Childhood Cross-ethnic Exposure Predicts Political Behavior Seven Decades Later: Evidence from Linked Administrative Data

Jacob R. Brown[*], Ryan D. Enos[†], James Feigenbaum[‡], and Soumyajit Mazumder[§]

Abstract

Does contact across social groups influence sociopolitical behavior? This question is among the most studied in the social sciences with deep implications for the harmony of diverse societies. Yet, despite a voluminous body of scholarship, evidence around this question is limited to cross-sectional surveys that only measure the short-term consequences of contact, or to panel surveys with small samples covering short time periods. Using advances in machine-learning that enable large-scale accurate linkages across datasets, we examine the long-term determinants of sociopolitical behavior through an unprecedented individual-level analysis linking contemporary political records to the 1940 United States Census. These linked data allow us to measure the exact residential context of nearly every person in the United States in 1940 and connect this with the political behavior of those still alive over 70 years later. We find that, among white Americans, early-life exposure to Black neighbors predicts Democratic partisanship over 70 years later.

[*]Department of Government, Harvard University, jrbrown@g.harvard.edu

[†]Department of Government, Harvard University, renos@gov.harvard.edu

[‡]Department of Economics, Boston University and National Bureau of Economic Research, jamesf@bu.edu

[§]Department of Government, Harvard University, smazumder@g.harvard.edu

Introduction

The social and behavioral consequences of ethnic diversity are implicated in the long-term success of diverse societies and, consequently, are among the most important and longstanding topics across the social sciences. Scholars have argued that ethnic diversity leads to social inefficiencies, including discriminatory behavior (Alexander and Christia, 2011; Brewer, 1979; Enos, 2014; Sherif, 1961) and, in the aggregate, social and political instability (Alesina and Spolaore, 2005; Habyarimana et al., 2007; Putnam, 2007). Indeed, intergroup conflict may have been a crucial selective pressure in human evolution and is a nearly universal feature of human societies (Couzin et al., 2011; Lim, Metzler, and Bar-Yam, 2007), leading to a "liberal dilemma" (Enos, 2017) of an association between diversity and social inefficiency in an increasingly diversifying world.

As an antidote to this dilemma, scholars have long argued that interpersonal relationships across social groups can reduce prejudice and lead to more aggregate harmony (Allport, 1954; Pettigrew and Tropp, 2006; Ramos et al., 2019), especially when such contact occurs during adolescence. Such claims from psychology and other fields have been implicated in consequential jurisprudence (e.g, Clark and Clark (1939); Mussen (1950)) and public policies (Cook, 1985) and are among the most publicly influential social scientific theories (Paluck, Green, and Green, 2018).

Yet, despite the important implications and voluminous research, there are severe limitations to the evidence on the effect of inter-ethnic contact on long-term outcomes (Paluck and Green, 2009). To date, studies have relied almost entirely on experiments where the outcome is measured nearly immediately or a matter of days after the treatment, longitudinal surveys over short time periods, or cross-sectional surveys where the persistence of the association cannot be adequately measured (Hewstone et al., 2014; Paluck, Green, and Green, 2018; Wittlin et al., 2019), thus limiting the scientific and policy relevance of the findings and the long-term association between inter-ethnic contact and social harmony remains obscured.

To understand the longterm effects of intergroup contact, we undertake the first-large scale linking of full population, individual-level administrative data on adolescent inter-ethnic contact with records of later sociopolitical behavior. In doing so, we are able to construct a dataset that captures a number of early life experiences, including exact residential context, of nearly every child living in the United States in 1940 and, for those still living and registered to vote, to match this context to political behavior over six and, in many cases, seven decades later. This linking yields a dataset of over 650,000 individuals who were living in nearly every US county in 1940 and for whom we can observe political behavior much later in life. With these linkages, we are able to observe significant individual variation in racial diversity, socio-economic status, and other social indicators during childhood and link this variation to later individual political behavior allowing us to test for the long-term relationship between political behavior and having a different race neighbor in 1940.

Using an empirical strategy of increasingly fine-grained geographic comparisons to account for sorting at small levels of geography, we show that white individuals who had a Black neighbor in 1940, compared to white individuals who did not, are more likely to be associated with racially-liberal politics, as indicated by their association with left-wing political parties even as late as 2017. This relationship persists even when comparing whites living in the same neighborhood but with different levels of cross-racial exposure. We are further able to stratify our sample by individuals with like-age neighbors, by residential history, and other criteria indicating that the likely mechanism for this association is the influence of exposure to a neighbor of a different race rather than transmission of attitudes through parents' political attitudes or other unobserved variables.

In the US context, psychologists have argued that partisanship and racial attitudes are conditioned early in life and tend to be stable throughout a person's lifespan (Campbell et al., 1960; Tarman and Sears, 2005), with attitudes on inter-ethnic tolerance more strongly predicted by early, rather than later, life environments (Miller and Sears, 1986). Cross-ethnic exposure, including early in life, has been shown to liberalize short-term sociopolitical attitudes (Brown and Hewstone, 2005; Enos, 2016; Green and Wong, 2009; Putnam, 2007), perhaps because contact reduces the salience of group-based categories (Miller and Brewer, 1986) or because positive experiences with individuals are generalized to the out-group (Hewstone and Brown, 1986). The long-term effect of these early-life experiences on racial liberalism and spillover into other political attitudes could be why racial attitudes and partisanship are so highly correlated in the US (Erikson, MacKuen, and Stimson, 2002) and other countries (Sidanius and Pratto, 2001), with members of left-wing parties consistently displaying less ethnic and racial prejudice than members of right-wing parties (Jost et al., 2003).

