NBER WORKING PAPER SERIES

THE EFFECT OF CHILDHOOD ENVIRONMENT ON POLITICAL BEHAVIOR

Jacob R. Brown
Enrico Cantoni
Sahil Chinoy
Martin Koenen
Vincent Pons

Working Paper 31759
http://www.nber.org/papers/w31759

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
October 2023, Revised August 2026

We thank Ryan Enos for sharing the TargetSmart voter file data and Opportunity Insights for sharing the Infutor data. For suggestions that have improved this article, we are grateful to Raj Chetty, Ryan DeTamble, Rafael DiTella, Mitch Downey, Nathaniel Hendren, Matthew Weinzierl, Jennifer Wolak, as well as seminar and conference participants. We gratefully acknowledge generous financial support from the Andrew Carnegie Foundation. Any views expressed are those of the authors and not those of the U.S. Census Bureau. The Census Bureau has reviewed this data product to ensure appropriate access, use, and disclosure avoidance protection of the confidential source data used to produce this product. This research was performed at a Federal Statistical Research Data Center under FSRDC Project Number 2747 (CBDRBFY26- P2747-R12874). The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

At least one co-author has disclosed additional relationships of potential relevance for this research. Further information is available online at http://www.nber.org/papers/w31759

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2023 by Jacob R. Brown, Enrico Cantoni, Sahil Chinoy, Martin Koenen, and Vincent Pons. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.

The Effect of Childhood Environment on Political Behavior
Jacob R. Brown, Enrico Cantoni, Sahil Chinoy, Martin Koenen, and Vincent Pons
NBER Working Paper No. 31759
October 2023, Revised August 2026
JEL No. D72, P0

ABSTRACT

We estimate the effect of childhood environment on future political behavior. Linking national voter files to Census Bureau and commercial address data, we compare children who moved between the same counties but at different ages. Growing up in a county where their peers are 10 p.p. more likely to affiliate with the Republican Party makes children 4.7 p.p. more likely to do so. The effects are similar for Democratic affiliation and slightly larger for registration and turnout. Exposure during adolescence is especially influential and effects decay slightly after the first election. They reflect both state-level factors and more local forces.

Jacob R. Brown
Boston University
jbrown13@bu.edu

Enrico Cantoni
Università di Bologna
Department of Economics
enrico.cantoni@unibo.it

Sahil Chinoy
Harvard University
sahil.chinoy@gmail.com

Martin Koenen
Harvard University
martin_koenen@fas.harvard.edu

Vincent Pons
Harvard University
Harvard Business School
and CEPR
and also NBER
vpons@hbs.edu

1 Introduction

Our preferences and behavior are shaped by innate characteristics, current contexts, family upbringing, and the environments in which we grew up. Much of the work of social scientists has aimed to assess the relative contributions of these four sets of factors. In democracies, where the designation of political leaders is delegated to millions of citizens, special attention has been given to the forces that lead people to favor one party over another and to vote rather than abstain. Fowler et al. (2008) establish the importance of genetic factors by comparing correlations in political behavior among monozygotic and dizygotic twins. Other studies point to the intergenerational transmission of political attitudes within families (Jennings et al., 2009; Healy and Malhotra, 2013). Cantoni and Pons (2022) estimate the share of cross-state variation in turnout and party affiliation due to people's *current environment* by tracking changes in the behavior of voters who move across counties or states.

By contrast, existing evidence on the influence of *childhood environment* on later political behavior remains less definitive. Yet this influence may be no less important: childhood environments may help sustain democratic equilibria marked by strong participation and regular electoral turnovers, which in turn support legitimate policymaking and robust economic performance (e.g., Marx et al., 2025; Cantoni et al., 2025) – or instead entrench harmful trends such as polarization and populism (e.g., Azzimonti, 2018; Funke et al., 2023). Recent studies examine specific components of childhood environment, including racial context (Billings et al., 2021; Brown et al., 2021; Kaplan et al., 2025), economic adversity (Chyn and Haggag, 2023; Elder et al., 2024), school quality (Cohodes and Feigenbaum, 2025), and immigration (Bolotnyy et al., 2026), but we lack evidence on the combined effect of all such factors.

In this paper, we estimate the *aggregate* effect of where voters grow up during childhood and adolescence on their partisanship and political behavior as adults, identify the most formative ages, and uncover the mechanisms underlying these effects.

Our analysis exploits administrative voter file data from TargetSmart, which include the exact residential address, individual turnout, and, in states recording it, party affiliation of all registered voters in the United States from 2012 through 2021. These data, however, do not tell us where children grew up. We obtain that information through two complementary linkages. For analyses of registration and turnout, we link the voter file to individual-level administrative data from the U.S. Census Bureau, which provide parent-child links and residential address histories for the universe of eligible voters. To our knowledge, this is the first linkage between voter file records and individual-level Census data used to study nationwide turnout and voter registration. For analyses of party affiliation, which we could not conduct in the Census environment, we link the voter file to commercial address histories from Infutor, match young voters to their parents using shared last names and addresses, and trace parents' residential histories in Infutor. In both samples, we infer childhood locations from parents' address histories, assuming that children lived and moved with their parents

before turning 18.

The strategy we use to estimate the effect of childhood environment closely follows Chetty and Hendren (2018a), who measure childhood place effects on economic outcomes such as adult earnings. We focus on people who relocated once during childhood and compare individuals who moved between the same origin and destination counties but at different ages. Then, we ask whether the political behavior of those who moved earlier and thus spent more time in the destination county is more similar to the behavior of permanent residents in that place. For instance, comparing children who moved at age six with those who moved at age seven enables us to identify the effect of spending one additional year (the seventh) in the destination county. Importantly, we do not need to assume that children's potential outcomes are independent of where their parents choose to live, only that they are uncorrelated with the timing of moves.

Our main estimates measure effects in the first general election in which individuals are eligible to vote after turning 18. We find that children who move to the destination county at a younger age are more strongly influenced by that environment. Overall, our results imply that spending one's entire childhood (ages 0 to 19) in a county where permanent residents are 10 percentage points more likely to register in their first eligible election increases one's own likelihood of registering by 6.87 percentage points. The corresponding effect is 5.75 percentage points for turnout. Party affiliation effects are only slightly smaller: when permanent residents are 10 percentage points more likely to affiliate with the Republican or Democratic Party, one's own probability of affiliating with these parties increases by 4.69 and 4.09 percentage points, respectively. In other words, the environment in which people grow up makes their future partisanship and political participation 40 to 70% more similar to the behavior of the people they grow up around. The influence of place persists well beyond the first election, with a slower decay over time for party affiliation than for registration and turnout.

We address concerns about endogenous move timing by showing that our estimates are robust to including family fixed effects, thereby only using variation across siblings who moved at the same time but spent different portions of their childhood in the same origin and destination. We also show that movers' outcomes are better predicted by the outcomes of permanent residents of the same age, gender, or race, and infer that our results are unlikely to be driven by shocks affecting both parents' decision to relocate and their children's future political behavior.

A separate concern, specific to the party affiliation results, is the risk of sample selection bias. Thanks to the use of Census data, the sample used to study registration and turnout approximates the universe of voting-eligible citizens, including those who never register and therefore never appear in the voter file. The Census data also include comprehensive parent-child links. By contrast, the sample used to study party affiliation primarily includes individuals who

appear on the voter rolls at least once, and excludes young voters whom we fail to link to their parents or whose parents we fail to link to the Infutor address data. Reassuringly, 80% of voting eligible people of the relevant birth cohorts in the Census data do appear in the voter records, and estimated exposure effects do not vary much with the registration rate in movers' destination counties. They also remain very similar when we focus on voters who turn 18 shortly before Election Day and are thus more likely to live with their parents, and to be linked to them, around the time they register for their first election.

To examine when childhood environment matters most, we compare exposure effects across different ages. We observe a steeper gradient in the correspondence between the outcomes of one-time movers and permanent residents during their teenage years relative to pre-teen years, suggesting that adolescence is a particularly formative period for political behavior. Since the effects of adolescent years are estimated based on children who move to the destination county at a later age and thus spend fewer years in that county before turning 18, these larger effects could also reflect diminishing marginal returns to exposure. We cannot use one-time movers to distinguish between these two interpretations, so we develop and implement a novel empirical strategy. We focus on individuals who move not just once but multiple times, and compare those who spend the same number of childhood years in a county, but at different ages. By holding total exposure constant, this design shuts down the diminishing marginal returns channel and isolates the importance of exposure timing. We find larger exposure effects during adolescence again. On average, the influence of one's environment between ages 13 and 19 is four times larger than the influence of one's environment between ages 0 and 12.

Finally, we examine the mechanisms driving childhood place effects. We first ask whether the effects on young adults' political behavior are mediated by socioeconomic outcomes, using sample restrictions that hold place effects on future earnings and educational attainment constant across origin and destination locations. The estimates remain nearly unchanged, suggesting that childhood environments affect future political behavior directly. We then compare effects on people who move across states, across counties within states, and across ZIP codes within counties, to assess the importance of factors operating at each geographic level. We find that state-level factors, such as electoral procedures and the party exerting power, account for much of the exposure effects on registration and turnout but for a smaller share of the effects on party affiliation. Young voters' partisanship is more strongly influenced by sub-state factors, such as schools and peer effects.

The paper makes three core contributions. First, we provide the most comprehensive evidence to date on the causal effect of childhood location on adult political behavior. Cantoni and Pons (2022) measure the immediate influence of the context in which voters currently live (including state voting rules and the political and economic environment) on their elec-

toral participation and party affiliation, finding particularly strong effects for young adults.[1] We study similar outcomes but focus instead on the long-term influence of the places where voters grew up. Our design enables us to estimate the aggregate effect of childhood environment, whereas most existing research on the childhood context and political behavior focuses on specific components (e.g., Billings et al., 2021; McNeil et al., 2023; Kaplan et al., 2025; Cohodes and Feigenbaum, 2025). Recent work by Chyn and Haggag (2023), Elder et al. (2024), and Knight and Zhang (2024) studies individuals displaced from disadvantaged neighborhoods during childhood, due to the demolition of Chicago housing projects and the Moving to Opportunity experiment, and finds contrasting effects of these relocations on adult participation. Our quasi-experimental strategy pools information from all voting-eligible citizens in the U.S., allowing us to study a much broader population. In addition, we study childhood environment effects not just on registration and turnout but also on partisan affiliation.

Second, we build on previous work in economics and political science documenting the importance of adolescence (e.g., Sears and Valentino, 1997; Pacheco, 2008; Neundorf and Smets, 2017; Ghitza et al., 2023) and of the context of one's first election (e.g., Mullainathan and Washington, 2009; Dinas, 2014) for the formation of durable partisan attachments and political behavior. Since age at move is perfectly collinear with exposure duration at destination, our baseline one-time movers design cannot separate age-specific effects of neighborhood exposure from diminishing marginal returns to exposure duration. We thus implement a multiple-movers design that holds exposure duration fixed, while varying the age of exposure to a mover's first destination. This multiple-movers design shows that early childhood environment influences future political behavior, but that the environment in which one lives during adolescence has much larger effects.[2] Together, these results provide comprehensive evidence that teenage years are especially formative for political behavior.

Third, we contribute to the broader literature documenting the large and long-lasting effects of childhood location on adult outcomes such as earnings, educational attainment, health, marriage, and criminal activity (e.g., Kling et al., 2007; Damm and Dustmann, 2014; Bell et al., 2019; Chetty and Hendren, 2018b; Chyn, 2018; Deutscher, 2020; Laliberté, 2021; Nakamura et al., 2022; Derenoncourt, 2022).[3] Such outcomes generally reflect a complex combination of beliefs, preferences, constraints, and choices. By contrast, we study behaviors that are more immediately under individuals' control – whether to register, vote, and affiliate with a political party – and that more directly express their political preferences. While individual preferences are often taken as given, our results thus shed light on the forces that

[1] Other work on the influence of contemporary context on adults' partisanship, political behavior, and a range of social and political attitudes includes Huckfeldt and Sprague (1987), Campbell (2006), Perez-Truglia (2018), Brown (2025), and Berry and Trounstine (2023).

[2] Chetty and Hendren (2018a) estimate a different multiple-movers specification and conclude from it that exposure effects on income and education are *not* driven by critical-age effects.

[3] See Chyn and Katz (2021) for a systematic review of the recent literature on neighborhood effects.

shape them in the long run. In addition, we show that the effects on political outcomes are not mediated by income or education, suggesting that childhood environment directly affects adult behavior. The disproportionate influence of adolescent exposure, which contrasts with the linear exposure effects on income and education documented by Chetty and Hendren (2018a), further supports this conclusion.

The remainder of the paper proceeds as follows. Section 2 describes the data and the construction of our samples, and Section 3 presents our empirical strategy. In Section 4, we show the main results as well as robustness checks addressing concerns about the identification assumption and the risk of sample selection bias. Section 5 compares exposure effects across ages using multiple-time movers, and Section 6 investigates the mechanisms responsible for childhood environment effects. Section 7 concludes by discussing the broader implications of the results.

2 Data and sample construction

Our analysis requires combining data on young voters' political behavior and their address histories.[4] We measure political behavior using voter file records from TargetSmart, described in Section 2.1. Because individuals only enter the voter file when they register, we do not observe their address histories directly in the voter file before they become eligible to vote at age 18. We thus merge the voter data with address history data covering the residential locations of voters' parents, and assume children lived with their parents during their youth.

We use two distinct methods to link children with their parents and to track parental address histories, yielding two different samples. The first sample ("Census sample") uses the voter file data merged with individual-level administrative data from the U.S. Census Bureau and encompasses the universe of *eligible* voters. This merge was done in cooperation with the Census Bureau. We observe parent-child links and residential address histories in the Census data, as described in Section 2.2. We use this sample to study effects on turnout and registration.

Because our data-linkage agreement with the U.S. Census Bureau covers only voter turnout and registration, we use a second sample ("partisanship sample") to study effects on Democratic and Republican Party affiliation. We construct this sample by merging the voter file data with commercially available address data from Infutor. Unlike in the Census sample, here we only observe individuals who appear at least once in the TargetSmart data. We link children to their parents using shared last name and address, and we use parents' residential address histories from Infutor, as outlined in Section 2.3.

[4]Henceforth, by "voters" we mean all individuals with the right to vote, including those who are eligible to vote but are not registered, and those who are registered but abstain.

Section 2.4 describes our main outcomes and provides summary statistics. Section 2.5 discusses possible sources of sample selection.

2.1 TargetSmart data

We obtained our voter data from TargetSmart, a vendor that collects public voter lists, combines them with commercial data, and sells the resulting data to campaigns, non-profits, and researchers. The TargetSmart data consist of yearly snapshots from 2012 to 2021 covering all registered voters in the country, along with some unregistered individuals. These snapshots contain each voter's name, residential address, gender, race, date of birth, registration status, turnout history, and party affiliation. Gender is recorded directly in the raw voter lists in many states and otherwise imputed or sourced from commercial data by TargetSmart.[5] Race is recorded directly for some or all years in eight states, and otherwise sourced from commercial data or imputed by TargetSmart based on voters' full names and local census demographics.[6] We observe the exact date of birth of the majority of individuals.[7] Thirty states and the District of Columbia record party affiliation in their voter data in all snapshot years, while the other twenty states do not offer voters the option of affiliating with a party when registering to vote.[8] Party affiliation has been repeatedly found to be one of the strongest predictors of an individual's likelihood to vote for Republican or Democratic candidates (e.g., Bartels, 2000; Green et al., 2002; Guo, 2023).

We construct a panel of all individuals appearing in the TargetSmart data by matching records that correspond to the same person across states and over time. Our cleaning and de-duplication procedure uses state-level administrative identifiers, personal information such as name and date of birth, turnout history, as well as TargetSmart's own linkage model, which incorporates external information such as the USPS National Change of Address Database. The full procedure is described in detail in Appendix C. There, we also present validation checks showing that the de-duplicated TargetSmart data closely align with independent sources on the number of registered voters, voter turnout, and Republican and Democratic vote shares in each state.[9]

[5]Overall, 53.8% of voter-years include gender information sourced directly from state records, and 14.9% include race information sourced directly.

[6]The states recording race directly in the voter file are Alabama, Florida, Georgia, Louisiana, North Carolina, South Carolina, Tennessee, and Texas.

[7]The recorded birth date for 21.3% of voters in the partisanship sample is January 1. It is likely that only the year of birth, rather than the exact date of birth, was recorded for these voters. In Appendix Table A1, we show that estimates excluding them are similar to the main results.

[8]The states recording party affiliation are Alaska, Arizona, California, Colorado, Connecticut, Delaware, the District of Columbia, Florida, Iowa, Idaho, Kansas, Kentucky, Louisiana, Massachusetts, Maryland, Maine, North Carolina, Nebraska, New Hampshire, New Jersey, New Mexico, Nevada, New York, Oklahoma, Oregon, Pennsylvania, Rhode Island, South Dakota, Utah, West Virginia, and Wyoming.

[9]Appendix Figure B1 plots the number of registered voters in TargetSmart in each state in 2021 against the 2020 Current Population Survey (CPS) estimates of registered voters. State-level counts are plotted separately for four age groups: voters 18 to 24 years old, 25 to 34, 35 to 44, and 45 to 64. For each group,

2.2 Census sample

Our registration and turnout analyses use voter file data merged with restricted-access U.S. Census Bureau data. These linked data offer two main advantages over the data underlying the partisanship sample. First, they cover all individuals, including all voting-eligible citizens who never register to vote and therefore never appear in the voter data. We can thus use voter registration as an outcome and estimate effects on turnout without restricting the analysis to individuals who appear in at least one yearly snapshot of the voter data. Second, the Census data contain information on parent-child links and address histories from a variety of government records, yielding comprehensive and high-quality information about each child's residential history before age 18, a critical input to our empirical strategy.

To construct the Census sample, we start with all individuals in the Census data who were born between November 3, 1992 and November 3, 2002 and who were citizens by the time of their 18th birthday. This time frame ensures that we observe these voters in the first general election for which they were eligible to vote, since the TargetSmart data cover the 2012 to 2020 election cycles.[10] Individuals are uniquely identified by a "protected identification key" (PIK). The Census Bureau used a probabilistic matching algorithm to assign each row in the TargetSmart data to a PIK based on name, gender, date of birth, and address.[11] As shown in Appendix Table A2, 94.9% of rows in the TargetSmart data corresponding to individuals born between 1992 and 2002 were successfully linked.

We identify a child's parents using the Census Household Composition Key, which contains probabilistic parent-child links, and the 2000 and 2010 decennial censuses, which record intra-household relationships. We then retrieve parents' address histories from the Master Address File Auxiliary Reference File (MAF-ARF), which records all addresses associated with each individual in federal administrative data, starting in 2000.[12] The final address panel measures yearly residential location at the ZIP code level, thus associating each parent-year with at most one ZIP code.[13]

counts in TargetSmart closely match those in the CPS but are generally larger, with slopes between 1.079 and 1.154. This slight difference is likely due to deadwood and imperfect de-duplication in the TargetSmart data. Appendix Figure B2 plots 2020 total votes in the TargetSmart data against turnout from state-level administrative reports collected by the United States Election Project (McDonald, 2022). The state-level correlation is again close to 1. Finally, Appendix Figure B3 plots the share of registered Democrats and Republicans in each state against the 2020 presidential vote share and also shows a high correspondence.

[10]Individuals born on November 2, 1992, turned 18 on Election Day and were therefore eligible to vote in the November 2, 2010 election. Those born one day later, on November 3, 1992, were the first cohort whose first age-eligible election was the 2012 election. Similarly, individuals born on November 3, 2002, were the last cohort whose first age-eligible election was the November 3, 2020 election.

[11]See Wagner and Lane (2014) for details on this algorithm.

[12]Source files can include IRS 1040 and 1099 files, the Selective Service System, Medicare Enrollment Database, Indian Health Service database, HUD Public and Indian Housing, and HUD Tenant Rental Assistance Certification System (Graham et al., 2017).

[13]We drop return moves occurring suspiciously close to an earlier move: if an individual moves from ZIP code A to ZIP code B and, within one year, relocates back to the original address in ZIP code A, we drop

Comparing residential histories recorded in MAF-ARF to those from commercial data (described in the next section), we find that when both sources record a move between the same origin-destination county pair, the Census data tend to record the move later. Appendix Figure A1 shows a histogram of the difference between the year a given move was recorded in the Census data versus the commercial data: in 45.9% of cases, the Census measured the move one year later, compared with 21.5% of cases when the two sources measured the move in the same year.[14] Because the modal difference is one year and it is unlikely that the commercial data record changes of address *before* voters relocate, we adjust move timing in the Census sample by subtracting one year from each move; for example, we consider moves that the MAF-ARF records in 2016 as having occurred in 2015.[15]

We then infer each child's address history from exactly one "primary" parent, chosen using the following lexicographic order: the parent with the longest address history, the parent with the most recent address observation, the female parent, and the youngest parent. We define children who share the same primary parent as siblings and assign them to the same family unit.

Finally, we define the two outcomes analyzed using the Census sample: registration and turnout. We consider an individual in the Census sample to be registered to vote in a given election year if they are linked to any record in the *following* year's TargetSmart snapshot and at least one of these linked records is recorded as having registered, or if the individual is recorded as having voted in the general election of the given election year. We consider an individual to have voted in a given election year if they are linked to *any* TargetSmart record indicating that they turned out in that election.[16] In these analyses, we use race, gender, and age obtained directly from the Census data (not from the voter file).

2.3 Partisanship sample

Because we could not include information about partisanship in the Census sample, we study effects on Democratic and Republican Party affiliation by constructing a second sample

this return move.

[14] A possible explanation for this pattern is that many observations in MAF-ARF come from tax returns, which are filed towards the beginning of the year, meaning that moves after Tax Day might not be recorded until the following year.

[15] Appendix Table A3 shows that our main estimates of exposure effects by age 19 are largely unaffected by this adjustment, but effects after age 20 are larger without it, likely reflecting the type of measurement error in move year described above.

[16] Since voter files may have been drawn or updated months before Election Day, we assess registration and party affiliation based on voter files from the year after the election. Unlike registration and party affiliation, which are updated as of the date a given TargetSmart snapshot was drawn, each snapshot reports an individual's entire voter turnout history up to that date. For example, the 2021 voter file includes turnout information about 2020, 2018, 2016, 2014, etc. For this reason, along with the fact that states can take several years to fully update their turnout records, we define turnout outcomes by checking whether *any* voter file snapshot linked to an individual indicates turning out in a particular election.

("partisanship sample") which merges the TargetSmart voter file data with commercially available address data. Our address data come from Infutor, a vendor that aggregates public and proprietary data such as phone books, property deeds, magazine subscriptions, and credit header files.[17] These data include 228 million individuals born between 1942 and 1986, corresponding to the cohorts of likely parents in our data. Each record indicates the person's full name, current residential address, date of birth,[18] and gender. Crucially, these data also include longitudinal address information covering more than 1.3 billion address observations from 1992 to 2021.

Infutor's coverage of the adult population is relatively comprehensive, especially since 2000 (Diamond et al., 2019). In Appendix B, we show that counts of Infutor records closely match adult population counts from the decennial census and that the data are representative across U.S. counties.[19] Coverage is particularly good for individuals aged 30 through 49, the age range in which parents in our sample are most likely to be observed during their children's childhood and adolescence.

To construct the partisanship sample, we start with all individuals born between November 1992 and October 2002 who ever appear in our TargetSmart data. We first link them to potential parents: people who have the same last name, have lived at the same address at some point, are at least 16 years older, and appear in either the TargetSmart or Infutor data.[20] We provide more details about this procedure in Appendix C. Overall, we link 63.0% of young voters in the TargetSmart data to a parent. Young voters not linked to a parent include those who never appear at an address where we also observe a parent and those who live at the same address as a parent but have a different last name.

We then construct Infutor address histories for parents. The address history is immediately available for parents identified in Infutor, but parents identified in TargetSmart first need to be matched to Infutor records based on name, date of birth, and address. Overall, 87.1% of parents identified in the partisanship sample are successfully matched to Infutor address history records. We build parents' address histories using only Infutor data, and, like in the

[17]For further information on the Infutor data and previous studies using them, see Diamond et al. (2019) and Pennington (2025).

[18]Since the Infutor data do not include the day of birth, by "date of birth" we henceforth mean the month and year of birth when we refer to these data.

[19]Appendix Figure B4 shows counts of active Infutor records – individuals appearing in the dataset by a given date and not flagged as deceased – divided by the total adult population in each decennial census from 1990 to 2020. Data coverage in 2000 and later is close to 100%. Further corroborating the good coverage of the Infutor data, Appendix Figure B5 reports estimated slopes and R^2 statistics from linear fits of county-level headcounts in the Infutor data on county-level total adult population in each decennial census. Both the slope and the R^2 are close to one in 2000 and later years. The estimated slope for 1990 is lower (0.504), but the R^2 remains high, suggesting that, even for earlier years, the geographic distribution of Infutor records closely tracks that of the adult population.

