling distance" (i.e. *spedis*). Spelling distance is a common function in statistical software that compares the letters in the name.[21]

We consider any value greater than 0.95 a high enough probability to be a match. We first do this for first names, maintaining an exact match on last name and DOB, and then again for last name. These generate an additional

[20] 83,652 students were matched to multiple voting records (usually two), creating 193,986 observations. In our main analyses, we randomly choose one of these matches but our results are entirely insensitive to alternatives analyses, largely because this impacted such a small fraction of the sample.

[21] Documentation for *spedis* in SAS, the statistical software we used, can be found here: https://support.sas.com/resources/papers/proceedings/proceedings/sugi25/25/cc/25p086.pdf.

27,557 matches (one percent of the education data).

Third and finally, we slightly loosen the criteria on birth date by using exact matches on everything previously described but the voting data has a missing birth day or birth month, but not both and not missing year, so there is no conflicting information. This generates 635,344 additional matches (nearly two percent of the education data).

In total, we matched approximately 27.7 million observations, accounting for almost 53.6 percent of the education sample.

B7. Assessing Match Quality

Next, we assess the quality of our match. To do so, we compare voting rates in our matched data for different groups and different elections to reported voting rates in the Voting and Registration dataset from the U.S. Census Bureau.[22] These data come from a supplement to the Current Population Survey. We focus on the voting rates among 18-24 year olds in the 2012 and 2016 elections, a time period that our matched data covers well.

In Table B1 through Table B3, we compare voting rates in the two datasets by election, age, race, sex, state, and sometimes combinations of those variables. Generally speaking, we find lower voting rates in our matched data than in the Census data, but the patterns across subgroups follow one another. The lower voting rates is expected, because our matching process is imperfect and somewhat conservative. And it is reassuring that the relative voting rates across subgroups generally match.

Table B1 shows voting rates in the two datasets by age and sex. The table shows that the voting increases with age (in both elections), according to the Census. Our matched data show a similar pattern, although at lower rates.

Table B1 also highlights the differences between the two samples by sex. Similar to the Census data, older males vote more than younger males in the matched data. However, this is not true for females, especially in the 2012 election. This is likely because we had difficulty matching to women who change their last name. This also explains why females vote at higher rates than males in both datasets for the youngest cohorts, before women typically get married and change their last name, but not so for older cohorts. For older cohorts, Census data suggest females are more likely to vote than males, but the matched data does not. These facts motivate some robustness tests that focus on males and recent cohorts.

Table B2 shows voting rates in the two datasets by race and sex, The table shows that Black people are the most likely to vote in 2012 in both elections and Asian people are least likely. Similarly, White people are most likely to vote in 2016 in both datasets and Asian people are the least likely. These similar patterns between the datasets is comforting.

Finally, voting rates in the two datasets are compared by state. It is immediately clear that our matched sample

[22]https://www.census.gov/topics/public-sector/voting.html

reflects the voting rates of the Census in some states better than others. This is partially because College Board has relatively low coverage in some states, like Mississippi. However, in states where College Board has substantial coverage, such as Virginia, the two voting rates are well aligned. This motivates a few additional robustness tests, including only using states where College Board has substantial coverage and only using states where the voting rates between these two datasets are well aligned.

TABLE B1—COMPARISON OF VOTING RATES IN CENSUS DATA AND MATCHED DATA IN 2012 AND 2016 ELECTIONS, BY AGE AND SEX.

		Voted in 2012 (Percent)			Voted in 2016 (Percent)		
	Census Age (in years)	High School Cohort	Census	Matched Data	High School Cohort	Census	Matched Data
Overall	18	2012	31.3	25.2	2016	32.3	30.0
	19	2011	33.9	25.2	2015	36.9	30.9
	20	2010	39.7	24.7	2014	40.1	31.0
	21	2009	37.5	24.8	2013	38.8	31.5
	22	2008	41.0	26.9	2012	41.4	33.0
	23	2007	40.0	26.7	2011	41.6	33.1
	24	2006	41.5	26.6	2010	43.8	33.5
Male	18	2012	27.8	23.2	2016	30.5	27.1
	19	2011	30.4	22.8	2015	34.4	27.2
	20	2010	34.4	22.7	2014	37.4	27.3
	21	2009	34.3	23.3	2013	35.8	27.9
	22	2008	38.3	26.3	2012	36.3	29.7
	23	2007	39.3	27.1	2011	38.5	30.4
	24	2006	37.1	28.2	2010	41.9	31.6
Female	18	2012	34.8	27.2	2016	34.3	32.8
	19	2011	37.5	27.5	2015	39.5	34.4
	20	2010	44.7	26.5	2014	42.7	34.5
	21	2009	40.9	26.2	2013	42.0	35.0
	22	2008	43.7	27.5	2012	46.1	36.3
	23	2007	40.8	26.4	2011	44.8	35.8
	24	2006	46.0	25.1	2010	45.8	35.2

Notes: The U.S. Census data are a nationally representative sample, accessed online here: https://www.census.gov/topics/public-sector/voting.html. The matched data includes College Board test-taker data linked to Data Trust, LLC's national voter records. High school cohort is the year of graduation, comes from College Board data, and is an approximation of age.