The long-term endurance of these effects has important implications: if racial exposure is associated with behavior over seven decades later, especially on a characteristic with as much overtime stability as partisanship, it may be that the influence of these early-life experiences are also present in intervening decades, when people are active in politics and the workforce, and that the influence may be present in a range of sociopolitical behaviors and attitudes (Erikson and Stoker, 2011). Partisanship can be characterized as a social identity (Green, Palmquist, and Schickler, 2004) that is closely tied to a person's self-image and, as such, is a powerful predictors of behavior (Fiske, 2000). Party membership has been shown to induce a range of behaviors, including the type of group-based bias that characterizes race-based social identities (Iyengar and Westwood, 2014). Indeed, partisanship in the US is among the strongest predictors of nearly every political attitude (Achen and Bartels, 2017) and many nonpolitical lifestyle choices (Carney et al., 2008; Chen and Rohla, 2018) and, thus, long-term influences on partisanship have sweeping implications.

Linking 1940 Census Data to Contemporary Voterfiles

Census Data

We draw data on early-life experiences from the 1940 US Decennial Census, which attempted to record every person living in the US in that year and likely reached around 99% of the population (Steckel, 1991), including non-citizens. Enumerators recorded information on name, age, gender, race, place of birth, years of education, labor earnings, employment status and many other characteristics. Individual census records are not available when initially recorded but are made available for public use in accordance with a statutory 72-year restriction to protect the privacy of respondents. The census data we draw on has been transcribed and organized into structured datasets through collaborative efforts of the Integrated Public Use Microdata Series (IPUMS) at the Minnesota Population Center (MPC) and Ancestry.com and FamilySearch.org. We access de-anonymized data through files deposited by MPC at the National Bureau for Economic Research. For 131,903,910 of the 132,164,569 individuals enumerated in the 1940 census, we observe first and last names, as well as ages and places of birth, which we use to match to contemporary voter data.

Voterfile data

In most US states, citizens must register to be able to vote. When doing so, they usually provide basic information, including their full name, address, and date of birth. In most, but not all, states, voters also declare a party affiliation when registering. Because we can directly observe party affiliation using these administrative data, we largely avoid problems of measurement error associated with the surveys common in this field of study (Zaller et al., 1992).

In three states, California, North Carolina, and Nebraska, citizens provide their place of birth when registering. This information facilitates very accurate matches and so we focus on people currently living in those states. This population, however, was living in nearly every county in the United States in 1940 (3088 of 3108 counties, 99.4%), providing almost blanket coverage of US locations during the time period. We measure contemporary partisanship at two different points in time by pooling data from California and North Carolina in 2005 and 2009, respectively, and again to these states and Nebraska in 2017. The 2005/2009 data comes directly from the state governments, while the 2017 data is aggregated by the commercial vendor L2. The inclusion of samples at two different points in times allows us to test the robustness of the results to different datasets, states, and across different points in time. Because attrition via death and other causes will change the composition of the sample between 2005 and 2017, measurement at different times points also allows us to understand the influence of attrition from mortality and other sources on our inferences.

Record Linkage

Many of the methods for record linkage rely on methods based on the matching of available variables, most commonly name, gender, race, and age. These methods are non-statistical in the sense that they do not rely on a probabilistic model. Recent advances in the record linkage methodology argues for the use of probabilistic algorithms such as Support Vector Machines (Goeken et al., 2011), regression (Feigenbaum, 2016), and Bayesian inference (Enamorado, Fifield, and Imai, 2017; Abramitzky, Mill, and Perez, 2018). The core advantage of the probabilistic model is that it allows the researcher to directly control the false-positive rate by tuning the acceptance threshold for match probabilities.

To construct our linked sample, we use the supervised machine learning procedure developed by Feigenbaum (2016).[1] The key strength of this algorithm is that it allows the researcher to take advantage of the value of human coders by generating a hand-linked

[1] Abramitzky et al. (2019) document that the current generation linking methods most commonly used by economic historians trace out a frontier, trading off false positives and false negatives in linking. The Feigenbaum (2016) method is on this frontier. Further, Abramitzky et al. (2019) highlight that substantive research conclusions tend to be robust to all commonly used linking methods.

training set. Census data from 1940 was recorded in cursive by enumerators and transcribed 72 years later, introducing many layers of measurement error. Humans are quite good at comparing lists of names and making judgments across imperfect links, implicitly weighing differences in first and last names, as well as differences in year of birth or implied age. The machine learning approach makes those implicit weights on various record link features explicit and then replicably applies those weights to link very large datasets at scale.