[20]For examples of other papers using last name, address, and date of birth to construct family links, see Hersh and Ghitza (2018) and Pons and Liegey (2019).

Census sample, construct a panel associating each parent with at most one ZIP code per year.

We assume that children move with their parents until their first appearance in the voter file data. For each young voter, we infer address history from exactly one "primary" parent, chosen following the same rules as in the Census sample. We also define families in the same way as in the Census sample.

Appendix Figure B6 plots the probability that parents identified in the partisanship sample move across any two states in a given year against the annual probability of a cross-state move from 2000 to 2019 in the American Community Survey (ACS), measured for a similar sample of adults with children. The fitted slope is 0.581 and the R^2 is 0.461, indicating high correspondence between interstate mobility measured in the partisanship sample and in the ACS.

We classify an individual's Democratic or Republican Party affiliation in a given election year based on the party with which the individual is registered in the following year's TargetSmart snapshot, analogous to how we define voter registration in the Census sample. The Democratic and Republican Party affiliation variables are set to zero if the individual is not registered or is registered but not affiliated with that party. Since we primarily use the partisanship sample to study place effects on party affiliation, we restrict it to individuals who, at the time of the election, live in states where party affiliation is recorded. For movers, we further restrict the partisanship sample to individuals moving between states that both record party affiliation.[21]

2.4 Summary statistics

The Census sample includes 21.9 million permanent residents, defined as voters whose parents never changed county between the year these voters were born and the year they reached age 26, and 7.3 million one-time movers, whose parents moved across counties exactly once before these voters turned 27.[22] Given the restriction to party-affiliation states, the partisanship sample includes 6.4 million permanent residents and 2.2 million one-time movers.[23] Note that permanent residents and one-time movers are defined based only on the address history of their parents, regardless of any residential move these young voters may make themselves after age 18.

[21] We do not make these restrictions to individuals living in or moving between states where party affiliation is recorded for the analyses underlying Appendix Figure A4, in which we compare place effects on turnout between the Census and partisanship samples.

[22] Observation counts from the Census sample are rounded throughout the paper to preserve confidentiality.

[23] In analyses shown in Section 5, we also include voters whose parents moved across counties two or three times. Furthermore, in Section 6.2, we study moves across ZIP codes, including some taking place within the same county.

Table 1 presents summary statistics. In both the Census and partisanship samples, the mean year of birth is 1997 or 1998 for both permanent residents (Panel a) and one-time movers (Panel b), and about 51% of individuals are male. The partisanship sample over-represents white people: 68.5% of permanent residents versus 57.7% in the Census sample, reflecting racial disparities both in voter registration and in the likelihood of appearing in the Infutor data, as well as differences in the demographic composition of states that record party affiliation and those that do not.[24] In both samples, movers are somewhat more likely to be white and less likely to be Black than permanent residents.

The partisanship sample also contains smaller family units. In that sample, permanent residents and movers are linked to parents with an average of 1.6 and 1.5 children, respectively, compared with 1.8 children for both groups in the Census sample. This pattern reflects the more comprehensive parent-child linkage available in the Census administrative data.

Table 1, Panel b presents additional statistics for one-time movers. In both the Census and partisanship samples, these individuals move, on average, at age 11 to a county located around 325 miles away. The differences between average outcomes measured among permanent residents of the same cohort as each mover in destination and origin counties, which we leverage to estimate childhood environment effects, are substantial. The average absolute origin-destination difference, measured in a mover's first eligible election, is 9.1 percentage points for registration, 7.6 percentage points for turnout, and 8.6 and 9.6 percentage points for Republican and Democratic affiliation.

2.5 Sample selection

While the Census sample approximates the universe of eligible voters, the partisanship sample does not.

First, it excludes young voters we fail to link to their parents or whose parents we fail to link to the Infutor address data. Second, the TargetSmart data primarily cover individuals who are registered on the voter rolls at least once during our sample period. That said, individuals who register to vote at some point are also included in years in which they were eligible to vote but unregistered. In addition, the data also include some individuals who never register, identified through commercial records provided by data aggregation firms. Table 1 shows that 80% of individuals in the Census sample are linked to the TargetSmart data.

To gauge the extent of nonrandom selection in the partisanship sample, Appendix Table A5 shows median county (Panel a) and census block group (Panel b) ACS characteristics using different weights: ACS estimates of counts of 18-to-29-year-old citizens, counts of young

[24] Appendix Table A4 reports summary statistics for the partisanship sample in all states, including those that do not record party affiliation.

Table 1: Summary Statistics for Permanent Residents and One-Time Movers

(a) Permanent residents

	Census sample			Partisanship sample		
	Mean	Std. dev.	Num. of obs.	Mean	Std. dev.	Num. of obs.
Birth year	1997	2.91	21,890,000	1997	2.89	6,394,724
Male	0.511	0.500	21,890,000	0.506	0.500	5,917,230
White	0.577	0.494	19,910,000	0.685	0.465	6,175,898
Black	0.130	0.336	19,910,000	0.089	0.284	6,175,898
Hispanic	0.213	0.409	19,910,000	0.174	0.379	6,175,898
Other race	0.080	0.271	19,910,000	0.052	0.222	6,175,898
Number of children in family	1.81	0.95	21,890,000	1.58	0.76	6,394,724
Linked to voter file	0.800	0.400	21,890,000	–	–	–
Registered in first election	0.575	0.494	21,890,000	–	–	–
Voted in first election	0.341	0.474	21,890,000	–	–	–
Republican in first election	–	–	–	0.186	0.389	6,394,724
Democrat in first election	–	–	–	0.257	0.437	6,394,724

(b) One-time movers

	Census sample			Partisanship sample		
	Mean	Std. dev.	Num. of obs.	Mean	Std. dev.	Num. of obs.
Birth year	1998	2.88	7,303,000	1997	2.87	2,176,791
Male	0.511	0.500	7,303,000	0.510	0.500	2,044,781
White	0.602	0.489	6,515,000	0.788	0.409	2,102,470
Black	0.129	0.335	6,515,000	0.062	0.240	2,102,470
Hispanic	0.177	0.382	6,515,000	0.105	0.306	2,102,470
Other race	0.091	0.288	6,515,000	0.046	0.210	2,102,470
Number of children in family	1.80	0.90	7,303,000	1.53	0.72	2,176,791
Age at move	10.8	6.77	7,303,000	11.4	7.87	2,176,791
Move distance (mi.)	321.0	552.0	7,297,000	327.2	601.6	2,176,791
Linked to voter file	0.793	0.406	7,303,000	–	–	–
Registered in first election	0.564	0.496	7,303,000	–	–	–
Voted in first election	0.325	0.468	7,303,000	–	–	–
Abs. exposure diff. in reg.	0.091	0.082	7,303,000	–	–	–
Abs. exposure diff. in voted	0.076	0.068	7,303,000	–	–	–
Republican in first election	–	–	–	0.201	0.400	2,176,791
Democrat in first election	–	–	–	0.235	0.424	2,176,791
Abs. exposure diff. in Rep.	–	–	–	0.086	0.076	2,176,791
Abs. exposure diff. in Dem.	–	–	–	0.096	0.082	2,176,791

Notes: The table reports summary statistics for the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and the partisanship sample. Panels a and b refer to, respectively, permanent residents (individuals whose parents never relocate across counties when the individuals are aged 0 through 26) and one-time movers (individuals whose parents relocate across counties exactly once over the same age range).

voters in the TargetSmart data, counts of young voters in the TargetSmart data linked to parents, and counts of young voters in the TargetSmart data linked both to parents and to Infutor. Especially at the county level, tabulated characteristics are remarkably similar whether we weight by ACS counts of 18-to-29-year-old citizens or by TargetSmart counts of young voters. By contrast, relative to the ACS, individuals in the linked samples (columns 3 and 4) live in places with slightly more white, high-income, and college-educated residents, as well as with lower population density. This pattern is consistent with white people and richer individuals having relatively higher registration rates (U.S. Census Bureau, 2021), which makes it easier to link young voters to parents registered at the same location in the TargetSmart data.[25]

We address concerns related to sample selection in Section 4.3, including by comparing turnout results from the Census sample to those obtained using the partisanship sample, exploiting the fact that this outcome can be studied in both samples. This comparison is informative because the two issues described above – undercoverage of unregistered individuals and imperfect parent-child links – are mitigated in the Census sample.

3 Empirical strategy

To estimate the effect of childhood environments, we measure the extent to which voters who moved with their families to a new county during childhood adopt political behavior similar to that of permanent-resident peers in that county. The stronger the place effect, the more closely movers' political behavior should resemble that of permanent residents. To identify these effects, we leverage variation in the timing of moves: we compare voters born in the same year, who move between the same origin-destination county pair, but spend different portions of childhood in the two counties. This empirical strategy follows Chetty and Hendren (2018a).

To fix ideas, focus on one outcome: whether an individual is affiliated with the Republican Party in the first general election in which they are eligible to vote. We define the exposure effect at age m as the impact of spending year m of childhood or adolescence in an area where the proportion of permanent residents affiliated with the Republican Party in their first eligible election is 1 percentage point higher.

The ideal experiment to identify this effect would be to randomly assign children to new counties starting at different ages for the rest of their childhood. For each assignment age m, we would then regress y_i, an indicator for whether mover i is affiliated with the Republican

[25]Appendix Figure A2 compares the partisanship sample to the broader population. Specifically, the figure reports estimates from bivariate regressions of the ratio between block group-level headcounts in the partisanship sample and ACS population counts on block group characteristics such as racial composition and household income.

Party in their first eligible election, on $\overline{y}_{ds}$, the average Republican Party affiliation among permanent residents of the same birth cohort s in destination county d:

$$y_i = \alpha + \beta_m \overline{y}_{ds} + \epsilon_i. \tag{1}$$

The exposure effect at age m would then be obtained by computing $\beta_m - \beta_{m+1}$. Of course, this experiment is not feasible, and real-world parents' residential choices likely correlate with children's potential outcomes: parents relocating to more Republican places may have a latent ideology θ_i that affects their children's likelihood to affiliate with the Republican Party even absent exposure to the destination county. Thus, estimating equation (1) in observational data would yield

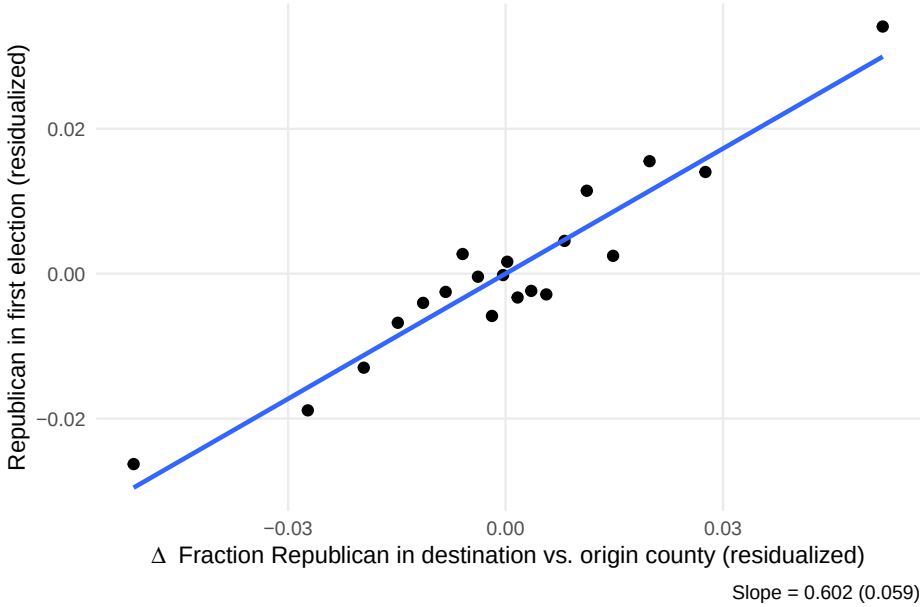
$$b_m = \beta_m + \delta_m,$$

where $\delta_m = \frac{Cov(\theta_i, \overline{y}_{ds})}{Var(\overline{y}_{ds})}$ is a selection term. The magnitude of δ_m is proportional to the covariance between parents' Republican ideology and the proportion of Republicans among permanent residents in the destination county. The more likely Republican-leaning families are to move to Republican places, the greater the selection.

To account for selection, we leverage variation in the timing of moves. We compare voters who spent different amounts of time growing up in similar counties because they moved there at different points in their childhood. In this setup, identification of exposure effects does not require that *where* families choose to move is unrelated to their children's potential outcomes. Instead, it only requires that the *timing* of the move is orthogonal to potential outcomes. Assuming that selection effects do not vary with the child's age at move, $\delta_m = \delta \ \forall m$, $\gamma_m = b_m - b_{m+1}$ is an unbiased estimate of the exposure effect at age m. Although this assumption is substantially weaker than assuming that location choice is unrelated to political ideology, it could still be violated. For example, parents with a stronger Republican ideology may be more likely to move to Republican places when their children are young or entering primary school. We return to validating the identification assumption in Section 4.2, after presenting our main results.

We now exploit this identification strategy to derive our baseline specification. We start by considering families who move when the child is six years old. In this subsample, we regress y_i, an indicator for whether child i registers as a Republican in the first eligible election, on $\Delta\overline{y}_{dos}$, the destination (d)-minus-origin (o) difference in the share of permanent residents from the same birth cohort s who register as Republicans in the first eligible election. The regressors also include origin-by-cohort fixed effects α_{os}, to control for common characteristics of voters born in the same year and coming from the same county of origin, and origin-by-destination fixed effects μ_{od}, to control for common characteristics of voters relocating along

Figure 1: Republican Party Affiliation in the First Eligible Election and County Exposure For Movers at Age Six



Notes: The figure plots the relationship between Republican Party affiliation and childhood context among movers who relocated at age six. The vertical axis shows the residualized likelihood that an individual who moved to a different county at age six registers to vote and affiliates with the Republican Party in their first eligible election. The horizontal axis shows the residualized destination-minus-origin difference in that outcome among same-cohort permanent residents. Both variables are residualized with respect to the fixed effects included in equation (2), namely origin-by-cohort fixed effects and origin-by-destination fixed effects. The black points show means for 20 bins with equal numbers of movers. The blue line is the fitted regression line and its slope corresponds to b_6 in equation (2). $N = 74{,}434$.

the same origin-destination county pair. We thus only exploit cross-cohort variation in the difference between origin and destination county:

$$y_i = \alpha_{os} + \mu_{od} + b_6 \Delta \overline{y}_{dos} + \epsilon_{1i}. \tag{2}$$

Figure 1 is a binned scatter plot of this relationship. Conditional on the fixed effects in equation (2), children who move at age six to a county where the share of same-cohort permanent residents registering as Republicans is 1 percentage point higher than in their origin county are 0.602 percentage points more likely to register as Republicans themselves.

As discussed above, b_6 does not have a causal interpretation because it combines exposure effects with selection. However, assuming that selection does not vary with age at move, differences between b_6 and analogous coefficients for other ages at move are causally inter-

pretable. We therefore repeat this exercise for children who moved at each age from 0 to 26.

We estimate these parameters simultaneously using the following regression:

$$y_i = \alpha_{osm} + \mu_{od} + \sum_{m=0}^{26} b_m(I_m = m)\Delta\overline{y}_{dos} + \sum_{s=1992}^{2001} \lambda_s(I_s = s)\Delta\overline{y}_{dos} + \epsilon_{2i}. \tag{3}$$

The coefficients of interest b_m are estimated in the first summation, which interacts indicators for age at move I_m with the destination-minus-origin difference in permanent residents' average outcomes $\Delta\overline{y}_{dos}$. Since we include origin-by-destination fixed effects (μ_{od}), our exposure effect estimates compare individuals who move between the same origin-destination county pair, but at different ages.

We note two differences relative to equation (2). First, we interact cohort indicators I_s with the exposure term $\Delta\overline{y}_{dos}$, omitting the 2002 cohort as the reference category, so that the b_m coefficients are identified purely from within-cohort variation. We could not include these terms in equation (2), where they would be collinear with $\Delta\overline{y}_{dos}$.[26] Second, equation (3) replaces origin-by-cohort fixed effects, α_{os}, with origin-by-cohort-by-age-at-move fixed effects, α_{osm}. By controlling flexibly for age at move, we account for the possibility that moving at different ages generates different disruptions that are independent of the characteristics of the origin or destination county. Our main regressions use heteroskedasticity-robust standard errors, since the number of years spent in the counties of origin and destination is determined at the individual level.

Under our identifying assumption that the unobserved determinants of children's probability of registering as a Republican are uncorrelated with age at move, $\gamma_m = b_m - b_{m+1}$ captures the causal effect of one additional year of exposure at age m. Specifically, it measures the effect on a voter's likelihood of registering as a Republican of spending that year in a destination county where same-cohort permanent residents are 1 percentage point more likely to register as Republicans than same-cohort permanent residents in the origin county.

[26] The inclusion of these terms controls for differential measurement error across cohorts. For instance, our ability to observe moves at early ages differs across cohorts because the quality of our address history data varies over time. Furthermore, for the youngest cohorts in the sample, we cannot observe moves at older ages, since our data stop in 2021.

4 Results

4.1 Main results

4.1.1 Graphical analysis

We now report our main results on the effects of childhood location on young voters' political behavior in their first eligible election. We begin with a graphical analysis: we plot the b_m coefficients from estimating equation (3) on the sample of one-time cross-county movers who relocated between ages 0 and 26. Differences between adjacent coefficients from ages 0 to 19 indicate how one additional year spent in the destination county changes the correlation between movers' behavior and the outcomes of that county's permanent residents. Since elections take place every other year, an individual's first eligible election occurs at age 18 or 19. For voters aged 20 or older when their parents move, first-eligible election outcomes are measured before the move and, by construction, they cannot be affected by the destination county. We should therefore not expect meaningful changes across the coefficients for moves after age 20. However, the *level* of these coefficients remains informative: in the absence of any destination effect, this level reflects the selection term δ discussed in Section 3.

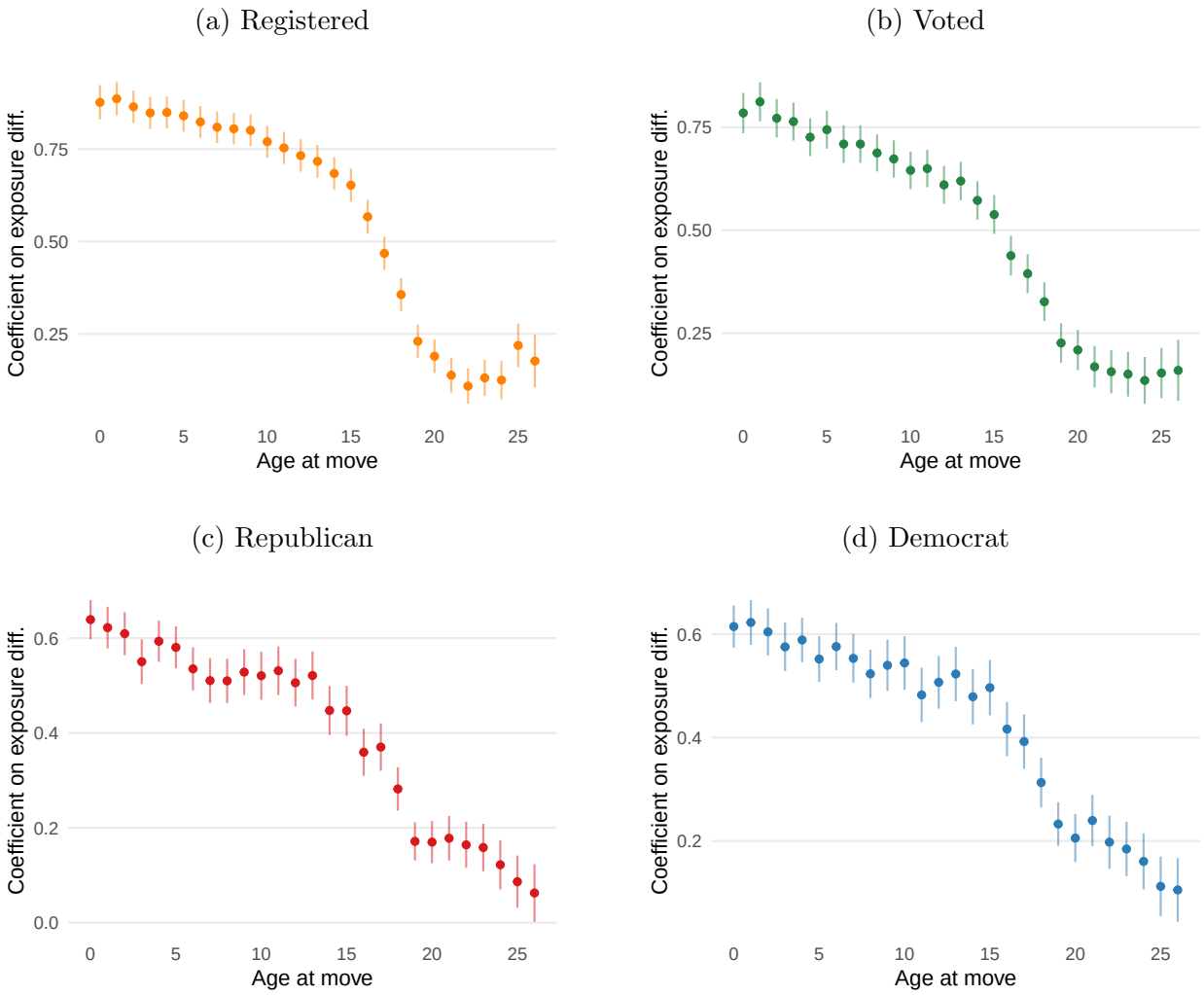
The results are shown in Figure 2. Each panel corresponds to a different outcome: registration (Panel a), turnout (Panel b), Republican Party affiliation (Panel c), and Democratic Party affiliation (Panel d).

All plots exhibit three key features. First, all coefficients are positive and statistically significant, including those for moves at ages 20 and above. For instance, in Panel a, $b_{20} = 0.189$ indicates that voters moving at age 20 to a place where permanent residents are 10 percentage points more likely to register to vote in their first eligible election are themselves 1.89 percentage points more likely to register. Since these voters' outcomes predate moving, the positive coefficients for moves after age 20 indicate selection: parents who move to areas with higher registration rates have unobserved characteristics that make their children more likely to register in adulthood. The same is true of parents moving to areas with higher turnout and to more Democratic or Republican areas. Our empirical strategy controls for this selection effect by comparing the magnitude of coefficients across ages at move.

Second, in all four panels, b_m decreases nearly monotonically as m increases from 0 to 20. In other words, children who move at later ages – and thus spend less time in the destination county – are less likely to mirror the behavior of same-cohort permanent residents in the destination than children who move at younger ages. This pattern is consistent with childhood places affecting people's political behavior in adulthood, beyond the influence of their family and innate characteristics.

While differences between adjacent coefficients identify the effect of one additional year of exposure to the destination county, the difference between b_0 and b_{20} captures the total effect

Figure 2: Nonparametric Estimates of Childhood Exposure Effects



Notes: The figure plots age-specific estimates of b_m, the coefficient on the origin-destination exposure difference for movers at age m, from equation (3). All outcomes are measured as of the first election in which an individual is eligible to vote. Vertical lines show 95% confidence intervals. Registration and turnout results use data from the Census sample (N = 7,303,000, FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample (N = 2,176,791).

of childhood environment. For registration (Panel a), we find that $b_0 = 0.876$ and $b_{20} = 0.189$, so this total effect is 0.687, implying that spending one's entire childhood in a county where same-cohort permanent residents are 10 percentage points more likely to register increases one's own probability of registering by 6.87 percentage points. The corresponding effect for turnout in the first eligible election is 5.75 percentage points. Assuming that these impacts also apply to non-movers, a claim we return to in Section 4.4, the effect of childhood location drives 69% of the cross-county variation in registration and 58% of the cross-county variation in turnout among young adults.

The influence of childhood places on partisanship is also very strong. As with registration and turnout, the coefficients in Panels c and d are positive and decline monotonically with age at move until age 20. The total effects of childhood places on Republican and Democratic Party affiliation are 0.469 and 0.409, which implies that spending one's entire childhood in a county where permanent residents are 10 percentage points more likely to affiliate with the Republican or Democratic Party increases one's own probability of affiliating with the corresponding party by 4.69 and 4.09 percentage points. These results imply that 40 to 50% of the cross-county variation in party affiliation among young registered adults is explained by the effect of where they grew up.

Third, while the b_m parameters decline with age at move throughout childhood, the decrease is more pronounced during the teenage years, suggesting that an additional year of exposure has a larger effect during adolescence than earlier in childhood.

4.1.2 Parametric estimates

To test formally whether exposure effects differ between childhood and adolescence, we now estimate separate linear trends in age at move for three age ranges: 12 and below, 13 to 19, and 20 and above.