TABLE B2—COMPARISON OF VOTING RATES IN CENSUS DATA AND MATCHED DATA IN 2012 AND 2016 ELECTIONS, BY RACE AND SEX.

	Voted in 2012 (Percent)		Voted in 2016 (Percent)	
	Census	Matched Data	Census	Matched Data
Total	38.0	25.7	39.4	31.8
Male	34.7	24.7	36.5	28.7
Female	41.3	26.7	42.4	34.9
Asian	20.1	17.2	25.2	25.9
Black	45.9	33.5	40.2	29.1
Hispanic	26.7	20.1	27.2	26.8
White	37.8	26.8	41.1	36.4
Male + Asian	18.1	15.3	21.2	21.8
Male + Black	41.4	26.5	36.5	22.1
Male + Hispanic	24.0	17.7	21.8	22.5
Male + White	34.6	28.0	38.3	34.9
Female + Asian	22.1	19.1	29.5	30.0
Female + Black	50.2	39.6	43.7	35.4
Female + Hispanic	29.7	22.3	32.7	30.7
Female + White	41.0	25.7	43.9	37.8

Notes: Census voting rates are for 18-24 year-olds; 2012 voting rates in matched data are for the 2006-2012 high school graduation cohorts; 2016 voting rates in matched data are for the 2010-2016 high school graduation cohorts. The U.S. Census data are a nationally representative sample, accessed online here: https://www.census.gov/topics/public-sector/voting.html. The matched data includes College Board test-taker data linked to Data Trust, LLC's national voter records. High school cohort is the year of graduation, comes from College Board data, and is an approximation of age.

TABLE B3—COMPARISON OF VOTING RATES IN CENSUS DATA AND MATCHED DATA IN 2012 AND 2016 ELECTIONS, BY STATE.

	Voted in 2012 (Percent)		Voted in 2016 (Percent)	
State	Census	Matched Data	Census	Matched Data
Mississippi	62.4	17.3	46.1	18.1
Minnesota	57.0	21.4	49.6	22.5
Wisconsin	53.3	24.7	45.6	21.2
Colorado	52.5	23.0	43.1	25.8
South Carolina	51.3	28.4	42.7	33.9
New Hampshire	50.0	17.8	*	21.1
Oregon	47.6	25.7	45.2	26.4
Iowa	46.6	41.1	35.5	52.9
Massachusetts	45.6	33.5	39.9	42.8
Rhode Island	45.2	27.2	*	35.9
North Carolina	45.1	16.8	44.8	18.5
Ohio	44.7	38.1	39.6	43.1
Michigan	43.5	17.1	36.1	22.3
Maine	42.8	13.7	48.8	14.3
Missouri	42.5	33.2	45.9	43.5
Maryland	42.1	33.3	48.0	40.1
Virginia	42.0	37.1	54.6	42.9
Montana	40.7	33.3	*	40.2
Delaware	40.5	21.0	*	22.0
Louisiana	40.4	11.6	49.2	14.9
Arizona	40.0	12.3	35.8	17.8
Pennsylvania	39.9	27.1	48.7	40.5
Nevada	38.2	27.3	37.6	32.3
Washington	38.2	33.7	42.2	38.7
Georgia	37.9	17.1	40.5	15.6
Kentucky	37.1	28.7	51.1	40.3
New Mexico	37.1	11.3	37.8	12.3
Connecticut	36.9	27.9	37.0	37.7
Florida	36.8	32.0	33.1	36.9
California	36.5	23.0	37.5	33.9
Nebraska	36.2	30.7	50.1	45.2
Indiana	35.7	24.8	40.8	31.9
Alabama	35.3	34.0	41.4	40.0
New York	35.0	27.1	34.6	33.2
New Jersey	34.9	32.4	35.9	38.4
Utah	34.8	28.2	41.6	38.3
Tennessee	34.0	26.9	29.9	32.7
South Dakota	32.7	12.3	*	16.6
Illinois	32.2	29.1	45.3	41.6
Kansas	30.1	26.5	33.8	36.1
Idaho	29.8	9.2	40.7	13.3
Oklahoma	27.2	21.9	32.4	34.2
Arkansas	24.3	26.0	33.1	30.9
West Virginia	22.6	28.1	32.2	38.0
Texas	22.5	19.1	27.3	24.1
Hawaii	22.1	8.0	20.4	10.1
District of Columbia	*	19.9	*	20.9
North Dakota	*	15.0	*	19.2
Alaska	*	12.4	*	15.1
Vermont	*	11.3	*	14.9
Wyoming	*	9.2	*	16.5

Notes: Census voting rates are for 18-24 year-olds; voting rates in matched data are for the 2006-2012 high school graduation cohorts. The U.S. Census data are a nationally representative sample, accessed online here: https://www.census.gov/topics/public-sector/voting.html. The matched data includes College Board test-taker data linked to Data Trust, LLC's national voter records. High school cohort is the year of graduation, comes from College Board data, and is an approximation of age. *indicates that Census data are not available.