We restrict our searches to men since the common practice of surname changes at marriage during the 20$^{\text{th}}$ Century makes it quite difficult to link women accurately. We then further subset the data to examine possible matches who were similar on first name, last name, and age, blocking on state of birth by looking at all men with matching states of birth, born within 5 years of the record in the voterfile, and with a first and last name within 0.3 in Jaro-Winkler string distance. A research assistant then attempts to link a random sample of the data by hand. We estimated a statistical model to predict which records the human linked and which records the human did not link, using a number of constructed features or variables based on the first and last name strings and the years of birth. Finally, we tuned a pair of hyperparameters to convert probit scores into matching rule decisions; the hyperparameters govern how strong a given match is both absolutely and relative to the next best alternative match. Trained on a small subset of the data, we then applied the linking algorithm to the rest of the data, creating matches from the voterfiles to the Census (see Supporting Information)

Overall, our linkage method performed quite well. The match rates—the shares of voter records of adequate age that we are confident in linking to a record in 1940—for California, Nebraska, and North Carolina were 52, 46, and 65 percent, respectively, for a sample 672,318 individuals in 2005/2009 and 259,762 in 2017 (760,337 unique individuals) who were living across the US in 1940 (Figure 1) . Almost half (46.6%) were living in California, Nebraska, or North Carolina, the states in which we examine contemporary behavior, but the mid-

Century population flows in the US are also evident, with the 1940 population also living in major metropolitan areas in the industrial Midwest and Northeast, such as Chicago, Detroit and New York, from which they moved to California or North Carolina at a later time. We restrict the study to whites, yielding 618,712 subjects in 2005/2009 and 238,353 in 2017 (699,554 unique individuals).

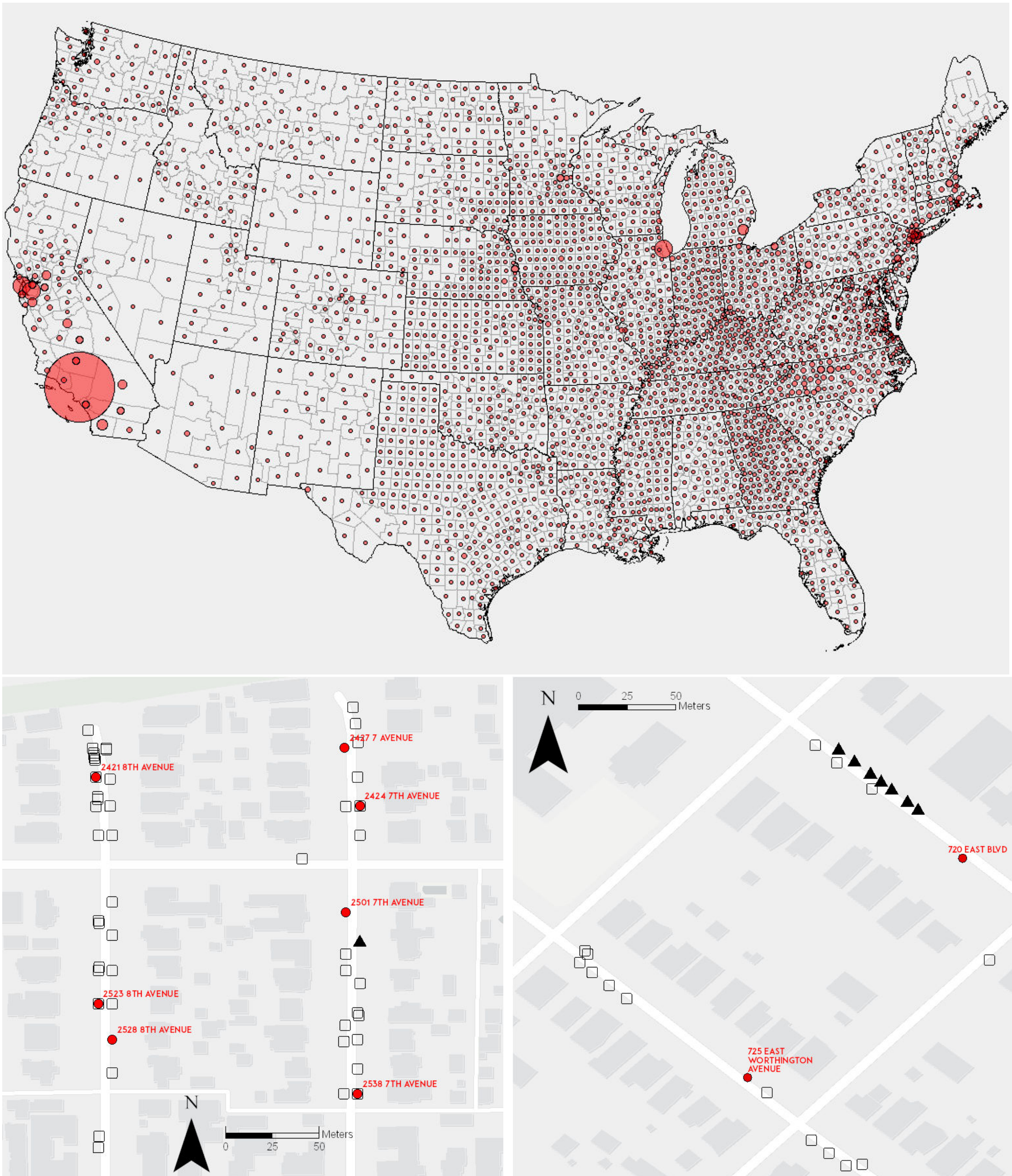
In 1940, 3.75% of the sample had an immediate Black neighbor. Because we can observe the exact location of individuals in 1940, we can distinguish between fine-grained differences in cross-ethnic exposure, even between people with different cross-ethnic exposure living in the same neighborhood (Figure 1). The proportion of registered Democrats in the linked sample, around 42%, is quite close to the proportion of individuals today who identify with Democrats from these two states and were born prior to 1940 (Figure S2).[2] A large majority of the sample were younger than ten years of age in 1940 and nearly the entire sample was younger than 20 years old (Figure S1). In the Supporting Information, we provide further descriptive statistics about our sample, documenting that our linked sample appears to be quite representative of the population of white children in 1940.