Formally, we estimate the following regression:

$$\begin{aligned} y_i =& \alpha_{osm} + \mu_{od} + \sum_{s=1992}^{2001} \lambda_s (I_s = s) \Delta \overline{y}_{dos} + I(m_i \leq 12)(\zeta' - \zeta m_i) \Delta \overline{y}_{dos} \\ &+ I(12 < m_i \leq 19)(\gamma' - \gamma m_i) \Delta \overline{y}_{dos} + I(m_i > 19)(\delta' - \delta m_i) \Delta \overline{y}_{dos} + \epsilon_{3i}. \end{aligned} \tag{4}$$

This specification includes the same controls as equation (3) but restricts the relationship between age at move m_i and the destination-minus-origin exposure difference to be piecewise linear.[27]

Table 2 reports estimates of the average yearly exposure effects ζ, γ, and δ. One additional

[27] When we need to parsimoniously compare yearly exposure effects over ages at move 0–19 across different subsamples, we estimate the following specification:

year of exposure between ages 0 and 12 to a county where the registration rate among same-cohort permanent residents is 1 percentage point higher increases a person's probability of registering in the first eligible election by 0.012 percentage points (column 1). The analogous estimate for turnout is 0.015 percentage points (column 3). Consistent with the patterns in Figure 2, yearly exposure effects between ages 13 and 19 are much larger: 0.082 percentage points for registration and 0.064 percentage points for turnout.

Exposure effects on party affiliation are also larger during adolescence than earlier in childhood: 0.054 versus 0.011 percentage points for Republican Party affiliation (column 5), and 0.048 versus 0.010 percentage points for Democratic Party affiliation (column 7). Differences between exposure effects in childhood and adolescence are statistically significant at the 1% level for all four outcomes. These results have two possible interpretations: either where people live during adolescence affects their political behavior in adulthood substantially more than where they live earlier in childhood, or there are diminishing returns to additional years of exposure. We return to this question in Section 5.

4.2 Robustness checks: differential age selection effects

4.2.1 Time-invariant family characteristics

The causal interpretation of our results rests on the assumption that move timing is uncorrelated with factors affecting young voters' political behavior other than location exposure. This assumption would be violated if families' latent political ideology were correlated with children's age at move, e.g., if more Republican families moved toward more Republican places when their children were younger. In this case, differences across the b_m coefficients would conflate the causal effect of place with differences in the partisanship of families who move when their children are different ages.

Comparing siblings To relax this assumption, we restrict both the Census and partisanship samples to young voters with at least one sibling and add family fixed effects to equation (4).[28] Consider, for instance, a family with two children that moves across counties when one child is 12 and the other is 15. The younger child therefore spends three more years in the destination before becoming eligible to vote. The family fixed effects specification identifies exposure effects from this within-family difference alone, while holding all time-invariant family characteristics constant, including latent political ideology.

$$\begin{aligned} y_i =& \alpha_{osm} + \mu_{od} + \sum_{s=1992}^{2001} \lambda_s (I_s = s) \Delta\overline{y}_{dos} + I(m_i \leq 19)(\zeta' - \zeta m_i)\Delta\overline{y}_{dos} \\ &+ I(m_i > 19)(\delta' - \delta m_i)\Delta\overline{y}_{dos} + \epsilon_{4i}. \end{aligned} \tag{5}$$

Appendix Table A6 reports full-sample estimates from this specification.

[28] Since counties of origin and destination are fixed within families, the family fixed effects specification technically replaces origin-by-destination fixed effects with family fixed effects.

Table 2: Parametric Estimates of Childhood Exposure Effects

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure effect, 0–12	0.0120***	0.0167***	0.0151***	0.0177***	0.0112***	0.0034	0.0100***	0.0022
	(0.0008)	(0.0020)	(0.0010)	(0.0021)	(0.0015)	(0.0058)	(0.0015)	(0.0058)
Exposure effect, 13–19	0.0817***	0.0837***	0.0643***	0.0635***	0.0543***	0.0461***	0.0483***	0.0583***
	(0.0021)	(0.0039)	(0.0025)	(0.0044)	(0.0038)	(0.0101)	(0.0041)	(0.0107)
Exposure effect, 20–26	0.0022	0.0051	0.0101**	0.0090	0.0175***	0.0045	0.0194***	0.0201*
	(0.0037)	(0.0068)	(0.0042)	(0.0071)	(0.0042)	(0.0109)	(0.0044)	(0.0111)
Difference $[13, 19]$ vs. $[0, 12]$	0.070	0.067	0.049	0.046	0.043	0.043	0.038	0.056
p-value $[13, 19] = [0, 12]$	0.0000	0.0000	0.0000	0.0000	0.0000	0.0005	0.0000	0.0000
Family fixed effects		✓		✓		✓		✓
Number of observations	7,303,000	4,092,000	7,303,000	4,092,000	2,176,791	894,201	2,176,791	894,201
Dependent variable mean	0.564	0.561	0.325	0.318	0.201	0.219	0.235	0.224
R^2	0.228	0.642	0.242	0.661	0.277	0.753	0.219	0.712

Notes: The table reports yearly childhood exposure effects, measuring the effect of spending an additional year of childhood in a place where the permanent residents are 1 percentage point more likely to register (columns 1–2), vote (columns 3–4), affiliate with the Republican Party (columns 5–6), or affiliate with the Democratic Party (columns 7–8). Estimates are based on equation (4). All outcomes are measured in the first election in which an individual is eligible to vote. In even-numbered columns, the underlying regressions control for family fixed effects, thus comparing siblings who make the same move at different ages and therefore have different amounts of exposure to the destination county. Each column also reports the difference between the estimates for ages 0 to 12 and ages 13 to 19, as well as p-values testing the null hypothesis that the two estimates are equal. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample.

We show the results in Table 2, columns 2, 4, 6, and 8. Since the samples are restricted to families with multiple children, they are substantially smaller than those used for the main results. Nonetheless, across all outcomes, the point estimates for exposure effects until age 19 are very similar to the main estimates. Moreover, estimated exposure effects after age 20 are small and statistically non-significant across all outcomes, except for Democratic Party affiliation, for which the estimate is marginally significant. For Republican and Democratic Party affiliation, the estimates for ages 0 to 12 are also non-significant.

As shown in Appendix Figure A3, non-parametric estimates that control for family fixed effects exhibit declining patterns similar to those in Figure 2. In Appendix Table A7, we further control for differences in sample composition by estimating specifications both with and without family fixed effects on the subsample of children with one or more siblings. The resulting estimates are remarkably similar, suggesting that time-invariant family characteristics, which are controlled for only in the family fixed effects specifications, are unlikely to bias our full-sample estimates.

4.2.2 Time-varying shocks

The previous section shows that our findings are robust to controlling for time-invariant family characteristics that may be correlated with both children's age at move and their political behavior in early adulthood. However, the family fixed effects specification does not address the possibility that *time-varying* shocks influence both parents' relocation decisions and children's subsequent political behavior. For example, consider a couple with two children that becomes richer and moves to a more Republican area. Relative to the older sibling, the younger child spends more time not only in the destination county, but also in a richer family environment, making it difficult to isolate the place effect.

To address this concern, we present a series of overidentification tests, reported in Figure 3 and Appendix Table A8. We test whether movers' outcomes are better predicted by the outcomes of permanent residents from similar cohorts than by those of older or younger cohorts, and, in a similar spirit, we exploit variation in permanent residents' outcomes across gender and race within cohort-by-county cells. The logic of these tests is that, if places shape political behavior, movers' outcomes should converge more strongly to those of permanent residents who are most similar to them in age, gender, and race. Indeed, these permanent residents are more likely to experience the local context similarly as them and to influence them through personal interactions. By contrast, unobserved time-varying shocks would not generate this specific pattern of convergence. For example, family income shocks correlated with the timing of the move should not cause male movers' outcomes to converge more strongly to those of male permanent residents than to those of female permanent residents. For tractability, in these analyses, we report exposure effect estimates over ages at move 0–19 based on equation 5.

Cohort For each mover, we first construct exposure differences using the outcomes of permanent residents in three cohort groups: those born two to four calendar years before the mover, those from the mover's cohort or born within one calendar year of it, and those born two to four calendar years after the mover. We then replace the exposure measures $\Delta\overline{y}_{dos}$ in equation (5) with these three measures, $\Delta\overline{y}_{dos}^{-4 \text{ to } -2}$, $\Delta\overline{y}_{dos}^{-1 \text{ to } +1}$, and $\Delta\overline{y}_{dos}^{+2 \text{ to } +4}$. By including the three exposure measures in the same regression, we let exposure based on nearby cohorts compete with exposure based on older and younger cohorts in predicting movers' outcomes.

Figure 3, Panel a shows that the exposure effects are concentrated on the outcomes of permanent residents from nearby cohorts. By contrast, the coefficients associated with older and younger cohorts are substantially smaller and often non-significant. This pattern is especially striking because permanent residents' outcomes are highly correlated across adjacent cohorts.[29] To explain this pattern, unobserved time-varying confounders would have to cause parents to move not merely to counties where children in general grow up to be Republicans, but to counties where children *born around the same time as their own child* grow up to be Republicans. Since registration, turnout, and party affiliation are only realized later in life, once children become eligible to vote, it is unlikely that families can accurately predict the evolution of political geography and base relocation decisions on this information.

Gender We then compute average outcomes within county-by-cohort cells separately for permanent residents of the mover's gender and of the opposite gender, and we replace the exposure measures $\Delta\overline{y}_{dos}$ in equation (5) with $\Delta\overline{y}_{dos}^{g}$ and $\Delta\overline{y}_{dos}^{-g}$. As shown in Figure 3, Panel b, movers' behavior converges substantially more toward outcomes of same-gender permanent residents than toward those of opposite-gender permanent residents. As in the cross-cohort comparisons, this pattern alleviates concerns about time-varying shocks simultaneously affecting parents' location decisions and children's future political outcomes.

Race Finally, we compute average outcomes within county-by-cohort cells separately for permanent residents of the mover's race and for permanent residents of other races, and we replace $\Delta\overline{y}_{dos}$ in equation (5) with $\Delta\overline{y}_{dos}^{r}$ and $\Delta\overline{y}_{dos}^{-r}$.[30] Once again, Figure 3, Panel c shows that outcomes of same-race permanent residents are much better predictors of movers' political behavior than outcomes of other-race permanent residents. Overall, the results from these overidentification tests alleviate concerns that time-varying shocks simultaneously

[29]The population-weighted autocorrelation between adjacent cohorts within counties is 0.885 for Republican Party affiliation and 0.852 for Democratic Party affiliation. Similarly, permanent residents' outcomes within county-by-cohort cells are strongly correlated across genders. The two-way correlation across genders is 0.845 for Republican Party affiliation and 0.870 for Democratic Party affiliation. Within county-by-cohort cells, permanent residents' outcomes are less correlated across races: the two-way correlation, comparing individuals of one race to individuals of all other races, is -0.081 for Republican Party affiliation and 0.223 for Democratic Party affiliation.

[30]We use racial categories of non-Hispanic white, non-Hispanic Black, Hispanic, and other non-Hispanic.

affecting parents' location decisions and children's future political outcomes could bias our estimates.

4.3 Robustness checks: sample selection bias

A second threat specific to our party affiliation results is sample selection bias. This bias could arise because the partisanship sample is mostly limited to individuals who appear registered to vote in at least one TargetSmart snapshot, and because the analysis further restricts that sample to individuals who can be linked to parents and whose parents' address histories can be recovered in the Infutor data. These concerns do not apply to the registration and turnout estimates from the Census sample. In what follows, we first provide high-level evidence that these two concerns are unlikely to bias our results, and then conduct more direct tests addressing each of them.

Effects on turnout in the Census and partisanship samples Since we observe turnout in both the Census and partisanship samples, we first check whether estimated exposure effects on that outcome differ across the two samples. Appendix Figure A4 compares nonparametric turnout effects across the Census and partisanship samples. Panel a uses the baseline specification with origin-by-destination fixed effects and Panel b uses the more restrictive specification with family fixed effects. In both panels, the estimates for the two samples exhibit the same declining pattern and overall shape. They also share similar levels: although the Census sample estimates are somewhat larger for earlier ages at move (e.g., $b_0 = 0.785$ versus 0.639 in Panel a), the two samples produce nearly identical estimates during adolescence (e.g., $b_{15} = 0.538$ versus 0.528). Further corroborating this similarity, the parametric estimates corresponding to equation (4), shown in Appendix Table A9, are very similar, especially for ages at move 13–19. These results provide a first piece of evidence suggesting that the two potential sources of sample selection bias in the partisanship sample are unlikely to drive our results. We now discuss these two sources in more detail.

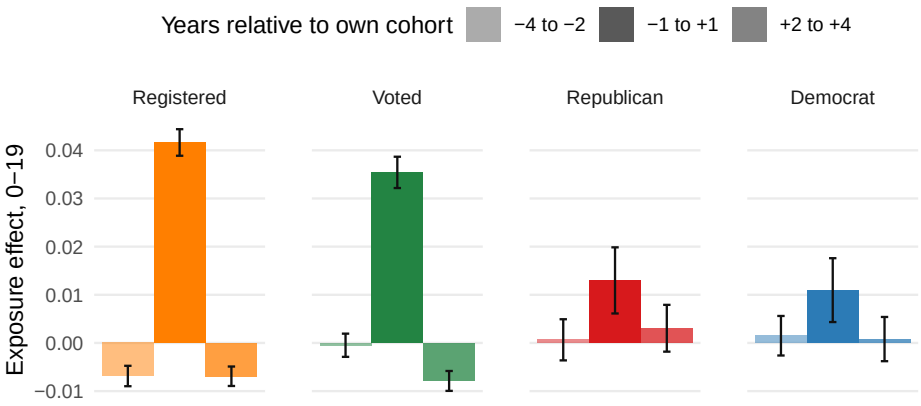
Focus on registered individuals The partisanship sample is mostly limited to individuals who appear registered to vote in at least one TargetSmart snapshot between 2012 and 2021.[31] If location exposure affects the likelihood of registering to vote, and therefore of appearing in the partisanship sample, our estimated effects on party affiliation may be biased.[32] On the one hand, politically disengaged voters who would otherwise remain unregistered may register – and therefore enter the sample – if they spend time in a location that increases

[31] In Appendix Table A10, we show that excluding individuals who never appear as registered voters, and were thus likely included by TargetSmart based on other datasets, yields virtually identical results to those in Table 2.

[32] Consistent with this concern, previous research shows that voter file analyses can be sensitive to differential post-treatment registration (Nyhan et al., 2017).

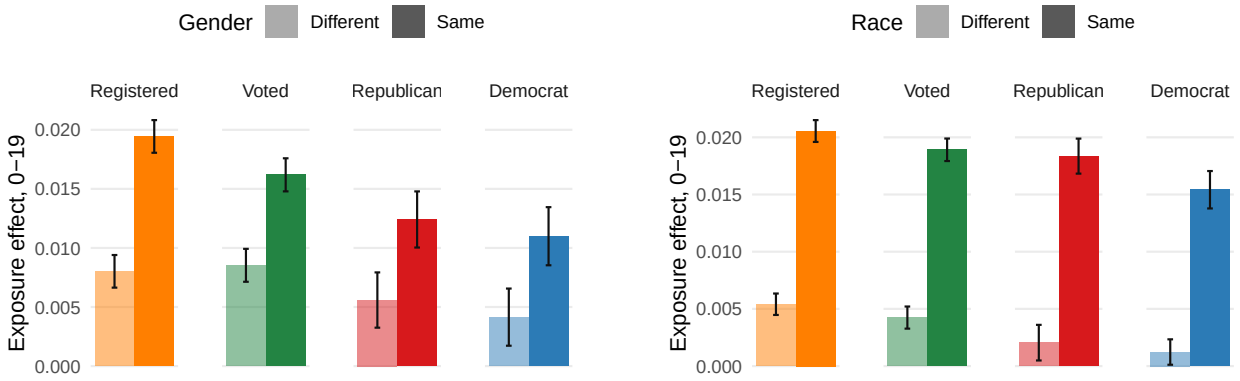
Figure 3: Overidentification Tests

(a) Cohort



(b) Gender

(c) Race



Notes: The figure shows overidentification tests that modify the exposure measures from the baseline specification. In Panel a, we replace the cohort-specific exposure measure in equation (5) with exposure measures based on three cohort groups of permanent residents: those born two and four calendar years before the mover, those from the mover's cohort or born within one calendar year of it, and those born two to four calendar years after the mover. In Panel b, we use separate exposure measures specific to permanent residents of the mover's gender and to permanent residents of the opposite gender. In Panel c, we use separate exposure measures specific to permanent residents of the mover's race and to permanent residents of other races, where race is collapsed into four categories: non-Hispanic white, non-Hispanic Black, Hispanic, and other non-Hispanic. We exclude voters with missing race information. Vertical lines show 95% confidence intervals. Appendix Table A8 reports these coefficient estimates along with standard errors. In Panel a, $N = 5{,}097{,}000$ for registration and turnout and $N = 1{,}539{,}043$ for party affiliation. In Panel b, $N = 7{,}301{,}000$ for registration and turnout and $N = 2{,}172{,}145$ for party affiliation. In Panel c, $N = 6{,}460{,}000$ for registration and turnout and $N = 1{,}972{,}105$ for party affiliation. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample.

electoral participation, biasing our estimates downward. That is, a county that fosters political engagement may draw into the electorate marginal voters who are less interested in politics and thus less likely to affiliate with a major party conditional on registration. On the other hand, counties leaning toward, say, the Republican Party may discourage Democratic-leaning voters from registering. Because these voters would not appear in the sample, our estimates would overstate the true size of location exposure effects on party affiliation.

We note, first, that the fact that 80% of individuals in the Census sample are linked to the TargetSmart data limits the scope for bias from this undercoverage. Second, we test whether exposure effects vary with voter registration rates in movers' destination counties. To do so, we group counties in party-affiliation states into quartiles based on registration rates among individuals aged 18 to 29,[33] and estimate exposure effects separately for movers relocating to destination counties in different quartiles. If conditioning the sample on voters who register to vote at least once biased our main estimates, we would expect effect sizes to vary systematically across the four quartiles.

Figure 4 shows no systematic or statistically significant variation in effect sizes for either Democratic or Republican affiliation. We interpret these results as evidence that restricting the partisanship sample to individuals who register at least once is not a major source of bias.

In addition, because Democratic and Republican Party affiliation are set to zero if an individual is not registered, as indicated in Section 2.3, effects on these outcomes could in principle partly reflect effects on registration. The fact that the large majority of individuals in the Census sample register at some point in time and that the effects on party affiliation remain similar in the highest-registration quartile, where there is less margin for registration to be affected, suggests that they rather capture effects on ideology.

Parent-child linkage in the partisanship sample Because party-affiliation analyses are limited to young people we successfully link to parents, they are subject to an additional source of potential selection bias: location exposure may affect the likelihood of moving away from home, for example to attend college, and, thus, of entering the partisanship sample. Young voters who leave home to attend college may be less likely to be observed at the same address as their parents.

To address this concern, we focus on voters who turn 18 shortly before Election Day. Many of these voters will register before the election and will therefore appear in the voter file at a younger age, on average, than voters who turn 18 shortly after Election Day and whose first eligible election is nearly two years afterward. This comparison is useful because young

[33] To construct these registration rates, we divide county-level TargetSmart counts of registered voters aged 18 to 29 by county-level ACS counts of same-age citizens. We choose the age range 18 to 29 rather than 18 to 26 because the ACS only reports these statistics for the former.

people are more likely to live with their parents at age 18 than at ages 19 or 20. As a result, movers who turn 18 shortly before an election should be more likely to register at the same address as their parents, increasing the likelihood that they appear in our linked sample regardless of their destination county, and reducing the risk of endogenous sample selection.

Figure 5, Panel a shows that, as expected, the partisanship sample contains substantially more people who turn 18 right before a general election rather than right after. We first count voters by the day on which they turn 18 relative to the 2012 election, focusing on a 100-day bandwidth around Election Day. We then repeat this exercise for all general elections through 2018 and stack the resulting voter counts.[34] Negative and positive values of the horizontal axis denote voters whose 18th birthday falls before or after the nearest Election Day, respectively. As expected, turning 18 right before an election greatly increases the number of voters (and thus the implied probability of) appearing in the partisanship sample.

We find little evidence that imperfect parent-child linkage drives our results, as estimates are largely unchanged when we focus on voters who turn 18 shortly before a general election. Figure 5, Panel b and Appendix Table A11 report exposure effects estimated on the subsample of voters who turned 18 in the 100 days before a general election between 2012 and 2020, and compare these estimates to those from the full partisanship sample.[35] If imperfect parent-child links biased our results, we would expect effects on the full sample to differ substantially from those for voters who turned 18 right before a general election. Reassuringly, point estimates for Republican and Democratic Party affiliation in the subsample and the full partisanship sample are virtually indistinguishable.

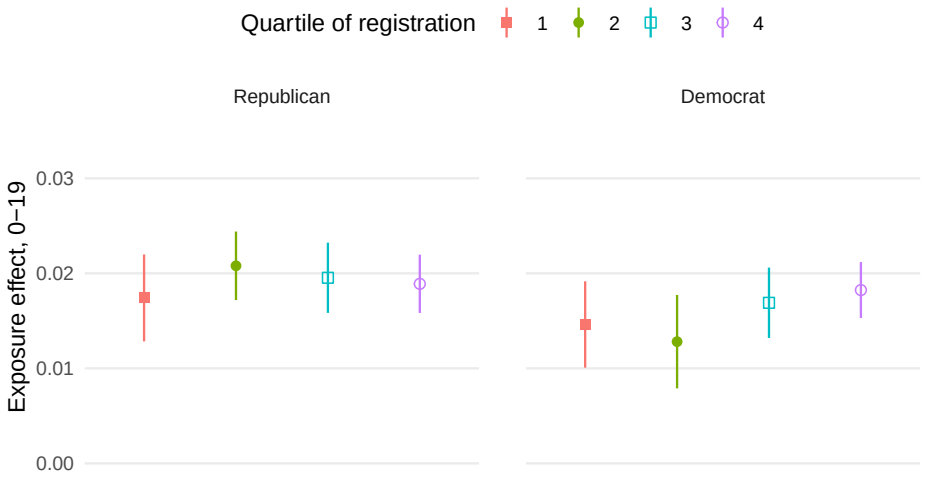
4.4 External validity

Setting aside internal validity concerns, the partisanship sample may still be unrepresentative of the broader U.S. population. As discussed in Section 2.4 and shown in Appendix Table A5, the partisanship sample differs from the U.S. voting-eligible population along some observable characteristics. For instance, the partisanship sample slightly over-represents census block

[34]We do not include the 2020 general election in this exercise because our sample does not include voters who are age-ineligible to vote in 2020.

[35]For this exercise, we exclude voters reported as born on the first of the month, as those birth dates could be accurate only to the month or year. Indeed, as discussed in Section 2.1, when the exact date of birth is unknown but the month and year are known, a voter's birthdate may be listed as the first of that month, and if only the birth year is known, the birth date field may indicate January 1st of that year. In Appendix Figure A5, we further restrict the sample to individuals born after the state-specific school entry cutoff, where available, to focus on individuals who are likely still in high school at the time of their first eligible election. We expect these individuals to be even more likely to live with their parents and, in turn, more likely to be linked to them. In this sample, we again obtain similar estimates, although precision is limited by the small sample size.

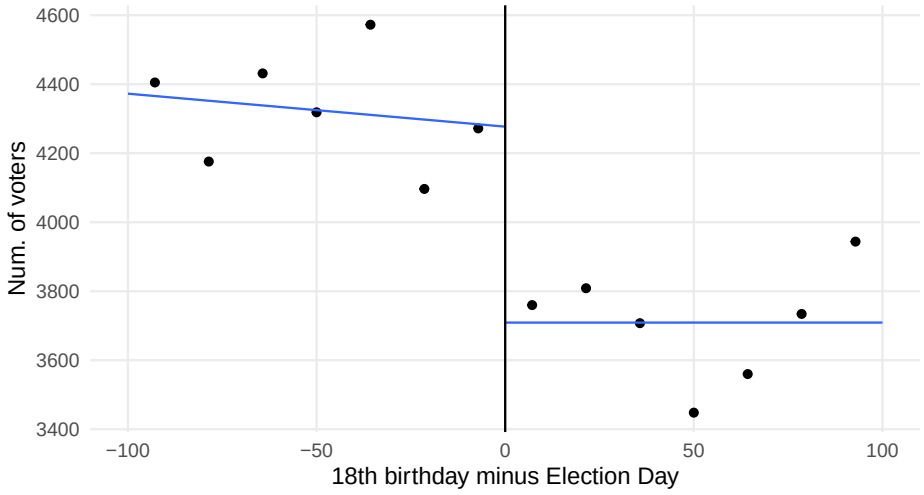
Figure 4: Party Affiliation Exposure Effects by Quartile of Voter Registration in the Destination County



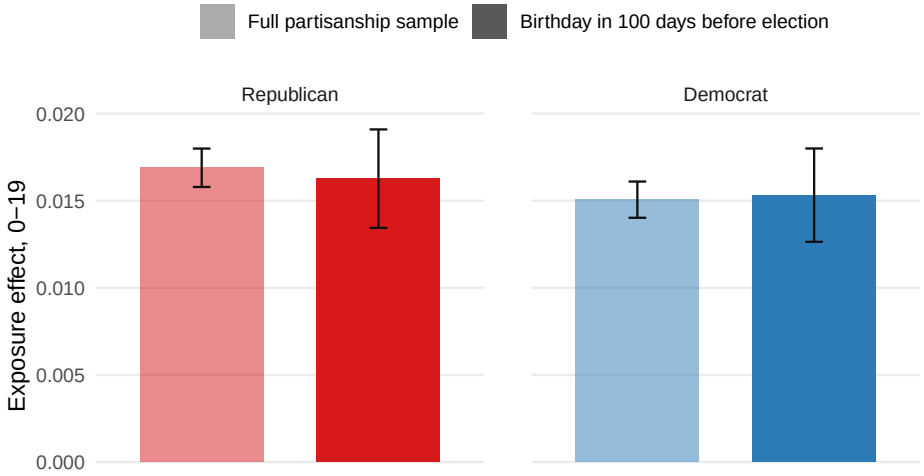
Notes: The figure plots parametric estimates of yearly exposure effects separately by quartile of voter registration among individuals aged 18 to 29 in the mover's destination county. To compute county registration rates, we divide county-level TargetSmart counts of registered voters aged 18 to 29 by county-level 2015–2019 ACS counts of same-age citizens (source table: Citizen, Voting-Age Population by Age). We choose the age range 18 to 29 rather than 18 to 26 because the ACS only reports these statistics for the former. Q1 and Q4 correspond to counties with the lowest and highest registration rates, respectively. Vertical lines show 95% confidence intervals. N = 2,173,988.

Figure 5: Party Affiliation Exposure Effects for Voters Newly Eligible Before Election Day

(a) Number of voters in the partisanship sample by birth date



(b) Exposure effects: full partisanship sample vs. specific birthday sample



Notes: Panel a shows counts of voters in the partisanship sample who turned 18 within a bandwidth of 100 days to the nearest general Election Day, 2012–2018. Negative values of the x-axis represent individuals who turned eighteen shortly before Election Day and who were therefore barely eligible to vote. By contrast, positive values represent individuals who turned eighteen shortly after Election Day and who therefore had to wait almost two years for their first eligible election. Panel a excludes the 2020 election because our voter file data end in 2021, meaning we do not observe individuals who turned 18 shortly after the 2020 election in any subsequent election. N = 890,438. Panel b plots parametric estimates of exposure effects on Republican and Democratic Party affiliation separately for the full partisanship sample and for the sample of voters who turned 18 in the 100 days before the closest general election. Vertical lines show 95% confidence intervals. For the full sample, N = 1,734,048. For the restricted sample, N = 244,821.

groups with high median incomes and those with large shares of white residents.

To address external validity concerns, we weight movers by the ratio of county-level ACS counts of citizens aged 18 through 29 to corresponding county-level counts of movers and permanent residents in the partisanship sample. We construct these weights separately using each mover's origin and destination counties. Appendix Figure A6 shows that, for both Republican and Democratic affiliation, parametric regression estimates using these weights are virtually identical to our baseline unweighted estimates.

4.5 Effect persistence

Our main results document the effects of childhood environment on political behavior in a voter's first eligible election. We now ask whether this influence persists in later elections. In the first eligible election, the influence of childhood locations may be attenuated by family background and genetic factors. In later elections, this influence may be further attenuated by post-childhood experiences, such as where individuals live and the friends they make as adults. We should therefore expect the effect of childhood location to decline over time.

Figure 6 and Appendix Table A12 present the persistence results. To ease comparisons, all estimates are based on the subsample of people whose outcomes are observed in each of the first three elections after they turn 18. Although exposure effects shrink slightly over time, they remain significant at the 1% level in the second and third eligible elections for all outcomes.

For registration, effects in the second eligible election are 75% as large as those in the first, while effects in the third eligible election are 69% as large. Turnout effects in the second and third elections are 76% and 80% as large as those in the first. Party affiliation outcomes exhibit even more persistence: exposure effects on Republican or Democratic Party affiliation in the second eligible election are 93% and 97% the size of those in the first, respectively. By the third eligible election, exposure effects on party affiliation remain 89% as large as those in the first election.

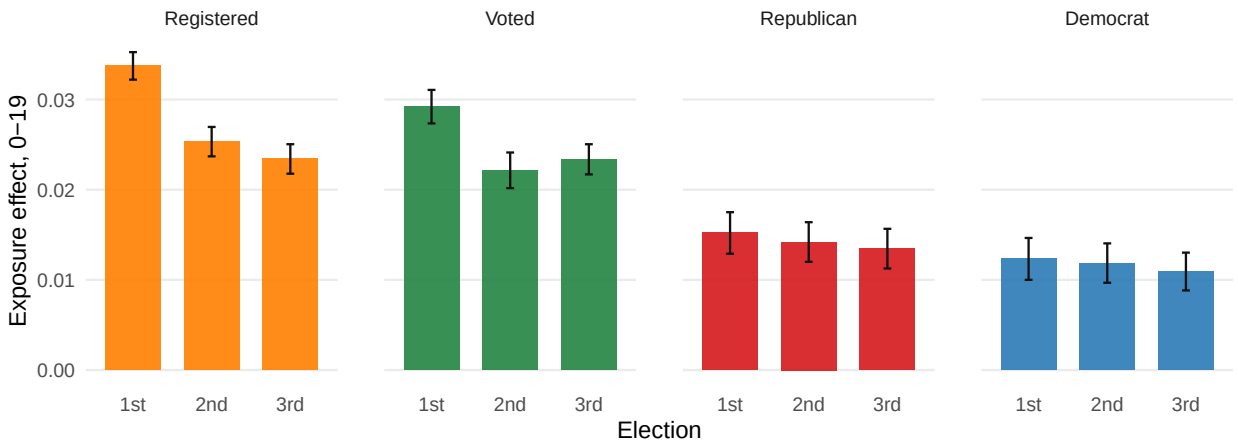
Thus, where one grows up influences political behavior even several years into adulthood, after people have experienced important events such as attending college and starting their first job.

5 Formative years

5.1 Empirical strategy

The literature on formative years posits that experiences at certain ages are especially consequential for later-life outcomes (e.g., Deutscher, 2020; Goldman and Hopkins, 2020; Ghitza

Figure 6: Persistence of Childhood Exposure Effects



Notes: The figure shows parametric estimates of exposure effects in the first, second, and third election for which an individual is eligible to vote. The sample is restricted to individuals for whom we can observe all three elections, that is, for birth cohorts 1992 to 1998. Vertical lines show 95% confidence intervals. Appendix Table A12 reports these coefficient estimates along with standard errors. Registration and turnout results use data from the Census sample (N = 4,194,000, FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample (N = 1,331,139).

et al., 2023; Neundorf and Smets, 2017; Pacheco, 2008; Sears and Valentino, 1997). In our setting, adolescence may be particularly consequential, since adolescents are generally more interested in politics than younger children, which may make them more susceptible to the influence of their environment (Janmaat and Hoskins, 2021). In addition, political events and experiences encountered during adolescence may be especially salient as voters approach their first opportunity to register and vote. The larger exposure effects between ages 13 and 19 than earlier in childhood, shown in Figure 2 and Table 2, are consistent with this hypothesis.

However, stronger effects during adolescence could also reflect diminishing marginal returns to exposure: an additional year in a county may matter more for children who have spent less time there. Indeed, an additional year of exposure after age 13 is estimated from movers with relatively short exposure to the destination, whereas an additional year in early childhood is estimated from movers with longer total exposure.

We cannot use one-time movers to distinguish between formative-years effects and diminishing marginal returns to exposure, because one-time movers' age at move and, thus, the ages at which they live in a given area, are perfectly collinear with length of exposure. For example, an individual who spends four years in the destination before their first eligible election at, say, age 19 must have lived there at age 15. By contrast, an individual with only

three years of exposure could not have lived in that destination at age 15, as they would have arrived at age 16. As a result, differences between the age-15 and age-16 coefficients in Figure 2 could reflect either the effect of increasing exposure length from three to four years or the effect of one year of exposure specifically at age 15, which may be particularly formative.

To overcome this limitation, we shift attention to individuals who move across counties two or three times between ages 0 and 26. Using this sample of "multiple movers," we assess whether exposure spells of a given length are more impactful during the teenage years than at earlier ages.[36] For example, we compare two individuals who spend four years in the same first destination, one during childhood and the other during adolescence, and see whether the latter behaves more similarly to that destination's permanent residents in the first eligible election.

Formally, we estimate the following specification:

$$\begin{aligned} y_i = \alpha_{os} + \sum_{a \in [0-12, 13-19]} I(a_{i1} = a) b_a \Delta \overline{y}_{dos_1} + \sum_{e_1=2}^{6} I(e_{i1} = e_1) \mu_{e_1} \\ + \sum_{j=2}^{3} \sum_{e_j=0}^{17} I(e_{ij} = e_j)(\mu_{e_j} + \nu_{e_j} \Delta \overline{y}_{dos_j}) + \epsilon_{5i}, \end{aligned} \tag{6}$$

where outcomes are defined as in the one-time movers specifications. The main regressor is the destination-minus-origin difference in permanent residents' average outcomes for the *first* move, $\Delta \overline{y}_{dos_1}$, which we interact with indicators for whether individual i lived in the first destination during childhood, ages 0–12, or adolescence, ages 13–19.[37]

The coefficients of interest, b_a, measure the extent to which movers who live in the first destination during age range a – childhood or adolescence – behave similarly to permanent residents in that destination, controlling for permanent residents' outcomes in the origin. Like in the one-time movers design, the individual b_a coefficients combine exposure effects with selection. However, assuming that the unobserved determinants of political behavior are uncorrelated with age at move, as in the one-time movers design, the difference between b_{13-19} and b_{0-12} captures the differential impact of adolescence relative to earlier childhood years.

The regression includes μ_{e_1}, μ_{e_2}, and μ_{e_3}, fixed effects for the number of years of exposure to the first, second, and third destinations before age 19, respectively. To allow exposure effects in the second and third destinations to vary flexibly with years spent there, we also interact

[36] Appendix Table A13 reports summary statistics for the sample of multiple movers.

[37] In the multiple-movers design, the origin county is defined as the first county in which we observe an individual, which typically corresponds to the county of birth. Thus, when constructing exposure differences for the first, second, and third destinations, the origin county is held fixed.

the exposure-length fixed effects for these destinations with the corresponding destination-minus-origin differences in permanent residents' outcomes, $\Delta\overline{y}_{dos_2}$ and $\Delta\overline{y}_{dos_3}$.

Beyond these regressors, we use a more parsimonious set of fixed effects than in equation (3), given the smaller size of the multiple-movers sample. Specifically, we use origin-by-cohort fixed effects α_{os} instead of origin-by-cohort-by-age at move fixed effects, we omit origin-by-destination fixed effects, and we omit interactions between cohort fixed effects and destination-minus-origin differences in permanent residents' outcomes.

Finally, our preferred multiple-movers specification includes only individuals who spent two to six years in the first destination. Exposure spells shorter than two years may reflect measurement error in the address history (primarily in the partisanship sample), while six years corresponds to the length of the 13–19 age range. We also exclude individuals who lived in the first destination during both age ranges, 0–12 and 13–19, as well as those who appear to have returned to a previous county, since these cases may also reflect measurement error.

5.2 Results using multiple movers

Table 3 shows the b_{13-19} and b_{0-12} coefficients from equation (6) based on the sample of two- and three-time movers who spent two to six years in the first destination, either during ages 0 to 12 or 13 to 19.

We find substantially larger effects of location exposure during adolescence than earlier in childhood. Specifically, there is much stronger congruence between movers' and permanent residents' outcomes when movers live in their first destination during ages 13 to 19 rather than during ages 0 to 12. For registration in one's first eligible election, b_{13-19} is more than four times as large as b_{0-12} (0.438 versus 0.098); for turnout, it is nearly three times as large (0.395 versus 0.142). We observe a similar pattern for party affiliation outcomes: b_{13-19} exceeds b_{0-12} for both Republican Party affiliation (0.397 versus 0.262) and Democratic Party affiliation (0.418 versus 0.277).

While we restrict the analysis to movers who spend two to six years in the first destination either during ages 0–12 or during ages 13–19, the average length of exposure could still differ slightly between these two groups. To rule out this concern, Appendix Figure A7 shows separate estimates for exposure spells of 2, 3, 4, 5, and 6 years. For every exposure-spell length, movers exposed during ages 13–19 behave more similarly to permanent residents than movers exposed during ages 0–12.

Table 3: Multiple Movers, Effect of First Destination by Age of Exposure

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure effect, 0-12	0.098***	0.087***	0.142***	0.139***	0.262***	0.271***	0.277***	0.277***
	(0.004)	(0.005)	(0.005)	(0.006)	(0.012)	(0.014)	(0.012)	(0.013)
Exposure effect, 13-19	0.438***	0.416***	0.395***	0.379***	0.397***	0.388***	0.418***	0.407***
	(0.008)	(0.009)	(0.009)	(0.010)	(0.022)	(0.024)	(0.021)	(0.023)
Difference $[13, 19]$ vs. $[0, 12]$	0.340	0.328	0.253	0.240	0.134	0.118	0.141	0.130
p-value $[13, 19] = [0, 12]$	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Destination 2 & 3 FE		✓		✓		✓		✓
Number of observations	1,213,000	1,213,000	1,213,000	1,213,000	188,176	188,176	188,176	188,176
Dependent variable mean	0.566	0.566	0.322	0.322	0.205	0.205	0.225	0.225
R^2	0.095	0.146	0.112	0.159	0.136	0.244	0.115	0.195

Notes: The table reports exposure estimates from equation (6) based on the sample of two- and three-time movers who spent two to six years in the first destination, either during ages 0 to 12 or 13 to 19. Individuals who lived in the first destination during both age ranges are excluded. The estimates measure the extent to which young movers' political behavior aligns with the difference in permanent residents' outcomes between the origin and the first destination. Odd-numbered columns report estimates from our main multiple-movers specification, equation (6). In even-numbered columns, the underlying regressions include birth-cohort-by-second-destination-county fixed effects and birth-cohort-by-third-destination-county fixed effects. Each column also reports the difference between the effects of exposure to the first destination during ages 13–19 and ages 0–12, along with the corresponding p-value. This difference represents the effect of spending two to six years, in the first destination during ages 13–19 rather than during ages 0–12. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample.

Table 3 further reports the differences between adolescence- and childhood-specific exposure effects. Assuming that unobserved determinants of political behavior are uncorrelated with age of exposure to the first destination, these differences identify the differential effect of living in the first destination during adolescence rather than earlier in childhood. For registration, that differential impact is 0.340: spending two to six years between ages 13 and 19, rather than between ages 0 and 12, in a first destination where the permanent residents' registration rate is 10 percentage points higher than in the origin increases one's likelihood of registering by 3.4 percentage points. This difference is similar for turnout (0.253) and present but smaller for affiliation with the Republican Party (0.134) and with the Democratic Party (0.141). All of these differential effects are significant at the 1% level, indicating that location exposure during the teenage years influences future political behavior significantly more than exposure earlier in childhood.

Our multiple-movers estimates are robust to allowing for cohort-specific selection into the second and third destinations. For example, if younger cohorts systematically move to different second or third destinations than older cohorts, and those destinations differ in their political environments, estimates of first-destination exposure could partly reflect these later cohort-specific moves. To address this possibility, the estimates reported in columns 2, 4, 6, and 8 come from regressions that also include birth-cohort-by-second-destination-county fixed effects and birth-cohort-by-third-destination-county fixed effects. Reassuringly, all estimates remain nearly unchanged and significant at the 1% level.

We now show that the multiple-movers estimates can account for most of the non-linearity in the one-time-movers estimates, pointing to the especially strong effect of location exposure during adolescence rather than to diminishing marginal returns to exposure. We begin with the parametric one-time movers estimates in Table 2: for each outcome, we take the difference between estimated yearly exposure effects during adolescence and earlier childhood, and multiply this difference by four, the average number of years that multiple movers spend in the first destination. This calculation yields 0.279 for registration ($= 4 \times (0.0817 - 0.0120)$), 0.197 for turnout ($= 4 \times (0.0643 - 0.0151)$), 0.172 for Republican Party affiliation ($= 4 \times (0.0543 - 0.0112)$), and 0.153 for Democratic Party affiliation ($= 4 \times (0.0483 - 0.0100)$). These values are very similar to the differential effects of adolescence in the multiple-movers design shown in Table 3. The difference between the estimates in Table 2 for ages 13–19 and 0–12 can therefore be attributed to age-specific effects of location exposure. Because the estimates for the teenage years are at least four times as large, depending on the outcome, we infer that yearly exposure effects during adolescence are at least four times larger than those during earlier childhood.

6 Mechanisms

In this final section, we investigate the mechanisms responsible for the effects of childhood environments on adult participation and partisanship. We first ask whether childhood locations influence future political behavior directly or through socioeconomic outcomes such as income and education. We then compare the magnitudes of exposure effects for cross-state, cross-county, and cross-ZIP code moves to assess the role of explanatory factors varying at each of these levels.

6.1 Income and education

We first examine whether childhood environments influence political behavior through socioeconomic outcomes. This mechanism is plausible because income and education are strongly correlated with registration, turnout, and partisanship (e.g., Wolfinger and Rosenstone, 1980; Verba et al., 1995; Milligan et al., 2004; Gelman et al., 2008), and they are themselves strongly affected by childhood environments (Chetty and Hendren, 2018a; Chetty et al., 2026). In addition, a large body of evidence shows that *current* economic conditions are one of the key drivers of voting behavior (e.g., Fair, 1978; Lewis-Beck and Stegmaier, 2000; Autor et al., 2020).

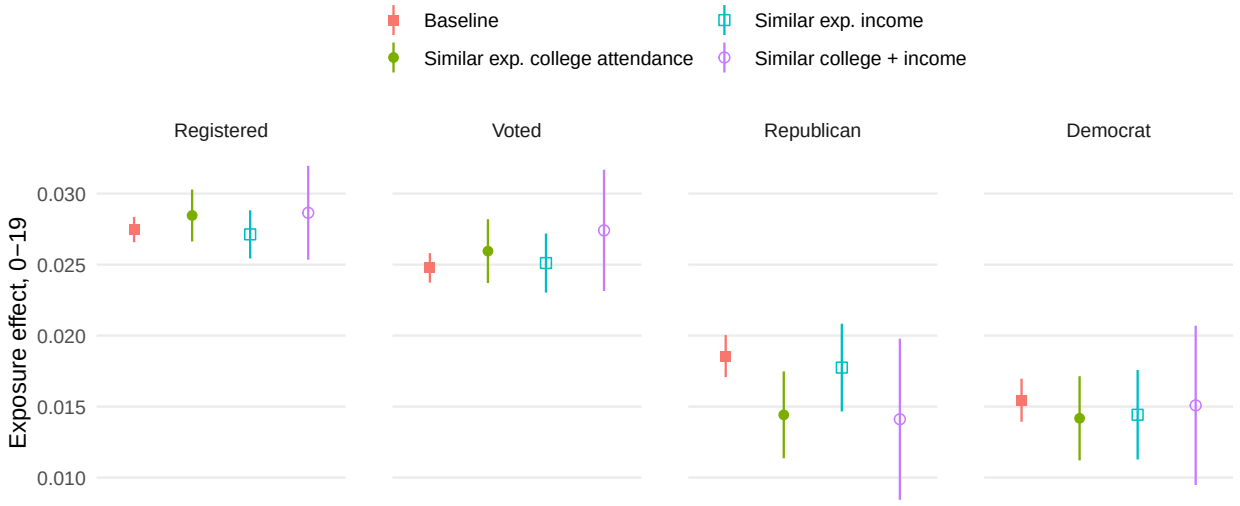
To test for this mediating relationship, we use Chetty et al. (2026)'s county-level data on expected adult income and college attendance for children with parents at the median of the national parent-income distribution. Specifically, we re-estimate our main parametric regression restricting the sample to movers whose origin and destination counties are in the same nationwide quartile of expected adult income, expected college attendance, or both. If education and income are important mediating factors, these sample restrictions should reduce the effects of childhood environment on political behavior.

Figure 7 shows that estimates from the restricted subsamples largely mirror those from the full sample. We infer that place effects on future partisanship and political participation are direct rather than mediated by effects on socioeconomic outcomes.

6.2 State, county, and ZIP code factors

Our effects may reflect the influence of the electoral context: rules governing how people register and vote, electoral competitiveness, the intensity of political campaigns, and which parties hold political power. For example, in battleground states, repeated exposure to competitive presidential elections and intense campaign activity may make future voters more likely to engage with politics (Hersh, 2015). Similarly, the partisan identity of state and local politicians and the policies they implement may shape voters' assessments of parties and, in turn, their party affiliation (Fiorina, 1981).

Figure 7: Exposure Effects for Moves Between Counties with Similar Expected Income and College Attendance



Notes: The figure plots parametric estimates of exposure effects, along with 95% confidence intervals, restricting the sample to movers whose origin and destination counties are in the same nationwide quartile of expected adult income, expected college attendance, or both. Full-sample estimates are also reported for comparison. County-level measures of expected college attendance and expected income are from Chetty et al. (2026) and are measured for individuals with parents at the median of the national parent income distribution. Registration and turnout results use data from the Census sample (N = 7,303,000, FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample (N = 2,175,432).

These factors vary primarily at the state level. Thus, to assess their importance, we compare exposure effects for people who move across states versus people who move across counties within the same state. Voters in the latter group are always exposed to the same state-level factors, regardless of the proportion of childhood spent in the origin and destination counties. Exposure effects for these voters therefore reflect only county-specific factors such as neighborhood demographics, the type and quality of schools, the local labor market, and the local political context. Naturally, differences in estimates between these two types of moves may also reflect differences between the types of families relocating across versus within states, and should thus be interpreted with caution.

As shown in Table 4, exposure effects on all outcomes are larger in the cross-state sample than in the within-state sample. The within-state effect on registration is about 74% the magnitude of the cross-state effect, while the corresponding figure for turnout is just 58%. For Republican and Democratic Party affiliation, the within-state effects are about 72% and 63% the size of the cross-state effects, respectively.

Table 4: Parametric Estimates of Exposure Effects for Cross-County Moves Within and Across States

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure effect, 0–19	0.0289***	0.0213***	0.0280***	0.0163***	0.0210***	0.0151***	0.0202***	0.0128***
	(0.0007)	(0.0009)	(0.0009)	(0.0009)	(0.0013)	(0.0012)	(0.0014)	(0.0012)
Across states	✓		✓		✓		✓	
Within state		✓		✓		✓		✓
Mean move distance (mi.)	743.2	48.6	743.2	48.6	828.7	48.7	828.7	48.7
Number of observations	2,868,000	4,435,000	2,868,000	4,435,000	777,259	1,399,532	777,259	1,399,532
Dependent variable mean	0.568	0.561	0.329	0.322	0.193	0.205	0.240	0.232
R^2	0.343	0.205	0.352	0.216	0.407	0.261	0.341	0.197

Notes: The table reports parametric estimates of yearly childhood exposure effects based on equation (5); that is, the effects of spending an additional year of childhood in a place where the permanent residents are 1 percentage point more likely to register (columns 1–2), vote (columns 3–4), affiliate with the Republican Party (columns 5–6), or affiliate with the Democratic Party (columns 7–8). For each outcome, the first column reports estimates restricting the sample to individuals who move across states, while the second column restricts the sample to individuals who move across counties within the same state. All outcomes are measured in the first election for which an individual is eligible to vote. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample.

These results suggest that state-level factors such as the electoral context and statewide economic conditions account for some of the exposure effects on registration and party affiliation, and for a larger share of the effect on voter turnout. However, even when state-level factors are held constant, the specific county in which people grow up affects their political behavior.

To better identify the types of factors responsible for exposure effects, we zoom in further and compare cross-county moves with within-county moves across ZIP codes.[38] Within-county moves hold fixed all county-level economic and political factors, including local labor market conditions and political culture.

As shown in Appendix Table A14, exposure effects remain substantial and, for all outcomes but Republican affiliation, statistically significant for moves across ZIP codes within the same county.[39] This finding suggests that political behavior in early adulthood is substantively affected by highly localized factors such as the towns and neighborhoods where children and adolescents live, the schools they attend, and their daily interactions with peers and adults.

7 Conclusion

Much of the academic literature and public commentary on political behavior focuses on its recent or contemporary determinants. We show that differences in political participation and party affiliation across individuals are also shaped to a large extent by the places in which they grew up.