Measuring Cross-ethnic Exposure

To measure cross-ethnic exposure, we build on methods for estimating segregation developed by Logan and Parman (2017), exploiting the ordering of households on the 1940 U.S. Census enumeration sheets. When compiling the Census, enumerators in 1940 went door to door down the street, recording information for each house. As a result, adjacent households are very likely to appear immediately next to each other on census pages, and we are able to identify the characteristics of an individual's immediate next door neighbor, as well as the characteristics of their neighbors two doors down, three doors down, and so on. With this information, we construct measures of racial exposure for each individual in our dataset,

[2] Authors' calculations based on the the 2016 Cooperative Congressional Election Study.

Figure 1: Maps of Linked Sample and Neighbors in 1940



Top panel are the counties represented in 1940 in the linked sample with dots scaled by population represented in the contemporary sample. Bottom left panel is a neighborhood in Central Los Angeles in 1940. Bottom right panel is a neighborhood in Charlotte, NC in 1940. In these figures, red squares represent the white households found in the linked sample, with addresses labeled. Black triangles represent Black households and squares represent all other households. On the left, the subject at 2501 7th Avenue has a Black neighbor, while subjects on 8th Avenue do not. On the right, the subject on East Boulevard has several Black neighbors, while the subject on Worthington Avenue has none.

specifically whether each individual had a neighbor of the opposite race, and their relative proximity to said neighbor.

The key assumptions for our measures of racial exposure are that proximity in census enumeration is a proxy for residential proximity in actual geography, and that geographic proximity serves as a reasonable proxy for racial exposure. This assumption is supported by evidence that census enumerators were exhaustive, in that they visited every recorded household, and that they visited and recorded households per their geographic ordering (Agresti, 1980; Magnuson and King, 1995; Grigoryeva and Ruef, 2015). Households geographically next door to each other are very likely to be recorded next to each other on census manuscript pages. Thus, as we move up or down the census reel away from an individual in our dataset, we locate increasingly more distant neighbors.[3]

We construct an indicator variable representing the ten most proximate pairs of neighbors for each individual in our dataset. Each indicator incorporates two variables because for any given position (closest neighbor, second closest neighbor, ...) there are two households that are equally proximate to the individual, depending on whether you look up or down the Census page. As such, for $k \in [1, 10]$, each indicator equals 1 if at least 1 of the 2 households located k doors down from the individual are inhabited by a head of the household who is Black. By measuring cross-ethnic exposure on the individual-level using administrative records, we avoid having to use survey or aggregate data that can be subject to measurement error, aggregation error, and ecological fallacies (Enos, 2017).

We measure the relationship for each $k \in [1, 10]$ neighbors in 1940 and 2005/2009 or 2017 partisanship by regressing partisanship on indicators for a Black neighbor, as well as variables for the age and education (to measure social status not captured by race (Goldin and Katz, 2009) and to control for this possible confounding with race) of both the subject

[3]We also assume that moving, say, 3 spaces up the Census manuscript page is on average the same relative position as moving 3 spaces down the page. The neighbor who lives 3 doors up the street is not considered more proximate than the neighbor living 3 doors down the street, and vice versa.

and each neighbor. We estimate the following model:

$$Y_{i,g} = \alpha + \sum_{k=1}^{10} \beta_k D_{k,i} + \gamma X_i + \lambda_g + \epsilon_{i,g}$$

where $Y_{i,g}$ measures Democratic partisanship in 2005/2009 or 2017 for individual i in geography g , α is the intercept, $D_{k,i}$ is a matrix of neighbor characteristics for neighbor k positions away on the Census page, including whether the neighbor is Black and the head of the household has at least a High School education, X_i is a matrix of individual-level covariates that may affect both residential location and eventual partisanship: individual age, family size, whether their family had resided in the same residence 5 years previous, and the high school education, income, age, employment status, hours worked on average per week and per week for the head of the household, and whether the head of the household was in the labor force (employed or unemployed but looking for work).[4] λ_g is the geographic fixed effect, and $\epsilon_{i,g}$ is the error term.

Using fixed-effects, we examine this relationship at increasingly small geographic levels, beginning with state, county, and enumeration district fixed effects. Thus, all comparisons are between individuals living in the same geography allowing us to account for residential sorting at increasingly fine-grained levels. Enumeration districts are sub-county geographic units defined as the area for which a Census enumerator could complete a count of the area's population for a given census year. We then proceed to areas defined by groups of Census pages, which capture geographic proximity because adjacent houses were recorded on consecutive page lines by enumerators (Agresti, 1980; Grigoryeva and Ruef, 2015; Magnuson and King, 1995). We define geographies by 10 and 5 page groups, with the smallest group

[4] In the Supporting Information (Figure S6), we compare balance of these covariates across our sample with and without next door black neighbors. At smaller geographic fixed effects most covariates are balanced across groups, but we include these covariates as controls to account for potential confounding in all specifications. We also show in Tables S3–S6 that are results are similar in models not including these covariates.

representing a very small geography, approximately several city blocks (Table S2).[5] Thus, when comparing individuals at these smallest levels of geography, we have increased confidence that the white individuals with and without Black neighbors are equal on observable characteristics and the difference in contemporary behavior is a result of exposure to these Black neighbors.