Our analysis relies on the first-ever combination of three comprehensive individual-level data sources: voter records, address histories, and Census data. The use of administrative records rather than self-reported information, together with the fact that our data cover all voting-eligible U.S. citizens, strengthens the reliability and external validity of our results. In addition, because we track parents' precise locations over time, we can measure the exact age at which children and adolescents move across counties and distinguish permanent residents, one-time movers, and individuals who move multiple times before adulthood.

Our main analysis compares individuals from the same cohort who move once between the

[38] For these analyses, moves are defined at the ZIP code level, rather than the county level. Therefore, one-time movers are now defined as individuals who moved exactly once across ZIP codes between ages 0 and 26, and they include some individuals who were considered permanent residents in the county-level analyses.

[39] To account for potential measurement error in exposure due to the small number of permanent residents in each ZIP-by-cohort cell, we use a split-sample instrumental variables approach. We randomly split permanent residents in each ZIP code into two groups (at the family level), then compute permanent resident outcome means in each ZIP-by-cohort cell, separately for both groups. We then regress the mean outcomes in the first group on the mean outcomes in the second, weighted by the total number of permanent residents in that ZIP-by-cohort cell across both groups. We use the fitted values from that regression as the exposure measure in our ZIP-code-level regressions.

same origin-destination county pair, but at different ages, and relates their adult political behavior to that of permanent residents in the destination county. This design provides the first estimates of the aggregate effect of geographic context at each age of childhood on the political behavior of young adults. Growing up in a county where same-cohort permanent residents are 10 percentage points more likely to be registered in their first eligible election increases one's own registration probability by 6.87 percentage points. The corresponding effects are 5.75 percentage points for turnout, 4.69 for Republican Party affiliation, and 4.09 for Democratic Party affiliation. These estimates imply that childhood environment accounts for 40 to 50% of observed cross-county variation in young voters' party affiliation, and for even larger shares of variation in registration and turnout.

Our complementary analyses using multiple movers show that the timing of exposure matters. Spending the same number of years in a given county has much larger effects on adult political behavior when those years occur during adolescence than earlier in childhood. This pattern contrasts with the effects of childhood locations on income or education, which previous work has shown to be linear in years of exposure.

Additional analyses suggest that effects on political behavior are unlikely to be mediated by changes in socioeconomic outcomes. Instead, childhood place effects operate through both broad political contexts and local social environments: state-level factors play a larger role for turnout, whereas more localized forces appear more important for party affiliation. The effects also persist well beyond the first election in which an individual is eligible to participate: turnout effects remain 80% as large by the third election, and party affiliation effects are even more persistent.

The large, direct, and durable influence of childhood places on politics has implications beyond individual political behavior. By transmitting geographic differences in turnout across generations, childhood environments may perpetuate inequalities in the attention that different places receive from elected officials. Childhood environments may also reinforce the geographic partisan segregation and the polarization of American politics. Partisan differences across places, our findings suggest, reflect not only where adults choose to live and how they adjust to local political context, but also where children happen to grow up.

References

Autor, David, David Dorn, Gordon Hanson, and Kaveh Majlesi, "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," *American Economic Review*, 2020, *110* (10), 3139–3183.

Azzimonti, Marina, "Partisan conflict and private investment," *Journal of Monetary Economics*, 2018, *93*, 114–131.

Bartels, Larry M., "Partisanship and Voting Behavior, 1952-1996," *American Journal of Political Science*, 2000, *44* (1), 35–50.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "Who Becomes an Inventor in America? The Importance of Exposure to Innovation," *The Quarterly Journal of Economics*, 2019, *134* (2), 647–713.

Berry, Christopher and Jessica Trounstine, "Does Where You Live Shape Your Politics? Measuring the Effect of Neighborhood on Political Identity and Behavior," 2023. Working Paper, University of Chicago.

Billings, Stephen B., Eric Chyn, and Kareem Haggag, "The Long-Run Effects of School Racial Diversity on Political Identity," *American Economic Review: Insights*, 2021, *3* (3), 267–284.

Bolotnyy, Valentin, Mayya Komisarchik, and Brian Libgober, "Backlash Against Repression: Evidence from Refugees Fleeing the Former Soviet Bloc," 2026. Working Paper.

Brown, Jacob R., "Partisan Conversion Through Neighborhood Influence: How Voters Adopt the Partisanship of their Neighbors," *The Journal of Politics*, 2025, *87* (4), 1482–1498.

_ , Ryan D. Enos, James Feigenbaum, and Soumyajit Mazumder, "Childhood cross-ethnic exposure predicts political behavior seven decades later: Evidence from linked administrative data," *Science Advances*, 2021, *7* (24), eabe8432.

Campbell, David E., *Why We Vote: How Schools and Communities Shape Our Civic Life*, Princeton, NJ: Princeton University Press, 2006.

Cantoni, Enrico and Vincent Pons, "Does Context Outweigh Individual Characteristics in Driving Voting Behavior? Evidence from Relocations within the United States," *American Economic Review*, 2022, *112* (4), 1226–1272.

_ , _ , and Jerome Schafer, "Voting Rules, Turnout, and Economic Policies," *Annual Review of Economics*, 2025, *17*.

Chetty, Raj and Nathaniel Hendren, "The Impacts of Neighborhoods on Intergenerational Mobility I: Childhood Exposure Effects," *The Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162.

_ **and** _ , "The Impacts of Neighborhoods on Intergenerational Mobility II: County-Level Estimates," *The Quarterly Journal of Economics*, 2018, *133* (3), 1163–1228.

_ **, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter**, "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility," *American Economic Review*, 2026, *116* (1), 1–51.

Chyn, Eric, "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children," *American Economic Review*, 2018, *108* (10), 3028–3056.

_ **and Kareem Haggag**, "Moved to Vote: The Long-Run Effects of Neighborhoods on Political Participation," *The Review of Economics and Statistics*, 2023, *105* (6), 1596–1605.

_ **and Lawrence F. Katz**, "Neighborhoods Matter: Assessing the Evidence for Place Effects," *Journal of Economic Perspectives*, 2021, *35* (4), 197–222.

Cohodes, Sarah and James J. Feigenbaum, "Why Does Education Increase Voting? Evidence from Boston's Charter Schools," Working Paper 29308, National Bureau of Economic Research 2025.

Damm, Anna Piil and Christian Dustmann, "Does Growing Up in a High Crime Neighborhood Affect Youth Criminal Behavior?," *American Economic Review*, 2014, *104* (6), 1806–1832.

Derenoncourt, Ellora, "Can You Move to Opportunity? Evidence from the Great Migration," *American Economic Review*, 2022, *112* (2), 369–408.

Deutscher, Nathan, "Place, Peers, and the Teenage Years: Long-Run Neighborhood Effects in Australia," *American Economic Journal: Applied Economics*, 2020, *12* (2), 220–249.

Diamond, Rebecca, Tim McQuade, and Franklin Qian, "The Effects of Rent Control Expansion on Tenants, Landlords, and Inequality: Evidence from San Francisco," *American Economic Review*, 2019, *109* (9), 3365–3394.

Dinas, Elias, "Does Choice Bring Loyalty? Electoral Participation and the Development of Party Identification," *American Journal of Political Science*, 2014, *58* (2), 449–465.

Elder, Elizabeth Mitchell, Ryan D. Enos, and Tali Mendelberg, "The Long-Term Effects of Neighborhood Disadvantage on Voting Behavior: The "Moving to Opportunity" Experiment," *American Political Science Review*, 2024, *118* (2), 988–1004.

Fair, Ray C., "The Effect of Economic Events on Votes for President," *The Review of Economics and Statistics*, 1978, pp. 159–173.

Fiorina, Morris P., *Retrospective Voting in American National Elections*, New Haven, CT: Yale University Press, 1981.

Fowler, James H., Laura A. Baker, and Christopher T. Dawes, "Genetic Variation in Political Participation," *American Political Science Review*, 2008, *102* (2), 233–248.

Funke, Manuel, Moritz Schularick, and Christoph Trebesch, "Populist Leaders and the Economy," *American Economic Review*, 2023, *113* (12), 3249–3288.

Gelman, Andrew, Boris Shor, Joseph Bafumi, and David Park, "Rich State, Poor State, Red State, Blue State: What's the Matter with Connecticut?," *Quarterly Journal of Political Science*, 2008, *2* (4), 345–367.

Ghitza, Yair, Andrew Gelman, and Jonathan Auerbach, "The Great Society, Reagan's Revolution, and Generations of Presidential Voting," *American Journal of Political Science*, 2023, *67* (3), 520–537.

Goel, Sharad, Marc Meredith, Michael Morse, David Rothschild, and Houshmand Shirani-Mehr, "One Person, One Vote: Estimating the Prevalence of Double Voting in U.S. Presidential Elections," *American Political Science Review*, 2020, *114* (2), 456–469.

Goldman, Seth K. and Daniel J. Hopkins, "Past Place, Present Prejudice: The Impact of Adolescent Racial Context on White Racial Attitudes," *The Journal of Politics*, 2020, *82* (2), 529–542.

Graham, Matthew R., Mark J. Kutzbach, and Danielle H. Sandler, "Developing a Residence Candidate File for Use With Employer-Employee Matched Data," Working Paper 17-40, Center for Economic Studies, U.S. Census Bureau 2017.

Green, Donald, Bradley Palmquist, and Eric Schickler, *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*, Yale University Press, 2002.

Guo, Fang, "Partisan Registration: A Truthful Statement or A Strategic Choice?," Working Paper 2023.

Healy, Andrew and Neil Malhotra, "Childhood Socialization and Political Attitudes: Evidence from a Natural Experiment," *The Journal of Politics*, 2013, *75* (4), 1023–1037.

Hersh, Eitan, *Hacking the Electorate: How Campaigns Perceive Voters*, Cambridge, UK: Cambridge University Press, 2015.

_ **and Yair Ghitza**, "Mixed partisan households and electoral participation in the United States," *PLOS ONE*, 2018, *13* (10), e0203997.

Huckfeldt, Robert and John Sprague, "Networks in Context: The Social Flow of Political Information," *American Political Science Review*, 1987, *81* (4), 1197–1216.

Janmaat, Jan Germen and Bryony Hoskins, "The Changing Impact of Family Background on Political Engagement During Adolescence and Early Adulthood," *Social Forces*, 2021, *101* (1), 227–251.

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

Kaplan, Ethan, Jörg L. Spenkuch, and Cody Tuttle, "A Different World: Enduring Effects of School Desegregation on Ideology and Attitudes," Technical Report, National Bureau of Economic Research 2025.

Kling, Jeffrey R., Jeffrey B. Liebman, and Lawrence F. Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, 2007, *75* (1), 83–119.

Knight, David Jonathan and Baobao Zhang, "Residential mobility and persistently depressed voting among disadvantaged adults in a large housing experiment," *Proceedings of the National Academy of Sciences*, 2024, *121* (20), e2306287121.

Laliberté, Jean-William, "Long-Term Contextual Effects in Education: Schools and Neighborhoods," *American Economic Journal: Economic Policy*, 2021, *13* (2), 336–377.

Lewis-Beck, Michael S. and Mary Stegmaier, "Economic Determinants of Electoral Outcomes," *Annual Review of Political Science*, 2000, *3* (1), 183–219.

Marx, Benjamin, Vincent Pons, and Vincent Rollet, "Electoral Turnovers," *Review of Economic Studies*, 2025, *92* (5), 3306–3339.

McDonald, Michael P., "Voter Turnout Data, 2020," 2022.

McNeil, Andrew, Davide Luca, and Neil Lee, "The long shadow of local decline: Birthplace economic adversity and long-term individual outcomes in the UK," *Journal of Urban Economics*, 2023, *136*, 103571.

Milligan, Kevin, Enrico Moretti, and Philip Oreopoulos, "Does education improve citizenship? Evidence from the United States and the United Kingdom," *Journal of Public Economics*, 2004, *88* (9-10), 1667–1695.

Mullainathan, Sendhil and Ebonya Washington, "Sticking with Your Vote: Cognitive Dissonance and Political Attitudes," *American Economic Journal: Applied Economics*, 2009, *1* (1), 86–111.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson, "The Gift of Moving: Intergenerational Consequences of a Mobility Shock," *The Review of Economic Studies*, 2022, *89* (3), 1557–1592.

Neundorf, Anja and Kaat Smets, "Political Socialization and the Making of Citizens," in "Oxford Handbook Topics in Politics," Oxford University Press, 2017.

Nyhan, Brendan, Christopher Skovron, and Rocío Titiunik, "Differential Registration Bias in Voter File Data: A Sensitivity Analysis Approach," *American Journal of Political Science*, 2017, *61* (3), 744–760.

Pacheco, Julianna Sandell, "Political Socialization in Context: The Effect of Political Competition on Youth Voter Turnout," *Political Behavior*, 2008, *30* (4), 415–436.

Pennington, Kate, "Rents, Movers, and Stayers: The Spillover Effects of Construction in San Francisco," Working Paper 2025.

Perez-Truglia, Ricardo, "Political Conformity: Event-Study Evidence from the United States," *Review of Economics and Statistics*, 2018, *100* (1), 14–28.

Pons, Vincent and Guillaume Liegey, "Increasing the Electoral Participation of Immigrants: Experimental Evidence from France," *The Economic Journal*, 2019, *129* (617), 481–508.

Sears, David O. and Nicholas A. Valentino, "Politics Matters: Political Events as Catalysts for Preadult Socialization," *The American Political Science Review*, 1997, *91* (1), 45–65.

U.S. Census Bureau, "Voting and Registration in the Election of November 2020," Report P20-585 April 2021.

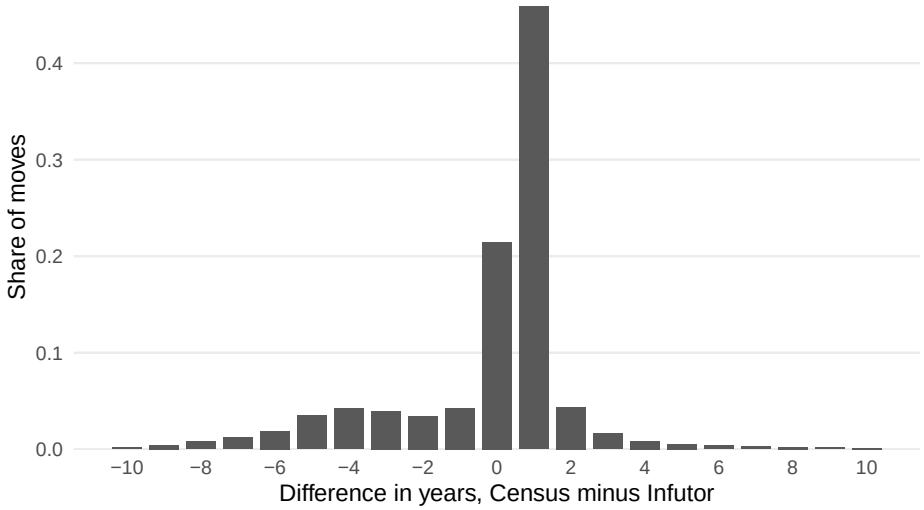
Verba, Sidney, Kay Lehman Schlozman, and Henry E. Brady, *Voice and Equality: Civic Voluntarism in American Politics*, Cambridge, MA: Harvard University Press, 1995.

Wagner, Deborah and Mary Lane, "The Person Identification Validation System (PVS): Applying the Center for Administrative Records Research and Applications' (CARRA) Record Linkage Software," Technical Report, Center for Economic Studies, US Census Bureau 2014.

Wolfinger, Raymond E. and Steven J. Rosenstone, *Who Votes?*, New Haven, CT: Yale University Press, 1980.

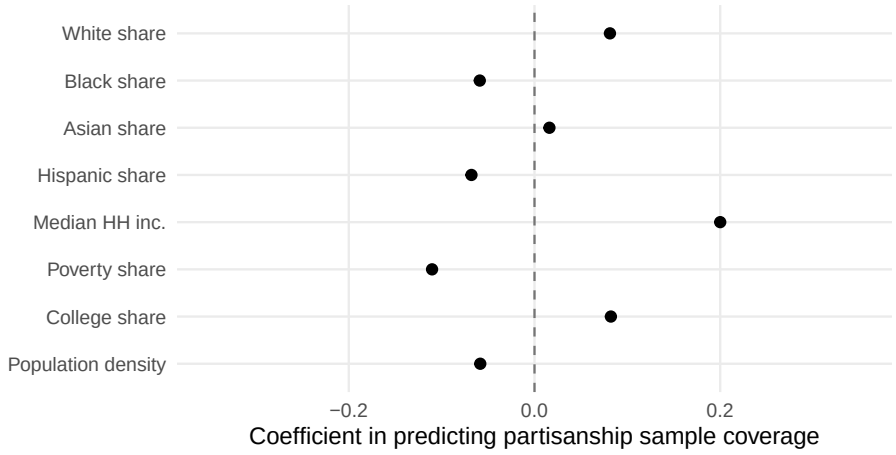
A Additional tables and figures

Figure A1: Difference in Move Timing in Census vs. Infutor Data



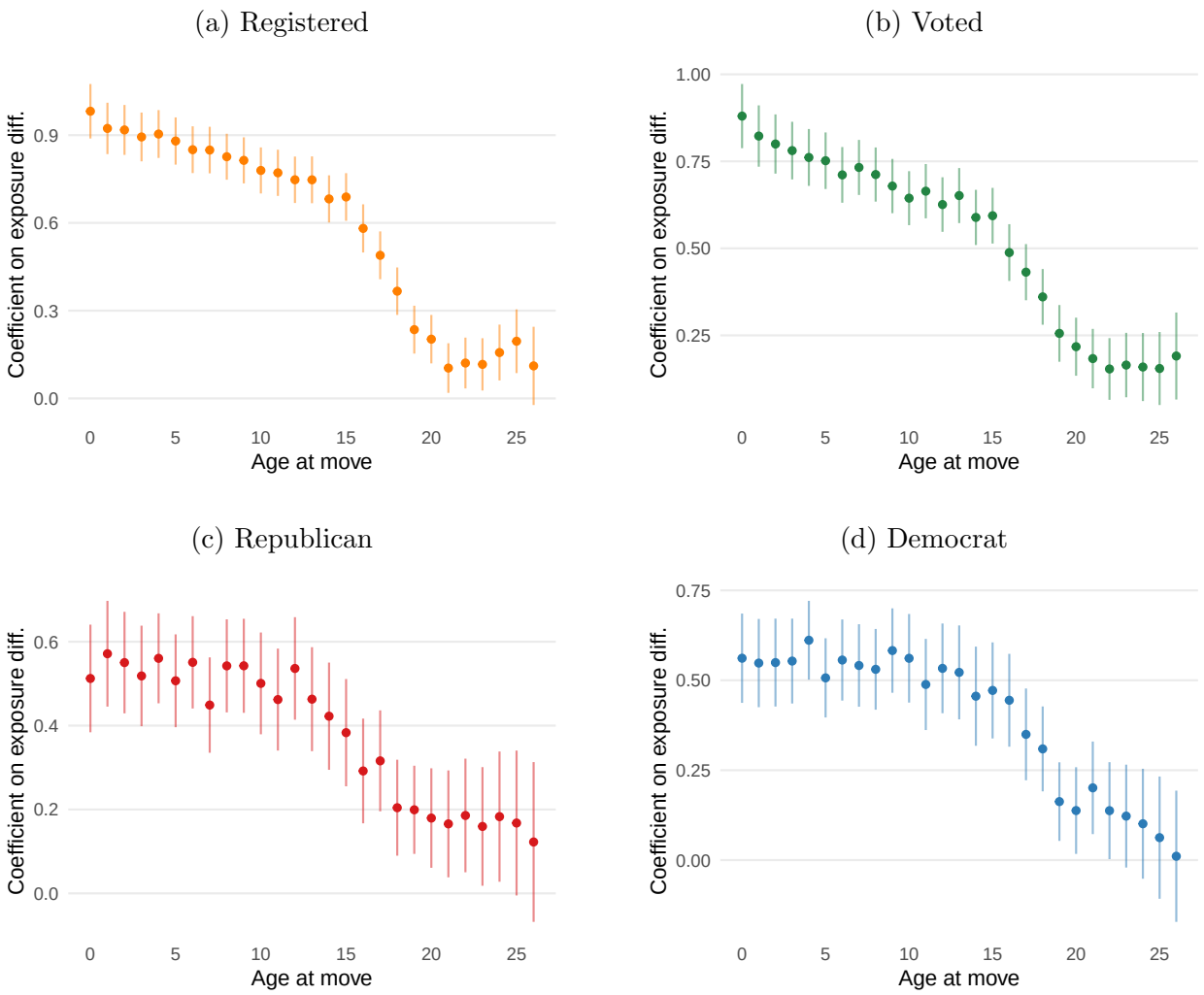
Notes: The figure plots the distribution of the difference, in calendar years, between when a cross-county move is recorded in the Census data versus when the same move is recorded in the Infutor data. Positive differences mean that a move is recorded earlier in Infutor than in Census data; negative differences mean the opposite. The figure shows only moves (N = 1,236,000) in which the identity of the mover, the origin county, and the destination county are the same in both sources of address data. Data are from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874).

Figure A2: Relationship Between Block Group Characteristics and Partisanship Sample Coverage



Notes: The figure plots estimated coefficients from census block group-level bivariate regressions. Each row corresponds to a separate independent variable based on 2015–2019 5-year ACS data. In all rows, the dependent variable is the ratio between block group-level headcounts of 18-to-29-year-olds in the partisanship sample and ACS population counts in the same age range. Individuals appearing in multiple annual snapshots are counted once in each, at their block group of residence in that year. All variables are standardized using ACS census block group population weights. A coefficient of zero means that the block-group covariate does not predict the ratio between counts in the partisanship sample and the ACS counts, indicating high correspondence across the two samples. Positive coefficients mean that the partisanship sample overcounts people in block groups with higher values of the corresponding ACS characteristics, while negative coefficients mean the opposite. Horizontal lines show 95% confidence intervals. N = 16,160,156.

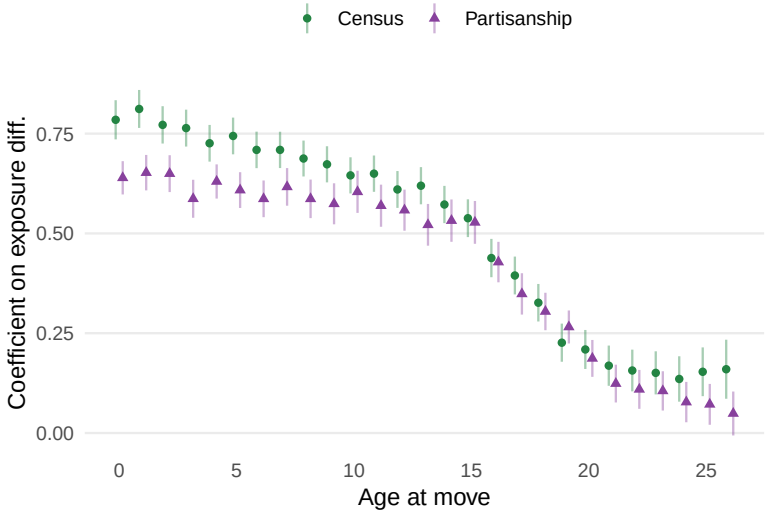
Figure A3: Sibling Comparisons: Nonparametric Estimates



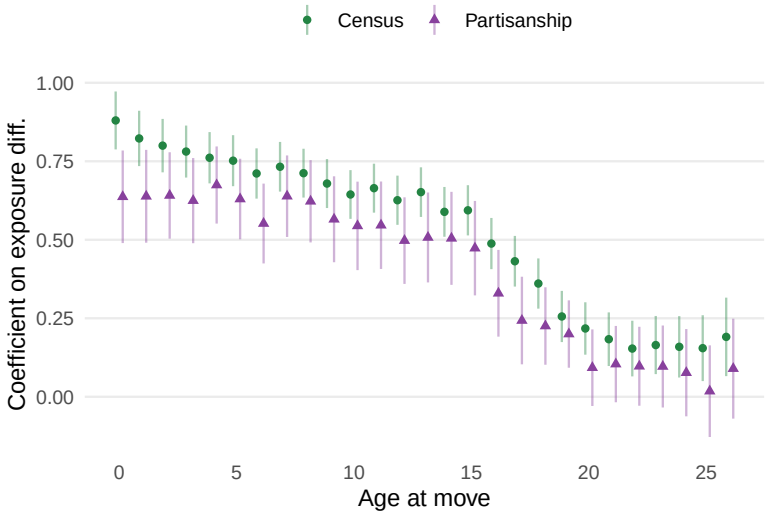
Notes: The figure plots age-specific exposure effects estimated using siblings of different ages with the same parent. See Section 4.2.1 for details. All outcomes are measured for the first election for which an individual is eligible to vote. Vertical lines show 95% confidence intervals. Registration and turnout results use data from the Census sample (N = 4,092,000, FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results are from the partisanship sample (N = 894,201).

Figure A4: Nonparametric Estimates of Childhood Exposure Effects on Turnout – Census vs. Partisanship Samples

(a) Origin-by-destination fixed effects

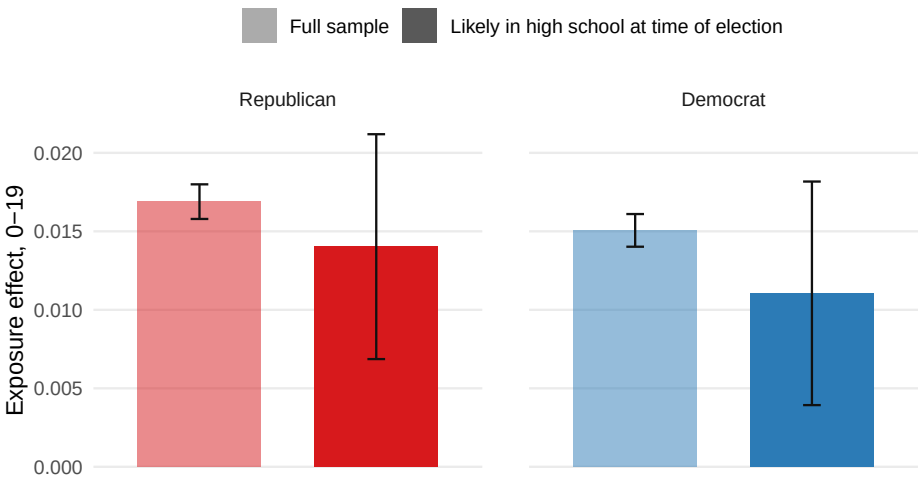


(b) Family fixed effects



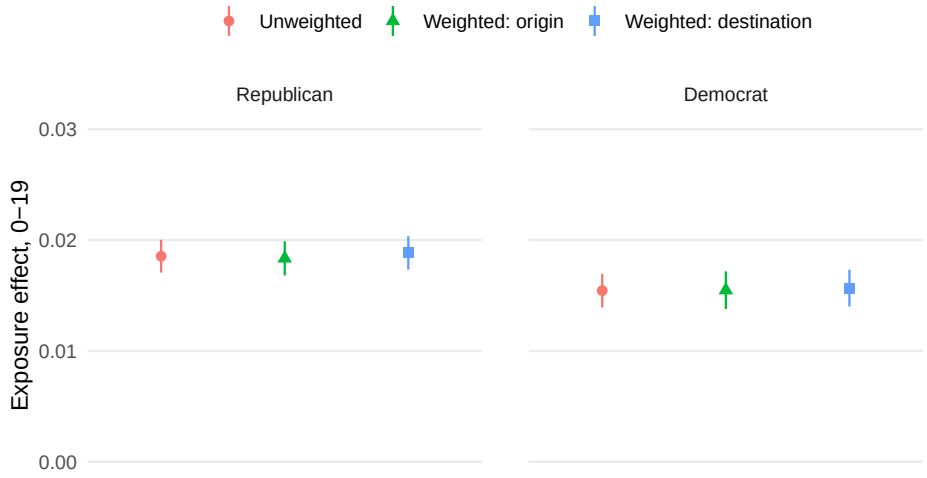
Notes: The figure plots age-specific exposure effects on turnout in an individual's first eligible election, separately for the Census sample (green circles) and the partisanship sample (purple triangles). Panel a shows estimates from the baseline regression with origin-by-destination fixed effects, while Panel b shows estimates using family fixed effects. Vertical lines show 95% confidence intervals. In Panel a, $N = 7{,}303{,}000$ for the Census sample and $N = 4{,}771{,}738$ for the partisanship sample. In Panel b, $N = 4{,}092{,}000$ for the Census sample and $N = 1{,}948{,}316$ for the partisanship sample. Data are from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and the partisanship sample.

Figure A5: Party Affiliation Exposure Effects – Full Sample vs. Newly Eligible Voters Likely to Be in High School



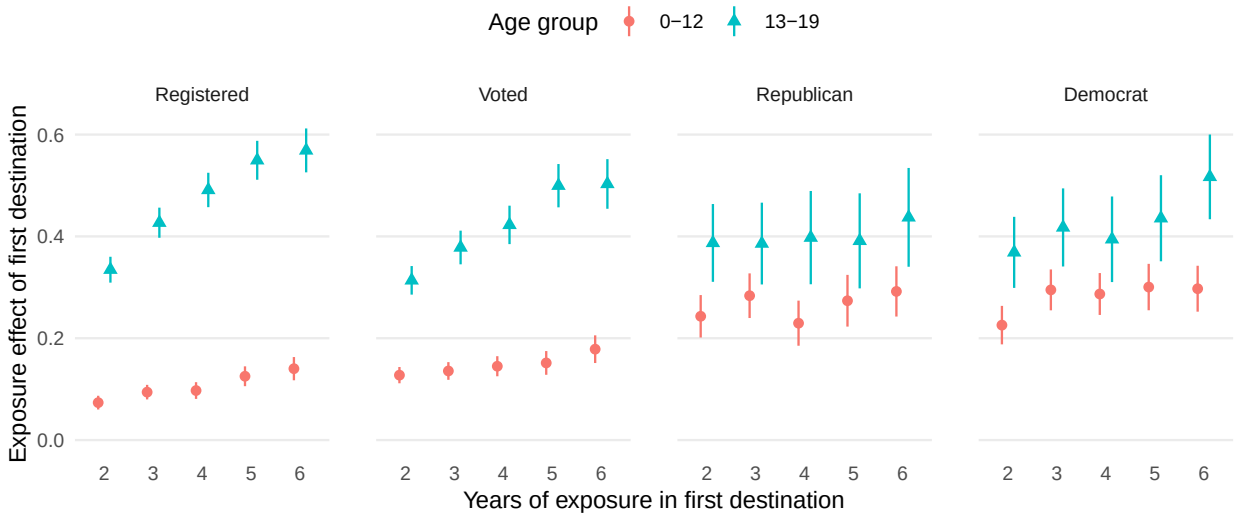
Notes: The figure shows parametric estimates of yearly exposure effects separately for the full partisanship sample (in lighter colors) and for a sample of individuals who were likely in high school at the time of their first eligible election (in darker colors). Specifically, the latter sample consists of voters who turned 18 in the 100 days before their first eligible election and who were born after the state-specific school entry age cutoff. Vertical lines show 95% confidence intervals. For the full sample, $N = 1{,}734{,}048$. For the restricted sample, $N = 36{,}977$.

Figure A6: Parametric Estimates of Childhood Exposure Effects – Weighting Observations to Match County Populations



Notes: The figure shows parametric estimates of yearly exposure effects after weighting observations to match county-level counts of 18-to-29-year-old U.S. citizens according to the 2015–2019 5-year ACS (source table: Citizen, Voting-Age Population by Age). The orange circles show the baseline (unweighted) estimates, the green triangles show estimates weighted by movers' counties of origin, and the blue squares show estimates weighted by movers' destination counties. Vertical lines show 95% confidence intervals. See Section 4.4 for details. $N = 2{,}176{,}736$.

Figure A7: Multiple Movers, Effect of First Destination by Age of Exposure, Separately by Years of Exposure



Notes: The figure plots exposure estimates from equation (6) based on the sample of two- and three-time movers who spent two to six years in the first destination, either during ages 0 to 12 or 13 to 19. Individuals who lived in the first destination during both age ranges are excluded. The estimates measure the extent to which young movers' political behavior aligns with the difference in permanent residents' outcomes between the origin and the first destination. The figure shows estimates separately by years of exposure in the first destination. Vertical lines show 95% confidence intervals. Registration and turnout results use data from the Census sample (N = 1,213,000, FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample (N = 188,176).

Table A1: Parametric Estimates of Exposure Effects – Exact Birth Date Partisanship Sample

	Republican		Democrat	
	(1)	(2)	(3)	(4)
Exposure effect, 0–12	0.0111***	0.0102	0.0112***	0.0094
	(0.0016)	(0.0063)	(0.0017)	(0.0066)
Exposure effect, 13–19	0.0511***	0.0428***	0.0485***	0.0539***
	(0.0042)	(0.0108)	(0.0047)	(0.0121)
Exposure effect, 20–26	0.0168***	0.0068	0.0219***	0.0350***
	(0.0045)	(0.0113)	(0.0049)	(0.0120)
Difference $[13, 19]$ vs. $[0, 12]$	0.040	0.033	0.037	0.045
p-value $[13, 19] = [0, 12]$	0.0000	0.0156	0.0000	0.0030
Family fixed effects		✓		✓
Number of observations	1,731,702	697,326	1,731,702	697,326
Dependent variable mean	0.193	0.211	0.245	0.235
R^2	0.273	0.751	0.214	0.708

Notes: The table reports yearly exposure effects on the partisanship sample, excluding voters whose recorded birth dates indicate that they were born on January 1. It is likely that only the year of birth, rather than the exact date of birth, was recorded for these voters. This restriction applies to both movers and permanent residents. The regression specification is equation (4), and the estimates can be compared to those in Table 2. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table A2: TargetSmart to Census Link Rates

Sample	Num. in TargetSmart	Num. linked to Census	Share linked
All rows	2,575,697,805	2,225,000,000	0.864
Registered	2,023,356,072	1,941,000,000	0.959
1992–2002 cohorts	207,628,037	197,000,000	0.949
Registered, 1992–2002 cohorts	200,347,706	191,800,000	0.957

Notes: The table reports link rates from TargetSmart to the Census data. Rates are computed at the row level, so each observation represents an individual in a TargetSmart annual snapshot. We show statistics for all TargetSmart rows, rows corresponding to registered individuals, rows corresponding to individuals born between 1992 and 2002 (according to the TargetSmart data), and rows corresponding to registered individuals in those cohorts. The first column reports the number of rows in the TargetSmart data, the second reports the rounded number of rows successfully linked to the Census data (assigned a PIK), and the third reports the ratio between linked rows and all TargetSmart rows. Linking was conducted by the U.S. Census Bureau. Data are from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874).

Table A3: Parametric Estimates of Exposure Effects – Move Timing Correction

	Registered				Voted			
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure effect, 0–12	0.0120***	0.0167***	0.0117***	0.0167***	0.0151***	0.0177***	0.0151***	0.0191***
	(0.0008)	(0.0020)	(0.0009)	(0.0020)	(0.0010)	(0.0021)	(0.0010)	(0.0021)
Exposure effect, 13–19	0.0817***	0.0837***	0.0618***	0.0616***	0.0643***	0.0635***	0.0498***	0.0470***
	(0.0021)	(0.0039)	(0.0021)	(0.0038)	(0.0025)	(0.0044)	(0.0024)	(0.0042)
Exposure effect, 20–26	0.0022	0.0051	0.0155***	0.0182***	0.0101**	0.0090	0.0162***	0.0181***
	(0.0037)	(0.0068)	(0.0031)	(0.0057)	(0.0042)	(0.0071)	(0.0035)	(0.0062)
Family fixed effects		✓		✓		✓		✓
Uncorrected			✓	✓			✓	✓
Number of observations	7,303,000	4,092,000	7,315,000	4,184,000	7,303,000	4,092,000	7,315,000	4,184,000
Dependent variable mean	0.564	0.561	0.565	0.564	0.325	0.318	0.326	0.322
R^2	0.228	0.642	0.229	0.640	0.242	0.661	0.243	0.661

Notes: The table reports parametric estimates of yearly exposure effects on registration (columns 1–4) and turnout (columns 5–8) in an individual's first eligible election. Columns 1–2 and 5–6 use the move timing correction that we employ in the main analysis, described in Section 2.2, which shifts moves one year into the past from when they are observed in MAF-ARF. Columns 3–4 and 7–8 do not use this correction. Columns 2, 4, 6, and 8 are from regressions that include family fixed effects. Estimates are based on equation (4). Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Data are from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874).

Table A4: Summary Statistics for Partisanship Sample – Including States That Do Not Record Party Affiliation

(a) Permanent residents

	Mean	Std. dev.	Num. of obs.
Birth year	1997	2.90	11,261,211
Male	0.507	0.500	10,557,457
White	0.700	0.458	10,893,783
Black	0.107	0.309	10,893,783
Hispanic	0.153	0.360	10,893,783
Other race	0.041	0.198	10,893,783
Number of children in family	1.57	0.75	11,261,211
Voted in first election	0.427	0.495	11,261,211
Republican in first election	0.186	0.389	6,394,724
Democrat in first election	0.257	0.437	6,394,724
State records party	0.568	0.495	11,261,211

(b) One-time movers

	Mean	Std. dev.	Num. of obs.
Birth year	1997	2.87	4,772,561
Male	0.510	0.500	4,528,495
White	0.808	0.394	4,623,374
Black	0.075	0.264	4,623,374
Hispanic	0.080	0.271	4,623,374
Other race	0.037	0.188	4,623,374
Number of children in family	1.52	0.73	4,772,561
Age at move	11.9	7.98	4,772,561
Move distance (mi.)	390.2	608.7	4,772,561
Voted in first election	0.446	0.497	4,772,561
Abs. exposure diff. in voted	0.064	0.061	4,771,738
Republican in first election	0.204	0.403	2,664,896
Democrat in first election	0.235	0.424	2,664,896
Abs. exposure diff. in Rep.	0.086	0.076	2,213,255
Abs. exposure diff. in Dem.	0.096	0.082	2,213,255
Orig. and dest. record party	0.464	0.499	4,772,561

Notes: The table reports summary statistics for the partisanship sample in all states, including those that do not record party affiliation. This table corresponds to Table 1, except Table 1 reports statistics for the partisanship sample limited to individuals living in the 30 states and the District of Columbia where party affiliation is recorded. Panels a and b refer to, respectively, permanent residents (individuals whose parents never relocate across counties when the individuals are aged 0 through 26) and one-time movers (individuals whose parents relocate across counties exactly once over the same age range).

Table A5: Median ACS Characteristics for Various Samples

(a) County

	ACS	TargetSmart	Linked to parents	Linked to Infutor
White share	0.748	0.743	0.751	0.751
Black share	0.083	0.087	0.083	0.083
Hispanic share	0.111	0.116	0.109	0.108
Asian share	0.034	0.037	0.037	0.036
College share	0.605	0.609	0.612	0.612
Median HH income	61,705	63,098	64,022	64,016
Poverty rate	0.138	0.137	0.13	0.129
Pop. density (per sq. mi.)	555	647	665	647
Num. of obs.	49,855,197	24,887,236	16,680,090	16,231,550

(b) Block group

	ACS	TargetSmart	Linked to parents	Linked to Infutor
White share	0.776	0.796	0.825	0.827
Black share	0.05	0.041	0.032	0.031
Hispanic share	0.089	0.081	0.07	0.069
Asian share	0.016	0.016	0.017	0.017
College share	0.602	0.616	0.641	0.642
Median HH income	59,773	64,740	71,094	71,250
Poverty rate	0.116	0.097	0.08	0.079
Pop. density (per sq. mi.)	3,246	2,805	2,445	2,421
Num. of obs.	49,855,197	24,881,086	16,675,663	16,227,215

Notes: The table reports weighted county- and block group-level median characteristics based on 2015–2019 ACS 5-year data. In the first column, we weight each county or block group by ACS counts of 18-to-29-year-old citizens. In the second column, we weight observations by counts of individuals aged 18 to 26 in the 2015–2019 TargetSmart snapshots. Individuals who appear in multiple snapshots are counted once for each snapshot based on where they were registered to vote in that year. In the third column, we weight by the number of 18-to 26-year old individuals in the 2015–2019 TargetSmart snapshots who successfully link to a parent. In the last column, we weight by the number of 18-to-26-year-old individuals in the 2015–2019 TargetSmart snapshots linked to a parent for whom we are able to retrieve address data from Infutor. This corresponds to a subset of our analysis sample of permanent residents and movers for 2015 through 2019.

Table A6: Parametric Estimates of Exposure Effects – Two and Three Age Groups

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure effect, 0–19	0.0275***		0.0248***		0.0185***		0.0154***	
	(0.0005)		(0.0005)		(0.0008)		(0.0008)	
Exposure effect, 0–12		0.0120***		0.0151***		0.0112***		0.0100***
		(0.0008)		(0.0010)		(0.0015)		(0.0015)
Exposure effect, 13–19		0.0817***		0.0643***		0.0543***		0.0483***
		(0.0021)		(0.0025)		(0.0038)		(0.0041)
Exposure effect, 20–26	0.0054	0.0022	0.0122***	0.0101**	0.0190***	0.0175***	0.0206***	0.0194***
	(0.0037)	(0.0037)	(0.0042)	(0.0042)	(0.0042)	(0.0042)	(0.0044)	(0.0044)
Number of observations	7,303,000	7,303,000	7,303,000	7,303,000	2,176,791	2,176,791	2,176,791	2,176,791
Dependent variable mean	0.564	0.564	0.325	0.325	0.201	0.201	0.235	0.235
R^2	0.228	0.228	0.242	0.242	0.277	0.277	0.219	0.219

Notes: The table reports parametric estimates of yearly exposure effects, constraining the exposure effect to be the same for ages 0 to 19. The regression specification for columns 1, 3, 5, and 7 is equation (5). For columns 2, 4, 6, and 8, the regression specification is equation (4), and the reported estimates are identical to those in Table 2. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample.

Table A7: Parametric Estimates of Exposure Effects – Sibling Comparisons

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Exposure effect, 0–12	0.0119***	0.0167***	0.0146***	0.0177***	0.0097***	0.0034	0.0076***	0.0022
	(0.0012)	(0.0020)	(0.0014)	(0.0021)	(0.0028)	(0.0058)	(0.0027)	(0.0058)
Exposure effect, 13–19	0.0848***	0.0837***	0.0690***	0.0635***	0.0474***	0.0461***	0.0554***	0.0583***
	(0.0029)	(0.0039)	(0.0034)	(0.0044)	(0.0064)	(0.0101)	(0.0067)	(0.0107)
Exposure effect, 20–26	-0.0019	0.0051	0.0043	0.0090	0.0176**	0.0045	0.0206***	0.0201*
	(0.0051)	(0.0068)	(0.0058)	(0.0071)	(0.0069)	(0.0109)	(0.0072)	(0.0111)
Difference $[13, 19]$ vs. $[0, 12]$	0.0729	0.0670	0.0543	0.0458	0.038	0.043	0.048	0.056
p-value $[13, 19] = [0, 12]$	0.000	0.000	0.000	0.000	0.0000	0.0005	0.0000	0.0000
Family fixed effects		✓		✓		✓		✓
Number of observations	4,092,000	4,092,000	4,092,000	4,092,000	894,201	894,201	894,201	894,201
Dependent variable mean	0.561	0.561	0.318	0.318	0.219	0.219	0.224	0.224
R^2	0.270	0.642	0.281	0.661	0.360	0.753	0.284	0.712

Notes: The table reports parametric estimates of yearly exposure effects, restricting the samples to individuals with siblings. The regression specification is equation (4), and the estimates can be compared to those reported in Table 2. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample.

Table A8: Overidentification Tests

(a) Cohort

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
-4 to -2		-0.0069***		-0.0005		0.0007		0.0015
		(0.0011)		(0.0012)		(0.0022)		(0.0021)
-1 to +1	0.0303***	0.0416***	0.0264***	0.0354***	0.0170***	0.0130***	0.0132***	0.0110***
	(0.0006)	(0.0014)	(0.0007)	(0.0017)	(0.0010)	(0.0035)	(0.0011)	(0.0034)
+2 to +4		-0.0069***		-0.0079***		0.0031		0.0008
		(0.0010)		(0.0011)		(0.0025)		(0.0023)
Num. of obs.	5,097,000	5,097,000	5,097,000	5,097,000	1,539,043	1,539,043	1,539,043	1,539,043
R^2	0.239	0.239	0.243	0.243	0.285	0.285	0.222	0.222

(b) Gender

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Same gender	0.0262***	0.0194***	0.0228***	0.0162***	0.0169***	0.0124***	0.0147***	0.0110***
	(0.0004)	(0.0007)	(0.0005)	(0.0007)	(0.0007)	(0.0012)	(0.0007)	(0.0013)
Other gender		0.0080***		0.0085***		0.0056***		0.0041***
		(0.0007)		(0.0007)		(0.0012)		(0.0012)
Num. of obs.	7,301,000	7,301,000	7,301,000	7,301,000	2,172,145	2,172,145	2,172,145	2,172,145
R^2	0.230	0.231	0.244	0.244	0.282	0.282	0.227	0.227

(c) Race

	Registered		Voted		Republican		Democrat	
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Same race	0.0239***	0.0205***	0.0209***	0.0189***	0.0188***	0.0184***	0.0161***	0.0154***
	(0.0004)	(0.0005)	(0.0005)	(0.0005)	(0.0008)	(0.0008)	(0.0008)	(0.0008)
Other race		0.0054***		0.0042***		0.0020**		0.0012**
		(0.0005)		(0.0005)		(0.0008)		(0.0006)
Num. of obs.	6,460,000	6,460,000	6,460,000	6,460,000	1,972,105	1,972,105	1,972,105	1,972,105
R^2	0.242	0.242	0.257	0.257	0.266	0.266	0.224	0.224

Notes: The table reports estimates from the overidentification tests described in Section 4.2.2 and summarized in Figure 3. Each panel corresponds to a distinct test, exploiting measures of exposure defined based on permanent residents' cohorts (Panel a), gender (Panel b), and race (Panel c). For each outcome, the first column reports parametric estimates of the yearly exposure effect estimated using permanent residents' outcomes specific to the mover's cohort group, gender, or race. The second column adds exposure measures for the other cohort groups, gender, or races. Race is collapsed into four categories: non-Hispanic white, non-Hispanic Black, Hispanic, and other non-Hispanic. We exclude voters with missing race information. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and party affiliation results use data from the partisanship sample.

Table A9: Parametric Estimates of Exposure Effects on Turnout – Census vs. Partisanship Samples

	Census			Partisanship		
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure effect, 0–19			0.0248***			0.0169***
			(0.0005)			(0.0008)
Exposure effect, 0–12	0.0151***	0.0177***		0.0067***	0.0117**	
	(0.0010)	(0.0021)		(0.0016)	(0.0057)	
Exposure effect, 13–19	0.0643***	0.0635***		0.0499***	0.0588***	
	(0.0025)	(0.0044)		(0.0042)	(0.0114)	
Exposure effect, 20–26	0.0101**	0.0090	0.0122***	0.0202***	0.0083	0.0215***
	(0.0042)	(0.0071)	(0.0042)	(0.0042)	(0.0113)	(0.0042)
Family fixed effects		✓			✓	
Number of observations	7,303,000	4,092,000	7,303,000	4,771,738	1,948,316	4,771,738
Dependent variable mean	0.325	0.318	0.325	0.446	0.456	0.446
R^2	0.242	0.661	0.242	0.344	0.703	0.344

Notes: The table reports parametric estimates of yearly exposure effects on turnout in an individual's first eligible election, separately for the Census sample (columns 1–3) and the partisanship sample (columns 4–6). Estimates in columns 1–2 and 4–5 are based on equation (4); columns 2 and 5 include family fixed effects. Estimates in columns 3 and 6 are based on equation (5). Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Data are from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and the partisanship sample.

Table A10: Parametric Estimates of Exposure Effects – Registered Voters Partisanship Sample

	Republican		Democrat	
	(1)	(2)	(3)	(4)
Exposure effect, 0–12	0.0112***	0.0038	0.0098***	0.0023
	(0.0015)	(0.0058)	(0.0015)	(0.0058)
Exposure effect, 13–19	0.0539***	0.0439***	0.0483***	0.0589***
	(0.0038)	(0.0101)	(0.0041)	(0.0107)
Exposure effect, 20–26	0.0175***	0.0044	0.0190***	0.0205*
	(0.0041)	(0.0108)	(0.0044)	(0.0110)
Difference $[13, 19]$ vs. $[0, 12]$	0.043	0.040	0.038	0.057
p-value $[13, 19] = [0, 12]$	0.0000	0.0012	0.0000	0.0000
Family fixed effects		✓		✓
Number of observations	2,149,840	878,081	2,149,840	878,081
Dependent variable mean	0.203	0.222	0.238	0.227
R^2	0.279	0.756	0.221	0.715

Notes: The table reports parametric estimates of yearly exposure effects, restricting the partisanship sample to individuals who register to vote at least once between 2012 and 2021. This restriction applies to both movers (who are used for estimation) and permanent residents (who are used to define exposure). The regression specification is equation (4), and the estimates can be compared to those in Table 2. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table A11: Parametric Estimates of Exposure Effects – Newly Eligible Voters

	Republican		Democrat	
	(1)	(2)	(3)	(4)
Exposure effect, 0–19	0.0169***	0.0163***	0.0151***	0.0153***
	(0.0006)	(0.0014)	(0.0005)	(0.0014)
Barely 18-year-olds		✓		✓
Number of observations	1,734,048	244,821	1,734,048	244,821
Dependent variable mean	0.193	0.194	0.245	0.234
R^2	0.075	0.084	0.069	0.070

Notes: The table reports parametric estimates of yearly exposure effects, restricting the sample in two ways. In all columns, the regression excludes individuals who were born on the first day of any given month, as such dates of birth could be accurate only to the month or year, as discussed in Section 2.1. In columns 2 and 4, the sample is further restricted to individuals who turned 18 in the 100 days before Election Day and are therefore newly eligible to vote. The regression specification is equation (5). Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively.