Results

Whites with a next-door Black neighbor in 1940 are 1.5 to 4.6 percentage points more likely to be Democrats in 2005/2009, and 2.3 to 6.0 percentage points more likely in 2017 than whites without a Black next-door neighbor (Figure 2, results for neighbors $k > 1$ are in the Supporting Information).[6] The relationship is stable looking within increasingly small geographies and when accounting for other characteristics of the individuals in 1940, suggesting that the relationship is the result of exposure to a Black neighbor.

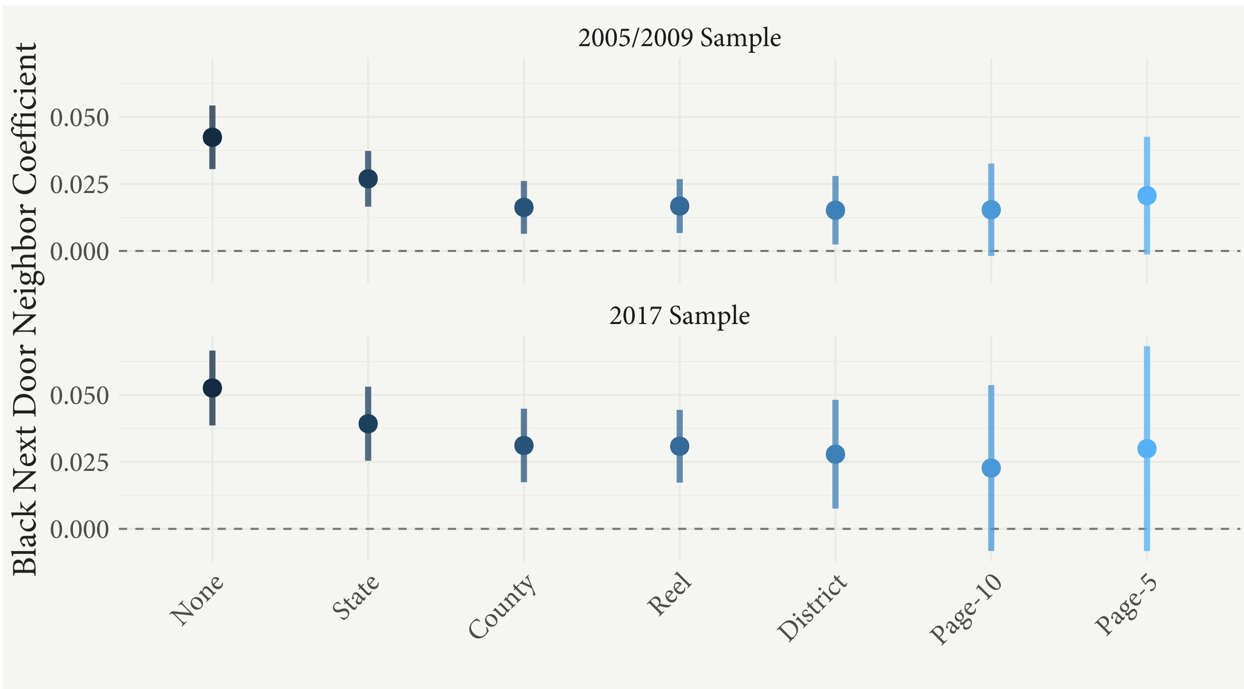
The Relationship Among Non-Movers, Non-Southerners, and Urban Dwellers

We examine several alternative explanations, other than cross-ethnic contact, for the association between childhood exposure and long-term partisanship. First the relationship could be driven by white people with racially progressive attitudes who selected to live next to Black people. Because we can compare individuals who lived within the same neighborhood, such geographic sorting is unlikely to drive the results, but to further investigate this possibility, we estimate specifications fitted to the subset our sample that has lived in their same residence for at least 5 years as recorded in the 1940 Census.

[5] In the Supporting Information (Figure S7), we present power analyses across fixed effect specifications.

[6] In the Supporting Information (Figure S10) we present the results with Republican partisanship as the outcome. A Black next door neighbor predicts decreased likelihood of being a Republican in the 2005/2009 and 2017 samples, and the results generally mirror the Democratic results in absolute magnitude and significance.

Figure 2: Black Next Door Neighbor and Democratic Partisanship Association



Points represent the coefficient of a Black next door neighbor on Democratic partisanship. Coefficients are from separate specifications with different geographic fixed effects, and are displayed in order of largest to smallest geographic comparison. Coefficient estimates are from models fitted to subset of whites in the data. Standard errors are clustered at the county level. Controls include individual age, family size, whether their family had resided in the same residence 5 years previous, as well as the high school education, income, age, employment status, labor force status, and hours worked on average per week and per week for the head of the household, and whether or not each neighbor has a head of the household with at least a high school education. The 2005/2009 sample consists of 512,558 individuals and the 2017 sample consists of 203,915 individuals. Differences between sample size in covariate models and full sample due to missingness across covariates.