Table A12: Persistence of Childhood Exposure Effects

(a) Registration and turnout

	Registered			Voted		
	1st	2nd	3rd	1st	2nd	3rd
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure effect, 0–19	0.0337***	0.0253***	0.0234***	0.0292***	0.0221***	0.0234***
	(0.0008)	(0.0008)	(0.0008)	(0.0009)	(0.0010)	(0.0009)
Number of observations	4,194,000	4,194,000	4,194,000	4,194,000	4,194,000	4,194,000
Dependent variable mean	0.536	0.616	0.656	0.297	0.264	0.380
R^2	0.256	0.247	0.245	0.277	0.256	0.238

(b) Party affiliation

	Republican			Democrat		
	1st	2nd	3rd	1st	2nd	3rd
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure effect, 0–19	0.0152***	0.0142***	0.0135***	0.0123***	0.0119***	0.0109***
	(0.0012)	(0.0011)	(0.0011)	(0.0012)	(0.0011)	(0.0011)
Number of observations	1,331,139	1,331,139	1,331,139	1,331,139	1,331,139	1,331,139
Dependent variable mean	0.184	0.209	0.209	0.206	0.257	0.295
R^2	0.297	0.290	0.285	0.232	0.231	0.231

Notes: The table reports parametric estimates of yearly exposure effects in the first, second, and third election for which an individual is eligible to vote. The sample is restricted to individuals whom we can observe in all three elections, that is, for birth cohorts 1992 to 1998. The regression specification is equation (5). Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and partisanship results use data from the partisanship sample.

Table A13: Summary Statistics for Multiple Movers

	Census sample			Partisanship sample		
	Mean	Std. dev.	Num. of obs.	Mean	Std. dev.	Num. of obs.
Birth year	1998	2.76	1,213,000	1997	2.90	188,176
Male	0.510	0.500	1,213,000	0.513	0.500	176,969
White	0.627	0.484	1,055,000	0.828	0.377	181,272
Black	0.128	0.334	1,055,000	0.047	0.212	181,272
Hispanic	0.148	0.355	1,055,000	0.075	0.264	181,272
Other race	0.097	0.296	1,055,000	0.049	0.216	181,272
Number of children in family	1.62	0.81	1,213,000	1.64	0.78	188,176
Two-time mover	0.691	0.462	1,213,000	0.793	0.405	188,176
Age at move 1	6.7	5.01	1,213,000	6.2	5.38	188,176
Age at move 2	10.4	5.60	1,213,000	10.5	6.51	188,176
Age at move 3	14.5	5.17	375,000	15.9	6.09	38,913
Move 1 distance (mi.)	493.5	699.0	1,206,000	383.9	638.7	188,176
Move 2 distance (mi.)	468.5	663.6	1,208,000	428.4	659.1	188,176
Move 3 distance (mi.)	476.6	656.8	374,000	503.2	692.7	38,913
Linked to voter file	0.784	0.411	1,213,000	–	–	–
Registered in first election	0.566	0.496	1,213,000	–	–	–
Voted in first election	0.322	0.467	1,213,000	–	–	–
Abs. exp. diff. in reg. (move 1)	0.105	0.095	1,213,000	–	–	–
Abs. exp. diff. in voted (move 1)	0.087	0.077	1,213,000	–	–	–
Abs. exp. diff. in reg. (move 2)	0.111	0.096	1,213,000	–	–	–
Abs. exp. diff. in voted (move 2)	0.090	0.079	1,213,000	–	–	–
Abs. exp. diff. in reg. (move 3)	0.116	0.099	375,000	–	–	–
Abs. exp. diff. in voted (move 3)	0.095	0.081	375,000	–	–	–
Republican in first election	–	–	–	0.205	0.404	188,176
Democrat in first election	–	–	–	0.225	0.418	188,176
Abs. exp. diff. in Rep. (move 1)	–	–	–	0.089	0.079	188,176
Abs. exp. diff. in Dem. (move 1)	–	–	–	0.102	0.087	188,176
Abs. exp. diff. in Rep. (move 2)	–	–	–	0.095	0.082	188,176
Abs. exp. diff. in Dem. (move 2)	–	–	–	0.104	0.088	188,176
Abs. exp. diff. in Rep. (move 3)	–	–	–	0.101	0.086	38,913
Abs. exp. diff. in Dem. (move 3)	–	–	–	0.106	0.090	38,913

Notes: The table reports summary statistics for the sample of multiple movers. We make the same sample restrictions as in Section 5: two- and three-time movers who spent two to six years in the first destination, either during ages 0 to 12 or 13 to 19. Individuals who lived in the first destination during both age ranges are excluded. Data are from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and the partisanship sample.

Table A14: Parametric Estimates for Cross-ZIP Code Moves

(a) Registration and turnout

	Registered			Voted		
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure effect, 0–19	0.0119***	0.0155***	0.0082***	0.0084***	0.0114***	0.0069***
	(0.0008)	(0.0029)	(0.0011)	(0.0007)	(0.0027)	(0.0010)
Across counties		✓			✓	
Within county			✓			✓
Mean move dist. (mi.)	134.2	287.3	6.5	134.2	287.3	6.5
Number of observations	8,153,000	3,707,000	4,446,000	8,153,000	3,707,000	4,446,000
Dependent variable mean	0.589	0.591	0.587	0.355	0.359	0.353
R^2	0.560	0.799	0.509	0.569	0.799	0.520

(b) Party affiliation

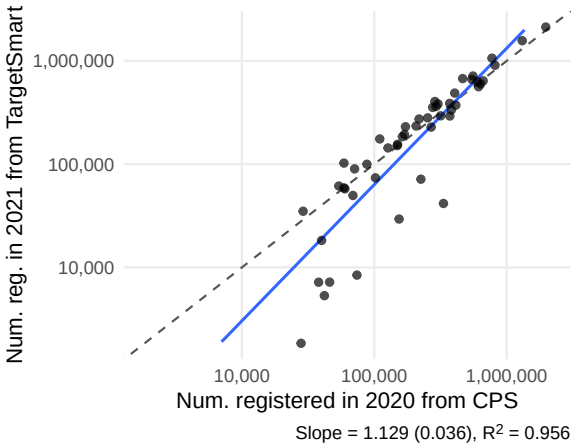
	Republican			Democrat		
	(1)	(2)	(3)	(4)	(5)	(6)
Exposure effect, 0–19	0.0052***	0.0084**	0.0025	0.0059***	0.0062*	0.0041**
	(0.0013)	(0.0038)	(0.0018)	(0.0013)	(0.0037)	(0.0017)
Across counties		✓			✓	
Within county			✓			✓
Mean move dist. (mi.)	137.2	308.1	7.2	137.2	308.1	7.2
Number of observations	1,819,079	388,088	1,110,371	1,819,079	388,088	1,110,371
Dependent variable mean	0.180	0.186	0.173	0.260	0.250	0.268
R^2	0.504	0.658	0.471	0.486	0.645	0.462

Notes: The table reports parametric estimates of yearly exposure effects, with exposure and moves defined at the ZIP code level instead of the county. For each outcome, the first column uses all cross-ZIP movers, the second column reports estimates restricting the sample to individuals who move across counties, and the third column restricts the sample to individuals who move within the same county. The regression specification is equation (5). To correct for measurement error when computing means within ZIP code-by-cohort cells with small numbers of permanent residents, we use a split-sample approach to constructing the exposure measures, detailed in Section 6.2. In Panel b, we report observation counts *after* dropping singletons, since standard errors are not computed properly otherwise. Heteroskedasticity-robust standard errors are in parentheses. ***, **, and * indicate significance at the 1%, 5%, and 10% levels, respectively. Registration and turnout results use data from the Census sample (FSRDC Project Number 2747, CBDRB-FY26-P2747-R12874) and partisanship results use data from the partisanship sample.

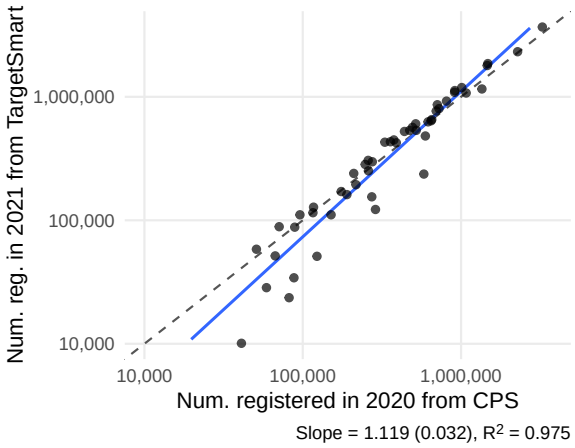
B Data quality checks

Figure B1: Registered Voters by State and Age, TargetSmart vs. Current Population Survey

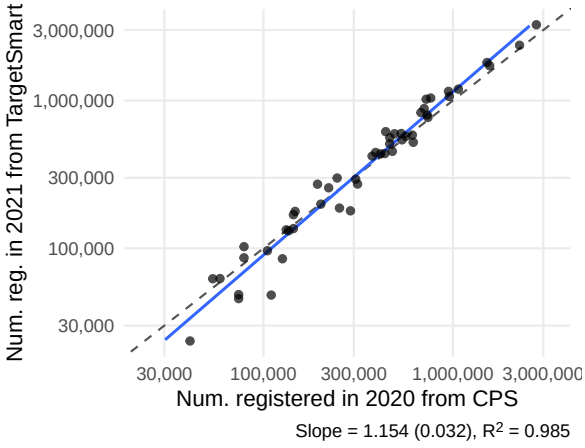
(a) 18 to 24



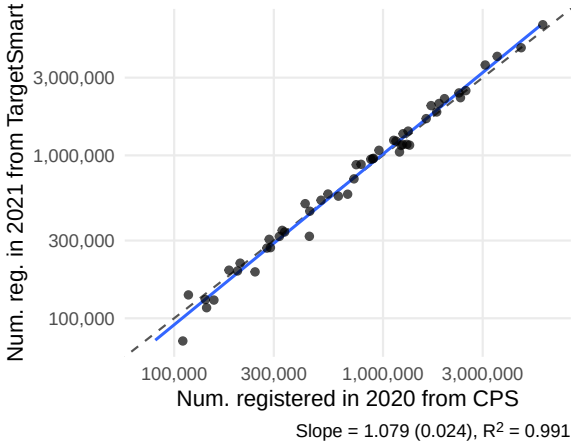
(b) 25 to 34



(c) 35 to 44

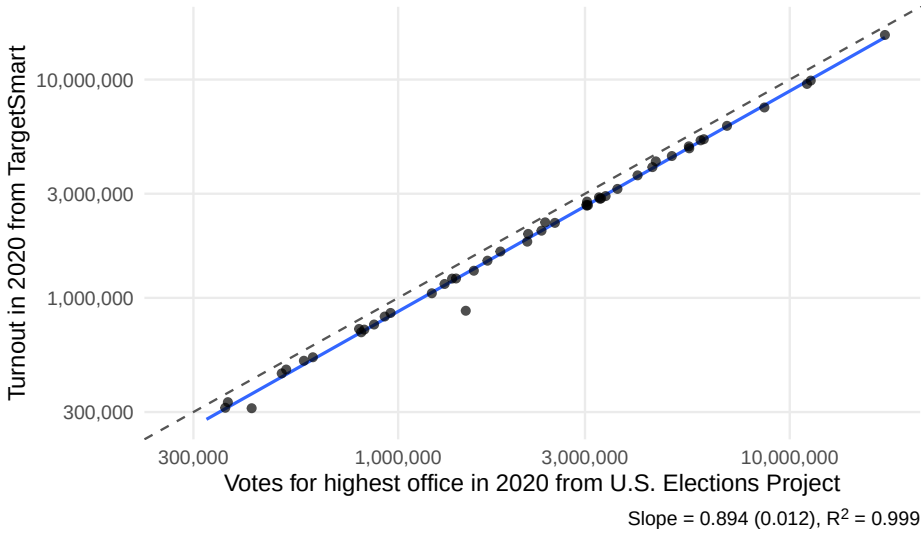


(d) 45 to 64



Notes: The figure plots the relationship between 2021 state-level counts of registered voters in the TargetSmart data (on the vertical axis) and estimated state-level counts of registered voters according to the 2020 CPS Voting and Registration Supplement (on the horizontal axis). Each panel plots this relationship for a different age group: 18 to 24 in Panel a, 25 to 34 in Panel b, 35 to 44 in Panel c, and 45 to 64 in Panel d. The scales are logarithmic. In each panel, the blue line is a linear fit, and we present its slope and R^2 in the bottom right. The dashed reference lines have slopes of 1.

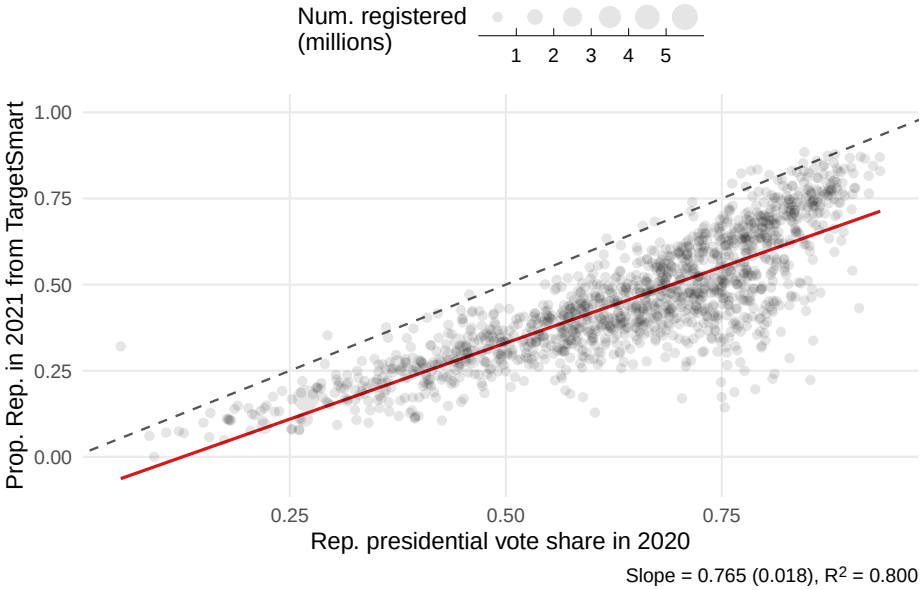
Figure B2: Turnout in 2020, TargetSmart vs. United States Election Project



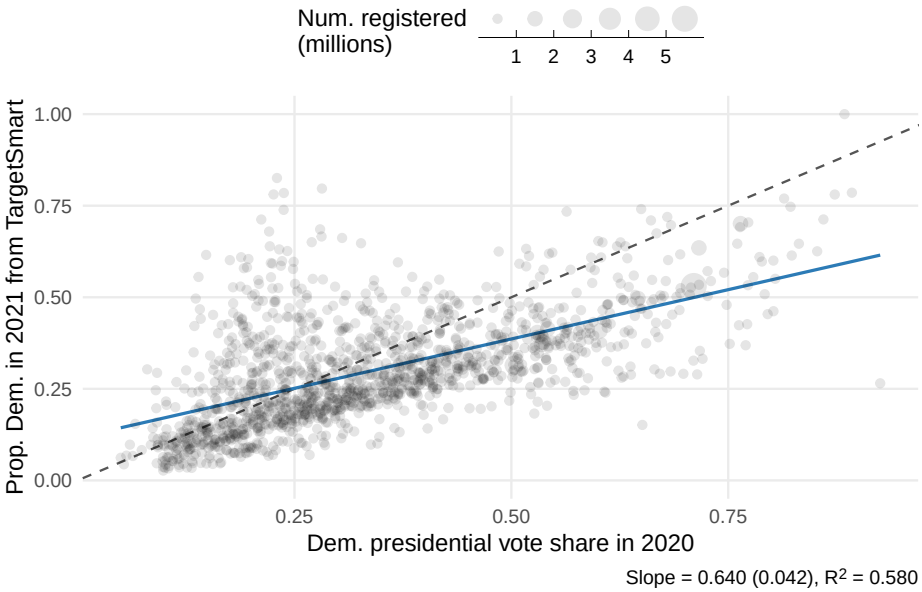
Notes: The figure plots the relationship between state-level counts of people who voted in the 2020 general election according to the 2021 TargetSmart voter file (on the vertical axis) and 2020 counts of ballots cast for the highest office on each state's ballot according to the United States Election Project (McDonald, 2022) (on the horizontal axis). The scales are logarithmic. The blue line is a linear fit, and we present its slope and R^2 in the bottom right. The dashed reference line has a slope of 1.

Figure B3: Share of Partisans in TargetSmart vs. Presidential Vote Share in 2020

(a) Republican

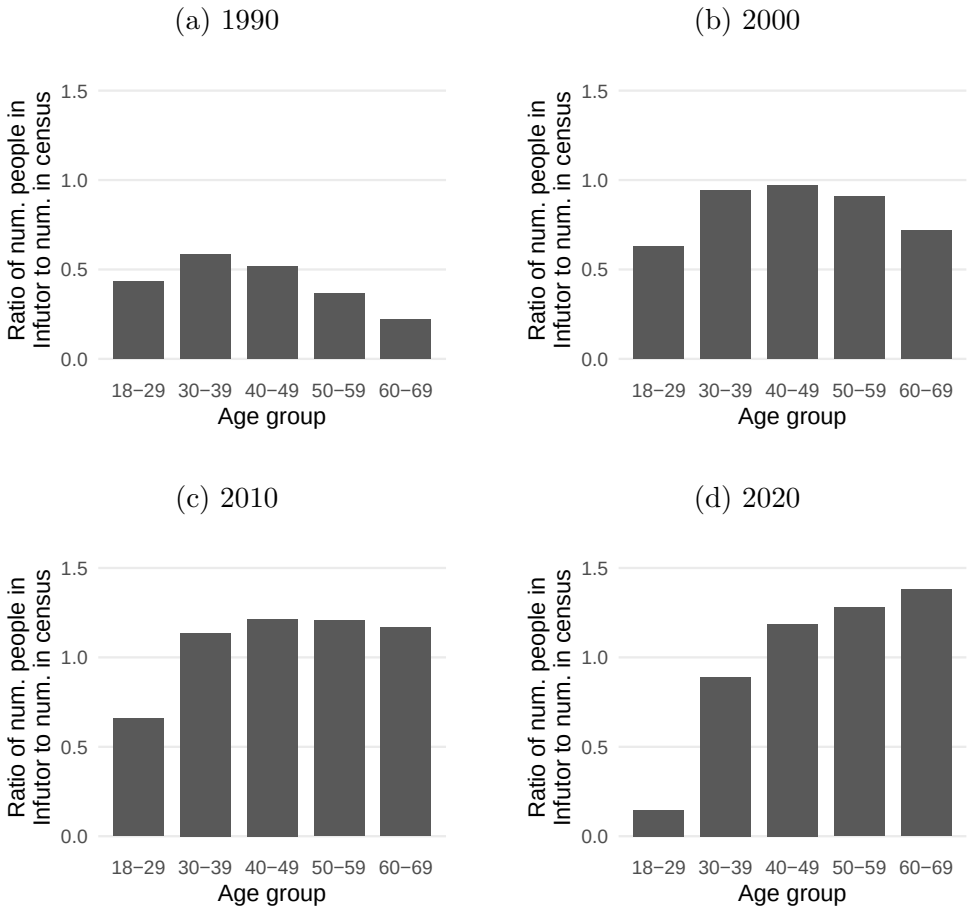


(b) Democrat



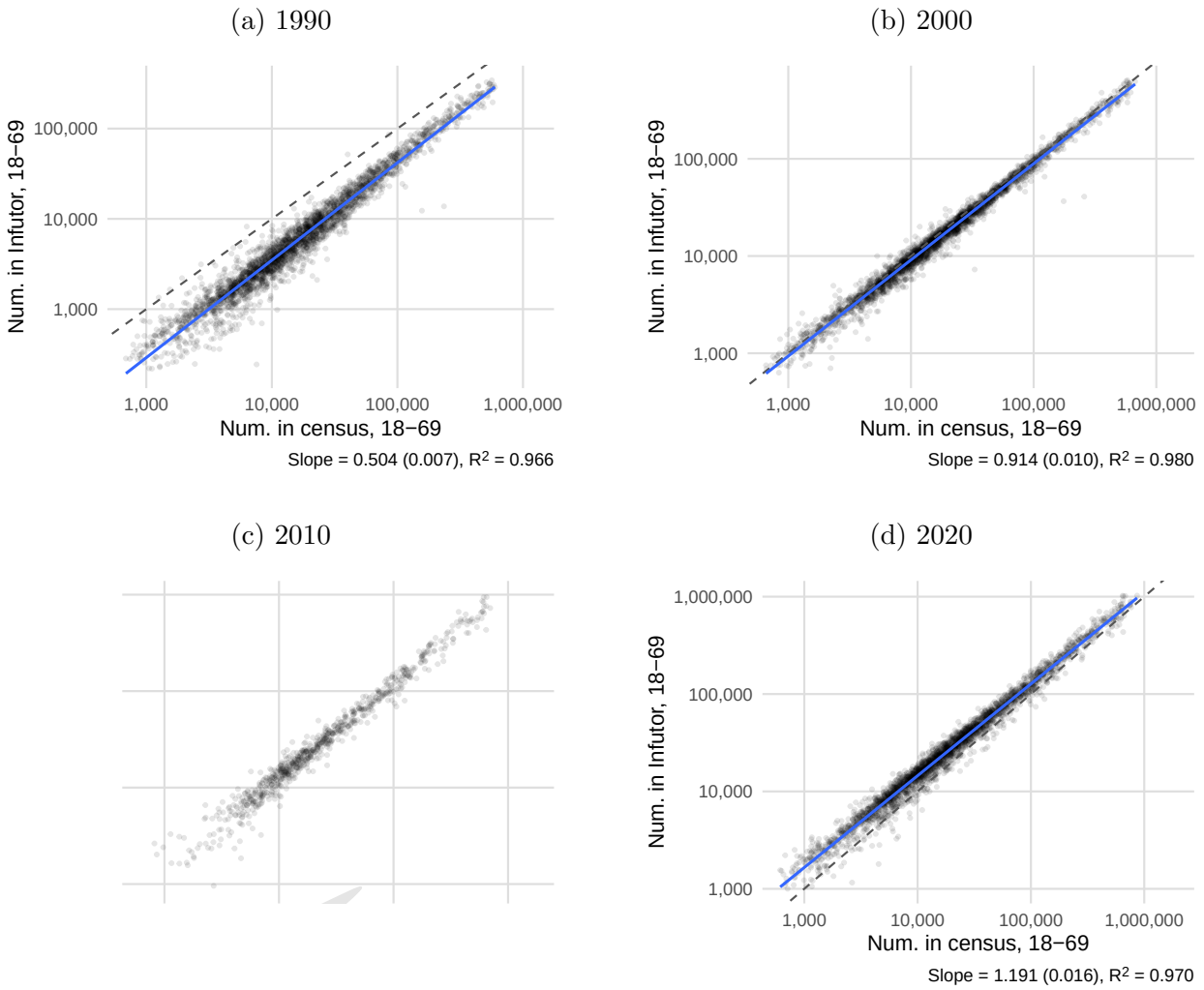
Notes: In Panel a, we show the proportion of people registered with the Republican Party in the 2021 TargetSmart voter file in each county on the vertical axis. On the horizontal axis, we show the proportion of the two-party vote for the Republican candidate (Donald Trump) in the 2020 general election in that county. In Panel b, we repeat this exercise for the Democratic candidate (Joe Biden). The red and blue lines are linear fits weighted by the number of registered voters in each county, and we present their slopes and R^2 in the bottom right of each panel. The dashed reference lines have slopes of 1.

Figure B4: Number of People in Infutor Relative to Census Population



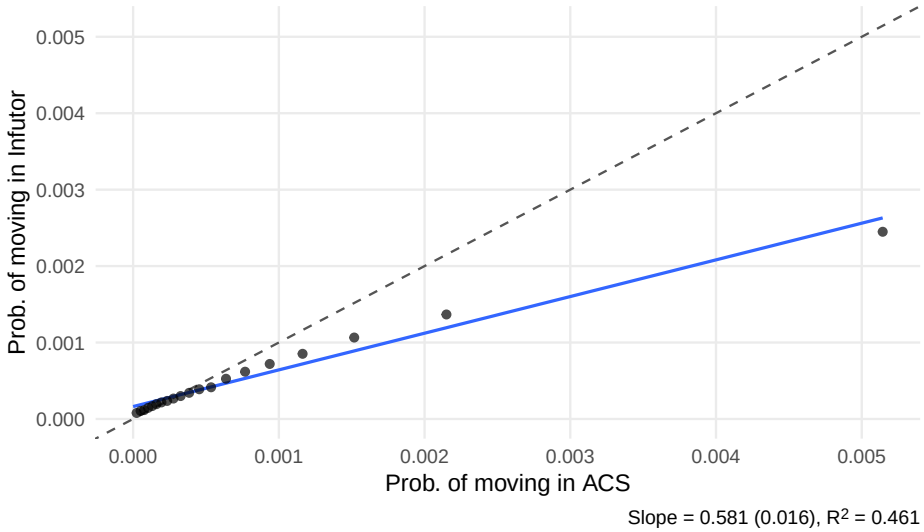
Notes: We show the ratio between the number of people in the Infutor data and the total number of people in the census by age group, for four decennial censuses: 1990 in Panel a, 2000 in Panel b, 2010 in Panel c, and 2020 in Panel d. We exclude individuals listed as "deceased" in Infutor.

Figure B5: County-Level Counts of Infutor vs. Census Population



Notes: On the vertical axis, we show the number of people between 18 and 69 in the Infutor data in a particular county, as of a particular decennial census year. On the horizontal axis, we show the number of people between 18 and 69 recorded in that year's census in the same county. We do this for four decennial censuses: 1990 in Panel a, 2000 in Panel b, 2010 in Panel c, and 2020 in Panel d. The scales are logarithmic. In each panel, the blue line is a linear fit, and we present its slope and R^2 in the bottom right. The dashed reference lines have slopes of 1. We exclude individuals listed as "deceased" in Infutor, and we exclude counties below the 1st percentile and above the 99th percentile of the Infutor or census population distributions.

Figure B6: Probability of Moving in Infutor vs. American Community Survey



Notes: On the vertical axis, we show across-state move probabilities in the Infutor data for parents of young voters. On the horizontal axis, we show the corresponding probabilities from the ACS. To construct move probabilities from the Infutor data, we focus on parents (either identified directly in the Infutor data or identified in the TargetSmart data and subsequently linked to Infutor), count the number of people moving across states in each year, and divide it by the number of individuals in the origin state. To construct move probabilities from the ACS, we use annual individual-level data from the ACS from 2000 to 2019 and construct a sample of individuals that corresponds to the parents in the Infutor data. Specifically, we subset to U.S. citizens residing in the U.S. for at least one year, who have at least one own child living in the same household, who are at least 15 years older than their oldest child, and whose children's ages align with the children's ages in our sample, i.e., their oldest child was born in the youngest cohort observed in our sample or before. To obtain move probabilities, we count the number of people moving across states within a given year and divide by the number of people matching the above criteria in the origin state, where both counts are weighted by the appropriate sampling weights. We then match the move probabilities from the two sources by origin, destination, and year. To construct the figure, we bin the probabilities into ventiles, weighting by the total number of people living in the origin state. The blue line is a linear fit, and we present its slope and R^2 in the bottom right. The dashed reference line has a slope of 1.

C Data construction

This appendix describes how we de-duplicate and process the TargetSmart and Infutor data, match parents with their children, and build and clean parents' residential histories, which we use to infer their children's residential histories.

C.1 Processing TargetSmart data

The raw TargetSmart data consist of 510 state-by-year snapshots, one for each of the 50 states and the District of Columbia, for each year between 2012 and 2021. These snapshots contain each voter's name, residential address, gender, race, date of birth, whether or not they were registered and voted in prior elections, and their party affiliation.

TargetSmart also provides a "voterbase ID" (henceforth VBID), which uniquely identifies each row for a given state and year, and an "exact track ID" (henceforth ETID), which represents the company's efforts to link individuals *across* years and states. We use VBID and ETID information, as well as voters' first name, middle name, last name, date of birth, and vote history to de-duplicate the TargetSmart data, so that for each state-by-year file, each observation used in the analysis represents each voter's most likely current record. We also use the same information to expand on TargetSmart's linkage: we further link records that likely correspond to the same individual, first within states, then across states.

C.1.1 Cleaning the raw TargetSmart state-by-year files

To begin, we clean and de-duplicate the raw TargetSmart state-by-year files as follows:

1. We remove hyphens and spaces from first and last names. We capitalize all letters of first and last names.
2. We recode invalid ZIP codes and census block group IDs as missing.
3. We use TargetSmart's information on whether a voter is found in Social Security death records to drop deceased voters.[C1]
4. We use TargetSmart's information from the United States Postal Service National Change of Address database to drop voters who no longer reside at their listed residence.
5. We de-duplicate records with the same ETID, first name, and last name, giving preference to the record whose registration status is "Registered" (versus "Unregistered"), whose voter status is "Active" (versus "Inactive," based on recent election participation), and with the most recent registration date. We do this because, for a given

[C1] Specifically, we drop voters indicated as "deceased" and who did not vote in the most recent general election before the year of the voter file.

individual, outdated records may exist alongside the active, most recent record in the same state voter file snapshot. When de-duplicating, we harmonize the vote history so that if any of the dropped records indicate that the person voted in a particular election, we maintain this turnout history in the record that we preserve for that individual.

6. We repeat the previous step, de-duplicating records with the same first name, middle initial, last name, date of birth, and ZIP code.[C2]

7. We drop any record where the voter's age is listed as under 18 and the individual is listed as "Registered."

C.1.2 Linking records within states

To link rows *within the same state* corresponding to the same individual but across multiple years – in other words, to assign a state unique identifier (henceforth "SUID") – we take the following steps:

1. We split the date of birth field into year, month, and day. The raw voter lists that TargetSmart receives from some states do not record the day of birth (only the month and year) and, in some cases, they only record year of birth. Even though TargetSmart uses commercial data to reconstruct the exact day of birth for voters from these states, the TargetSmart data feature high frequencies of dates of birth ending in `01` (i.e., indicating a voter born on the first day of a given month) and `0101` (i.e., indicating a voter born on January 1). When the date of birth ends in `01`, we set the *day* of birth to missing. When the date of birth ends in `0101`, we set the *month and day* of birth to missing.[C3]

2. The VBID uniquely identifies a row in each state-by-year snapshot, but when the same individual appears in multiple years in the same state, they typically appear with the same VBID. Therefore, we assume records that share a VBID are the same person, and assign them the same SUID. However, if TargetSmart assigned the same VBID to multiple rows where the first names, last names, and dates of birth are *all* different, or where the maximum difference in birth years is more than five years *and* the months and days of birth are different too, we break this link.

3. We group observations by ETID.

[C2] We consider groups of rows in which the middle initial is missing for one or more and the rest of these variables are identical to be duplicates.

[C3] We lose information by setting the month or day of birth to missing for some people who were actually born on the first of the month or on January 1. However, there is no reliable way to determine whether a date of birth ending in `0101` actually corresponds to a January 1 birthday, or whether it indicates that the month and day of birth are missing.

- If *at least one* of the first names, last names, and dates of birth is the same among all observations sharing the same ETID, there is only one record per year, and the maximum age difference is less than or equal to five years, then we assign the same SUID to all these observations.
- If not everyone in that set of observations shares *either* a first name, last name, or date of birth, we group them by name and date of birth and – as long as there is only one record per year – assign rows within each group the same SUID.

4. We group by first four letters of first name, last name, year of birth, and address. We require that:
   - Each record has non-missing information for all of the grouping variables.
   - Each record is unique *within a year* by these variables.

   If so, we assign these rows the same SUID.
5. We repeat the previous step using the following sets of grouping variables:[C4]
   - First four letters of first name, last name, year of birth, and address.[C5]
   - First name, last name, and address.
   - First name, last name, and date of birth.
   - First name, last name, and year and month of birth.
   - First name, last name, and year of birth.
   - First name, middle name, last name, and date of birth.
   - First name, middle name, last name, and year and month of birth.
   - First name, middle name, last name, and year of birth.

Appendix Table C1 summarizes the results of this procedure.

C.1.3 Linking records across states

To link rows corresponding to the same individual *across* states and assign them a nationally unique identifier (henceforth "UID"), we take the following steps.

1. We drop rows with missing first name, last name, or date of birth.

[C4] If month and day of birth are not part of the grouping variables, we require that, when non-missing, they do not conflict. We also require that middle initials, when non-missing, do not conflict, except if the set of grouping variables includes the address and year of birth, or if it includes the exact date of birth.

[C5] We only de-duplicate records if no more than 15 people ever lived at the same address, and if the difference between the maximum and minimum years of birth is no more than five years.

Table C1: Assigning State UIDs

Step	Num. unique	Prop. of original	Complete
Start (unique VBIDs)	436,736,545	1.000	-
Break links with different name/DOB	436,823,667	1.000	-
Break links born > 5 years apart	436,911,503	1.000	-
Link by ETID	385,073,781	0.882	-
Link by f4, last, birth yr., exact address	376,985,662	0.863	0.864
Link by first, last, exact address	367,761,419	0.842	0.993
Link by first, last, DOB	356,382,264	0.816	0.583
Link by first, last, birth yr./mo.	351,100,148	0.804	0.733
Link by first, last, birth yr.	345,771,012	0.792	0.870
Link by first, middle, last, DOB	345,703,988	0.792	0.514
Link by first, middle, last, birth yr./mo	345,563,884	0.791	0.636
Link by first, middle, last, birth yr.	344,888,280	0.790	0.754

Notes: We describe the results of our procedure to link records that correspond to the same person within a state in the TargetSmart data. The pipeline proceeds sequentially through each row of the table. "Num. unique" refers to the number of unique individuals at each step, and "prop. of original" refers to the ratio between the number of unique individuals at each step and the original number of unique individuals. "Complete" is the proportion of rows with non-missing information for the relevant variables at each step. "f4" refers to the first four letters of the individual's first name. Although we perform this procedure separately for each state, the table presents numbers aggregated across states.

2. As above, if the date of birth ends in `01`, we set the day of birth to missing. If the date of birth ends in `0101`, we set the month and day of birth to missing.

3. We group by ETID and check that the group's total vote count for every election is at most one.[C6] If so, we assign these rows the same UID.

4. We group by first name, last name, and date of birth. We require that:

   - Each record has non-missing information for all grouping variables.
   - Each record is unique by these variables within state.
   - The group has a record from at least two states.
   - The maximum vote count among records in the group for any election is one.

   If so, we assign these rows the same UID.

[C6]To compute the "vote count" for a given election, we take all unique SUID-by-ETID observations with the same ETID (harmonizing vote history by SUID) and sum an indicator equal to one if the observation is indicated as having voted. We therefore require that, for each person we propose to de-duplicate, at most one record linked to that person is recorded as having voted in each election. In other words, we do not de-duplicate rows if this would imply that the same individual voted more than once in the same election in two different states, which would be illegal. We impose this restriction because double voting is extremely rare: Goel et al. (2020) estimate that about 0.02% or less of votes cast in 2012 were double votes.

Table C2: Assigning National UIDs

Step	Num. unique	Prop. of original	Complete	Unique
Start (unique SUIDs)	344,888,280	1.000	-	-
Link by ETID	331,025,637	0.960	-	-
Link by first, last, DOB	320,991,306	0.931	0.619	0.995
Link by first, last, birth yr./mo.	317,964,337	0.922	0.752	0.968
Link by first, last, birth yr.	313,785,450	0.910	0.876	0.863
Link by first, middle, last, DOB	313,711,203	0.910	0.539	0.998
Link by first, middle, last, birth yr./mo.	313,280,345	0.908	0.641	0.996
Link by first, middle, last, birth yr.	311,512,774	0.903	0.747	0.986

Notes: We describe the results of our procedure to link records that correspond to the same person across states in the TargetSmart data. The pipeline proceeds sequentially through each row of the table. "Num. unique" refers to the number of unique individuals at each step, and "prop. of original" refers to the ratio between the number of unique individuals at each step and the original number of unique individuals. "Complete" is the proportion of rows with non-missing information for the relevant variables. "Unique" is the proportion of rows with non-missing information for the relevant variables that are unique within state by those variables. At each step, in order to be linked to another observation, we require that an observation has complete information for the relevant variables and is unique within each state by those variables.

5. We repeat the previous step using the following sets of grouping variables:[C7]

    - First name, last name, and date of birth.
    - First name, last name, and year and month of birth.
    - First name, last name, and year of birth.
    - First name, middle name, last name, and date of birth.
    - First name, middle name, last name, and year and month of birth.
    - First name, middle name, last name, and year of birth.

Appendix Table C2 summarizes the results of this procedure.

C.2 Processing Infutor data

The raw Infutor data consist of 51 files, one for each of the 50 states and the District of Columbia. The data contain the name, gender, year and month of birth, and up to 10 addresses associated with each individual/row, along with a date (year and month) at which Infutor considers that address effective.

[C7] As above, if month and day of birth are not part of the grouping variables, we require that, when non-missing, they do not conflict. We also require that middle initials, when non-missing, do not conflict, *except* if the set of grouping variables includes exact date of birth.

Table C3: Infutor De-Duplication

Step	Num. unique	Prop. of original
Start (unique IDs)	644,033,829	1.000
De-duplicate by f4, last name, address, birth yr.	623,140,262	0.968
De-duplicate by first name, last name, address	571,289,316	0.887

Notes: We describe the results of our procedure to de-duplicate records that correspond to the same person within the Infutor data. The pipeline proceeds sequentially through each row of the table. "Num. unique" refers to the number of unique individuals at each step, and "prop. of original" refers to the ratio between the number of unique individuals at each step and the original number of unique individuals. "f4" refers to the first four letters of the individual's first name.

We use two methods to de-duplicate the Infutor records, which are uniquely identified (across all address observations for a given individual) by "persistent ID" (henceforth PID). First, we group observations by first four letters of first name, last name, year of birth, and address. We require that:

- There are no more than 10 distinct PIDs in the group.
- Name suffixes, when non-missing, do not conflict.
- No more than 15 people ever live at the address.

If so, we assign these rows the same PID. Then, we repeat the same procedure, this time grouping by first name, last name, and address. In addition to the above restrictions, we require that:

- Middle initials, when non-missing, do not conflict.
- The difference between the maximum and minimum years of birth is no more than five years.

Appendix Table C3 summarizes the results of this procedure.

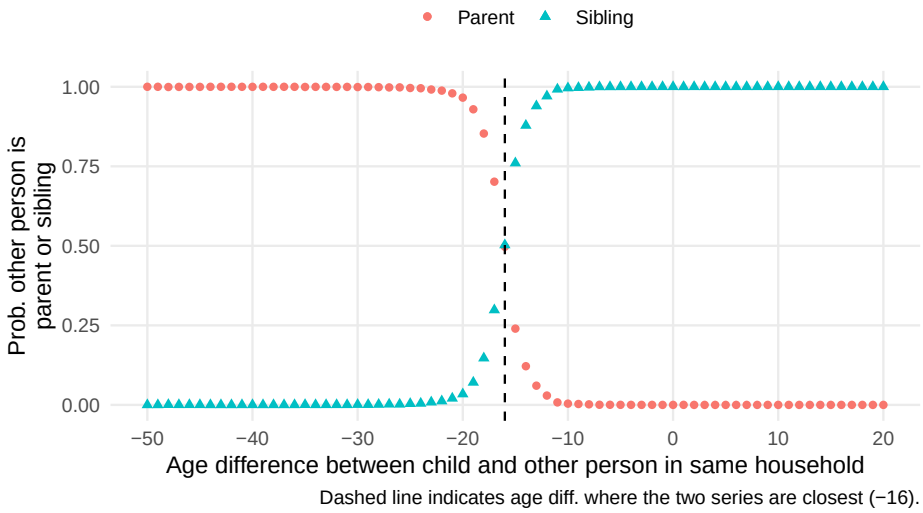
C.3 Linking children to parents

To link children to their parents, we employ two methods, using, respectively, (1) residential addresses from TargetSmart and Infutor, and (2) TargetSmart's household identifier field. We proceed as follows:

Method 1: Address

1. We stack the TargetSmart and Infutor addresses.
2. We merge children with potential parents by address (but not including unit or apartment number) and first four letters of the last name. We keep these matches only

Figure C1: Probability of Family Relationship vs. Age Difference



Notes: We use data from the 2021 ACS to show the probability that an individual living in the same household as a child is a "likely parent" or "likely sibling" by age difference between the child and the other person. Specifically, we subset the full ACS to individuals who are citizens, at most 21 years old, and who are classified as children. We then interact these individuals with all other people in the same household who are either classified as household head or spouse, individuals whom we refer to as "likely parents," or children, whom we refer to as "likely siblings." For a given age difference between two individuals, we then calculate the probability that these two individuals have either a parent-child relationship or a child-child relationship.

if:

- Parents are at least 16 years older than their children.[C8]
- Either there are fewer than 15 people who ever live at that address, or (1) the child and parent's last names match exactly, (2) the child and parent's addresses have the same unit number, and (3) fewer than 15 people share that address, unit number, and last name combination.

Method 2: Household identifier

1. We merge children with potential parents by TargetSmart's household ID. We keep these matches only if:
   - Parents are at least 16 years older than their children.
   - Fewer than 15 people share that household ID.

[C8] According to ACS data, this is the age difference at which it becomes more likely that two individuals in the same household have a parent-child relationship as opposed to a sibling relationship. See Appendix Figure C1.

Table C4: Linking Children to Parents

	Num.	Prop. of total
Total number of children	37,058,371	1.000
Linked using TargetSmart only	3,474,911	0.094
Linked using TargetSmart only, by address	3,155,076	0.085
Linked using TargetSmart only, by HH ID	3,249,468	0.088
Linked using Infutor only	1,259,912	0.034
Linked using TargetSmart and Infutor	18,289,016	0.494
Linked using TargetSmart or Infutor	23,350,299	0.630

Notes: We describe the results of our procedure to link children to their parents. "Num." refers to the number of unique individuals in each category, and "prop. of total" refers to the ratio between the number of individuals in each category and the total number of children, i.e., those who are born between 1992 and 2002. The groups are not mutually exclusive – that is, an individual can be linked using more than one method. "HH ID" refers to the household identifier assigned by TargetSmart.

Appendix Table C4 summarizes the results of these linking processes. Overall, 63.0% of individuals in the birth cohorts in our analysis, 1992-2002, are linked to at least one parent. We explain later how, in cases in which a child is linked to multiple parents, we choose which parent we use to infer the child's location history.

C.4 Creating the ZIP code panel

C.4.1 Merging Infutor and TargetSmart data

Next, we use the Infutor data to reconstruct young voters' residential histories by assembling residential histories for their parents. To do this, for children whom we link to a parent in TargetSmart (and not Infutor directly), we need to identify the parent's Infutor address history. We do this as follows:

1. As previously, in TargetSmart, if the date of birth ends in `0101`, we set the month of birth to missing. In Infutor, if the date of birth ends in `01`, we set the month of birth to missing.[C9]

2. We merge the set of records identified as parents in the TargetSmart data with the full Infutor address history by first name, last name, address, and month and year of birth.

3. We repeat the previous step using the following sets of merging variables. In each round, all rows are candidates to be merged (not only previously un-merged rows). We keep all the unique Infutor PIDs matched to a given parent in the TargetSmart data in any of the merge rounds.[C10]

[C9] Recall that the Infutor data only report month and year of birth.

[C10] For each merge round, the eligible TargetSmart rows must be uniquely identified by the merging variables.

- First name, last name, address, and month and year of birth.
- First name, last name, address, and year of birth.
- First four letters of first name, first four letters of last name, and address.
- First four letters of first name, first four letters of last name, month and year of birth, and ZIP code.
- First four letters of first name, first four letters of last name, year of birth, and ZIP code.
- First name, last name, and ZIP code.

4. We remove TargetSmart-Infutor matches with conflicting dates of birth. Specifically, we drop matches if the Infutor date of birth falls outside a 12-month window around the TargetSmart date of birth.[C11]

Appendix Table C5 summarizes the results of these merges. Overall, 87.1% of TargetSmart UIDs are successfully merged to an Infutor PID.

C.4.2 Creating ZIP code histories for parents

Our analysis uses location information from the Infutor data alone, not TargetSmart. Therefore, we drop the 1.8% of children whose parents were identified only in the TargetSmart data and not linked to the Infutor data. The remaining children – who constitute 61.2% of the total number of 18 to 26 year olds in the TargetSmart data – have at least one parent either identified directly in the Infutor data or identified in the TargetSmart data and successfully linked to the Infutor data.[C12]

We define moves at the ZIP code or county level: We do not analyze moves to different addresses in the same ZIP code. Thus, we build a panel of parents' residential histories at the ZIP code level by stacking all ZIP codes associated with all Infutor records linked to a parent. We require that each parent is observed at only one ZIP code within a month. Since

However, we allow one TargetSmart ID to be merged to multiple Infutor PIDs. We require that middle initials, when non-missing, do not conflict, except if the set of grouping variables includes exact address and year of birth.

[C11] Like for TargetSmart, Infutor dates of birth are sometimes approximate (e.g., January of the birth year). Thus, using the exact month and year of birth to resolve conflicts may lead to incorrectly dropping too many matches. We treat such records differently depending on whether there are other matched Infutor records with a non-missing date of birth that conflicts with TargetSmart's. Specifically, if all non-missing Infutor dates of birth are within ±12 months of TargetSmart's, we also keep Infutor records with missing date of birth; otherwise, we drop them (along with Infutor records with non-missing dates of birth conflicting with the TargetSmart date of birth).

[C12] If a young voter is linked to the same parent in the Infutor data both directly and indirectly through the TargetSmart-Infutor merge, we prioritize the latter, so we preserve information from the voter file data about parent demographics and political behavior. We also require that each Infutor PID is matched to no more than one TargetSmart UID.

Table C5: Merging TargetSmart and Infutor

	Num.	Prop. of total
Total number of TS parents	36,278,395	1.000
Merged using first, last, birth yr./mo., address	21,566,903	0.594
Merged using first, last, birth yr., address	25,962,928	0.716
Merged using f4, l4, address	29,193,648	0.805
Merged using f4, l4, birth yr./mo., ZIP	22,909,512	0.631
Merged using f4, l4, birth yr., ZIP	27,568,466	0.760
Merged using first, last, ZIP	29,577,404	0.815
Merged using any method	31,598,036	0.871

Notes: We describe the results of our procedure to merge parents identified in the TargetSmart data to Infutor. "Num." refers to the number of unique individuals merged using that set of variables, and "prop. of total" refers to the ratio between the number of individuals in that group and the total number of parents we identify in the TargetSmart data. The groups are not mutually exclusive – that is, an individual can be merged using more than one set of variables. "f4" refers to the first four letters of the individual's first name and "l4" refers to the first four letters of the last name.

Infutor combines multiple sources of address information, the data may contain duplicate address observations with slightly different dates or variants of the way the address was recorded. Thus, for each parent, we de-duplicate the ZIP code history as follows:

1. If a person is observed first in ZIP code A and then in ZIP code B in the same month, and they are next observed in ZIP code A, we drop the observation in ZIP code B.
2. For remaining duplicates by month, we keep the ZIP code that appears more often in the future. For example, if a person is observed in ZIP codes A and B in the same month and is later observed three times in ZIP code A but only once in ZIP code B, we drop the first, conflicting observation in ZIP code B.
3. For remaining duplicates by month, we keep the ZIP code that is geographically closer to the future modal ZIP code, that is, the one that has the smaller centroid-to-centroid geodesic distance to the ZIP code that appears most often in the list of ZIP codes after that month.

Finally, we drop return moves that happen suspiciously close – within one year – to earlier observations. That is, if an individual moves from ZIP code A to ZIP code B and relocates back to the exact same address in ZIP code A within one year, we drop this return move from B to A.

Ultimately, our analysis uses a location panel measured yearly, not monthly. Thus, if a person is associated with two or more ZIP codes in the same year, we keep the one closest to November (the month in which elections occur). If there are two ZIP codes in a given year,

one in October or November and one in December, we keep the former ZIP code for the current year and the December ZIP code for subsequent years, until the next move.

C.4.3 Creating ZIP code histories for children

Finally, we use the processed Infutor ZIP code histories for parents to infer children's residential histories. Since a child may have been matched to multiple parents, for each child, we choose one parent from whom to infer the address history using the following lexicographic rules: (1) the parent with the longest address history (time between the first and last Infutor address observations), (2) the parent who was most recently observed, (3) the female parent, and (4) the younger parent.[C13]

Here, we also create a family ID: we say that children are in the same family if they share the same "primary" parent (from whom we infer the address history), except if this would imply that more than 12 children are part of the same family.

[C13]We first standardize identifying information for each parent by choosing the modal value of name, date of birth, and gender. We prioritize information from TargetSmart over Infutor.