Second, in the first half of the 20^{th} century, racially conservative whites tended to be Democrats and were clustered around areas with large Black populations (Key, 1949), and, thus, the results may reflect the persistence of this old-fashioned Democratic party identification rather than racial liberalism.

Lastly, because distances between neighbors will vary based on population density and this likely influences frequency of contact, we compare individuals in our samples living in urban areas in 1940 to those living in rural areas.[7]

For each of these three moderating variables, we estimate specifications on subsets defined by each variable (Figure 3). Across both samples, the predictive effect of a Black next door neighbor remains positive and significant in each of these sub-samples meaning we can rule out these variables as drivers of the relationship between early childhood cross-ethnic exposure and later life partisanship.

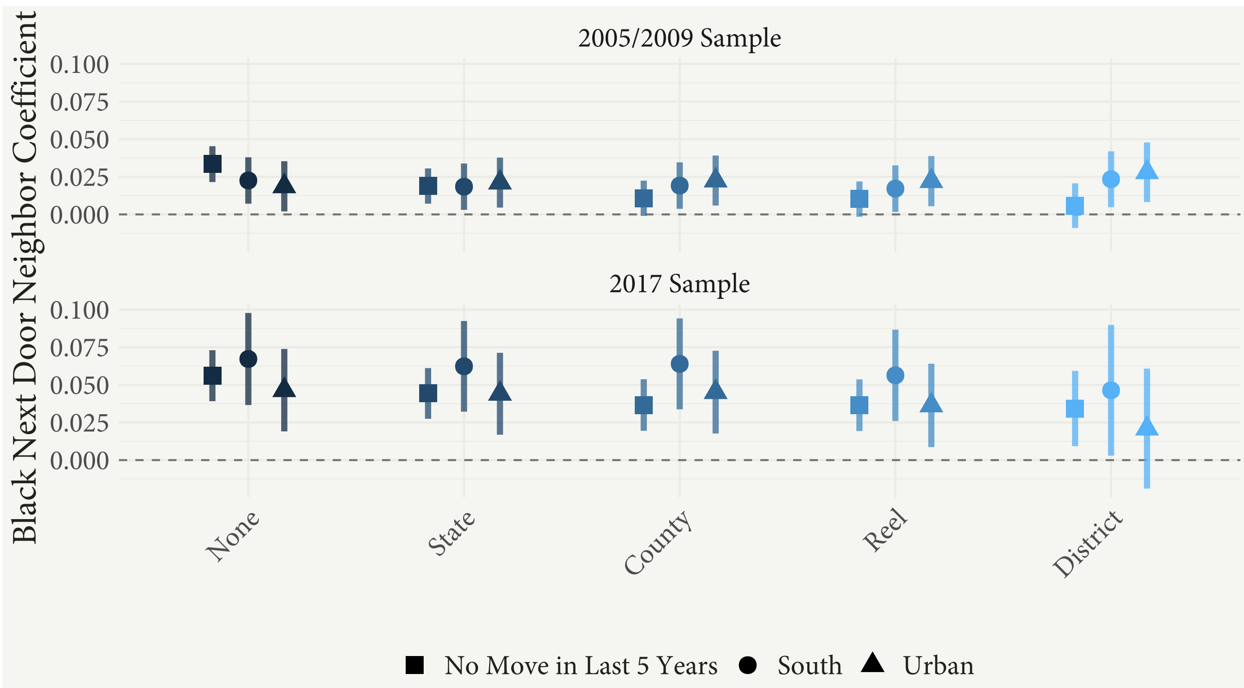
Additionally, in the Supporting Information, we also show (Figure S9) that the result is stable if we limit the sample to individuals who live next door to neighbors with children similar in age to the individual, consistent with the mechanism of cross-ethnic interaction producing durable effects. And we find that Black neighbors in 1940 are not predictive of living in neighborhoods with larger Black populations later in life (Figure S17). This indicates that the connection between early life context and future partisanship is more likely a product of the lasting impression left by cross-ethnic contact at a young age, rather than a product of greater long-term exposure to diversity.

Discussion

We have demonstrated that cross-ethnic contact early in life is associated with sociopolitical behavior over sixty and seventy years later and that this relationship can be observed even when comparing subjects who lived in the same neighborhood, but who had different levels

[7] The 1940 Census classifies a household as being in an urban area if it is located in a city or incorporated place of at least 2,500 inhabitants.

Figure 3: Black Next Door Neighbor Association on Sub-Samples



Points represent the coefficient of the Black next door neighbor on future Democratic partisanship. Coefficients are from separate specifications with different geographic fixed effects, and are displayed in order of largest to smallest geographic comparison. Coefficient estimates are from models fitted to subset of whites in the data who have either not moved in the last 5 years, lived outside the South in 1940, or lived in an urban area in 1940, respectively. Standard errors are clustered at the county level. Controls include individual age, family size, whether their family had resided in the same residence 5 years previous, as well as the high school education, income, age, employment status, labor force status, and hours worked on average per week and per week for the head of the household, and whether or not each neighbor has a head of the household with at least a high school education.

of exposure. Furthermore, the association is consistent across different regions of the country and different levels of density and for people who had and had not recently moved during childhood, suggesting that socialization early in life has a longterm effect on sociopolitical attitudes.

Most research on contemporary diversity and political behavior uses aggregate data, such as a measure of county-level demographics, to estimate ecological correlations, with the assumption that local diversity increases the probability of cross-ethnic contact. There are reasons to believe that a direct measure of exposure, such as the one we use here, is superior to or, at least, distinct from measures of contact taken from aggregate data. In fact, further analysis in the Supporting Information (Figures S15 and S16, and Tables S7 and S8) supports our finding that one's closest neighbors, rather than aggregate racial context, have a singular relationship with future partisanship. However, the relationship between direct cross-ethnic exposure, early-life aggregate diversity, and future political behavior deserves future investigation, especially since recent research has demonstrated a relationship between percent African American in high school and downstream racial attitudes for white Americans (Goldman and Hopkins, 2020). And, of course, future research examining the nature and valence of these cross-ethnic contacts will help with our understanding.

Remarkably, the relationship between early life experiences and partisanship persists over this long period, despite the many intervening life-experiences during this period—including the population of study being one of the more economically and geographically mobile populations in American history and with a large portion of the men experiencing the social disruption of military service. This persistence suggests a dominating socializing force of early life experiences. The longterm persistence of this relationship suggested that, despite the short-term social inefficiencies associated with diversity, there may be long-term positive effects for social harmony.

References

Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James J Feigenbaum, and Santiago Pérez. 2019. Automated linking of historical data. Technical report National Bureau of Economic Research.

Abramitzky, Ran, Roy Mill, and Santiago Perez. 2018. "Linking Individuals across Historical Sources: A Fully Automated Approach.".

Achen, Christopher H, and Larry M Bartels. 2017. *Democracy for realists: Why elections do not produce responsive government*. Princeton University Press.

Agresti, Barbara F. 1980. "Measuring Residential Segregation: In Nineteenth-Century American Cities." *Sociological Methods & Research* 8(4): 389–399.

Alesina, Alberto, and Enrico Spolaore. 2005. *The size of nations*. Mit Press.

Alexander, Marcus, and Fotini Christia. 2011. "Context Modularity of Human Altruism." *Science* 334(6061): 1392–1394.

Allport, Gordon W. 1954. *The Nature of Prejudice*. Cambridge, MA: Addison-Wesley.

Brewer, Marilynn B. 1979. "In-group bias in the minimal intergroup situation: A cognitive-motivational analysis." *Psychological bulletin* 86(2): 307.

Brown, Rupert, and Miles Hewstone. 2005. "An integrative theory of intergroup contact." *Advances in experimental social psychology* 37(37): 255–343.

Campbell, Angus, Phillip E. Converse, Warren. E. Miller, and Donald. E. Stokes. 1960. *The American Voter*. Chicago: University of Chicago Press.

Carney, Dana R, John T Jost, Samuel D Gosling, and Jeff Potter. 2008. "The secret lives of liberals and conservatives: Personality profiles, interaction styles, and the things they leave behind." *Political Psychology* 29(6): 807–840.

Chen, M Keith, and Ryne Rohla. 2018. "The effect of partisanship and political advertising on close family ties." *Science* 360(6392): 1020–1024.

Clark, Kenneth B, and Mamie K Clark. 1939. "The development of consciousness of self and the emergence of racial identification in Negro preschool children." *The Journal of Social Psychology* 10(4): 591–599.

Cook, Stuart W. 1985. "Experimenting on social issues: The case of school desegregation." *American Psychologist* 40(4): 452.

Couzin, Iain D, Christos C Ioannou, Güven Demirel, Thilo Gross, Colin J Torney, Andrew Hartnett, Larissa Conradt, Simon A Levin, and Naomi E Leonard. 2011. "Uninformed individuals promote democratic consensus in animal groups." *science* 334(6062): 1578–1580.

Enamorado, Ted, Benjamin Fifield, and Kosuke Imai. 2017. "Using a Probabilistic Model to Assist Merging of Large-scale Administrative Records.".

Enos, Ryan D. 2014. "Causal effect of intergroup contact on exclusionary attitudes." *Proceedings of the National Academy of Sciences* 111(10): 3699–3704.

Enos, Ryan D. 2016. "What the Demolition of Public Housing Teaches Us About the Impact of Racial Threat on Political Behavior." *American Journal of Political Science* 60(1): 123–142.

Enos, Ryan D. 2017. *The Space Between Us: Social Geography and Politics.* New York: Cambridge University Press.

Erikson, Robert S, and Laura Stoker. 2011. "Caught in the Draft: The Effects of Vietnam Draft Lottery Status on Political Attitudes." *American Political Science Review* 105(02): 221–237.

Erikson, Robert S., Michael MacKuen, and James A. Stimson. 2002. *The Macro Polity.* Cambridge studies in political psychology and public opinion Cambridge: Cambridge University Press.

Feigenbaum, James J. 2016. "A Machine Learning Approach to Census Record Linkage.".

Fiske, Susan T. 2000. "Stereotyping, Prejudice, and Discrimination at the Seam Between the Centuries: Evolution, Culture, Mind, and Brain." *European Journal of Social Psychology* 30(3): 299–322.

Goeken, Ron, Lap Huynh, Thomas Lenius, and Rebecca Vick. 2011. "New Methods of Census Record Linking." *Historical Methods* 44: 7–14.

Goldin, Claudia Dale, and Lawrence F Katz. 2009. *The race between education and technology.* harvard university press.

Goldman, Seth, and Daniel J. Hopkins. 2020. "Past Place, Present Prejudice: The Impact of Adolescent Racial Context on White Racial Attitudes." *Journal of Politics* .

Green, Donald P, and Janelle S Wong. 2009. "Tolerance and the contact hypothesis: A field experiment." *The political psychology of democratic citizenship* pp. 1–23.

Green, Donald P., Bradley Palmquist, and Eric Schickler. 2004. *Partisan Hearts and Minds.* New Haven, CT: Yale University Press.

Grigoryeva, Angelina, and Martin Ruef. 2015. "The Historical Demography of Racial Segregation." *American Sociological Review* 80(4): 814–842.

Habyarimana, James, Macartan Humphreys, Daniel N Posner, and Jeremy M Weinstein. 2007. "Why does ethnic diversity undermine public goods provision?" *American Political Science Review* 101(4): 709–725.

Hewstone, Miles, and Rupert Brown. 1986. "Categorization effects on ingroup and outgroup perception." In *Contact and conflict in intergroup encounters*, ed. Miles Hewstone, and Rupert Brown. Oxford: Basil Blackwell.

Hewstone, Miles, Simon Lolliot, Hermann Swart, Elissa Myers, Alberto Voci, Ananthi Al Ramiah, and Ed Cairns. 2014. "Intergroup contact and intergroup conflict." *Peace and Conflict: Journal of Peace Psychology* 20(1): 39.

Iyengar, Shanto, and Sean J Westwood. 2014. "Fear and Loathing Across Party Lines: New Evidence on Group Polarization." *American Journal of Political Science* .

Jost, John T, Jack Glaser, Arie W Kruglanski, and Frank J Sulloway. 2003. "Political conservatism as motivated social cognition." *Psychological bulletin* 129(3): 339.

Key, V O. 1949. *Southern Politics in State and Nation*. University of Tennessee Press.

Lim, May, Richard Metzler, and Yaneer Bar-Yam. 2007. "Global pattern formation and ethnic/cultural violence." *Science* 317(5844): 1540–1544.

Logan, Trevon D., and John M. Parman. 2017. "The National Rise in Residential Segregation." *The Journal of Economic History* 77(1): 127–170.

Magnuson, Diana L., and Miriam L. King. 1995. "Comparability of the Public Use Microdata Samples: Enumeration Procedures." *Historical Methods: A Journal of Quantitative and Interdisciplinary History* 28(1): 27–32.

Miller, Norman, and Marilynn B Brewer. 1986. "Categorization effects on ingroup and outgroup perception." In *Prejudice, discrimination, and racism*, ed. John F Dovidio, and Samuel L Gaertner. San Diego, CA: Academic Press.

Miller, Steven D, and David O Sears. 1986. "Stability and change in social tolerance: A test of the persistence hypothesis." *American Journal of Political Science* pp. 214–236.

Mussen, Paul H. 1950. "Some personality and social factors related to changes in children's attitudes toward Negroes." *The Journal of Abnormal and Social Psychology* 45(3): 423.

Paluck, Elizabeth Levy, and Donald P Green. 2009. "Prejudice reduction: What works? A review and assessment of research and practice." *Annual review of psychology* 60: 339–367.

Paluck, Elizabeth Levy, Seth A Green, and Donald P Green. 2018. "The contact hypothesis re-evaluated." *Behavioural Public Policy* pp. 1–30.

Pettigrew, Thomas F, and Linda R Tropp. 2006. "A meta-analytic test of intergroup contact theory." *Journal of personality and social psychology* 90(5): 751.

Putnam, Robert D. 2007. "E pluribus unum: Diversity and community in the twenty-first century the 2006 Johan Skytte Prize Lecture." *Scandinavian political studies* 30(2): 137–174.

Ramos, Miguel R., Matthew R. Bennett, Douglas S. Massey, and Miles Hewstone. 2019. "Humans adapt to social diversity over time." *Proceedings of the National Academy of Sciences* 116(25): 12244–12249.

Sherif, Muzafer. 1961. *Intergroup conflict and cooperation: The Robbers Cave experiment.* Vol. 10 University Book Exchange Norman, OK.

Sidanius, Jim, and Felicia Pratto. 2001. *Social dominance: An intergroup theory of social hierarchy and oppression.* Cambridge University Press.

Steckel, Richard H. 1991. "The quality of census data for historical inquiry: A research agenda." *Social Science History* 15(4): 579–599.

Tarman, Christopher, and David O. Sears. 2005. "The Conceptualization and Measurement of Symbolic Racism." *The Journal of Politics* 67(03): 731–761.

Wittlin, Natalie M, John F Dovidio, Sara E Burke, Julia M Przedworski, Jeph Herrin, Liselotte Dyrbye, Ivuoma N Onyeador, Sean M Phelan, and Michelle van Ryn. 2019. "Contact and role modeling predict bias against lesbian and gay individuals among early-career physicians: A longitudinal study." *Social Science & Medicine* 238: 112422.

Zaller, John R et al. 1992. *The nature and origins of mass opinion.* Cambridge university press.