Who Marries Whom?
The Role of Segregation by Race and Class*

Benjamin Goldman, Cornell University †
Jamie Gracie, Harvard University
Sonya R. Porter, U.S. Census Bureau

January 2025
Click here for most recent version

Abstract

Americans rarely marry outside their race or class group, a pattern with well-documented implications for inequality and intergenerational mobility. Limited exposure—or interactions with members of other groups—may partly explain these low intergroup marriage rates. We instrument for exposure using variation in childhood neighborhoods based on whether other race and class groups had more opposite-sex children of similar age. Exposure increases interclass (high- and low-parent-income) marriage but has no detectable effect on interracial (White and Black) marriage. A spatial marriage market model predicts that residential segregation—one of many forms of exposure—accounts for more than one third of marital sorting by class but less than 5% by race.

*Any opinions and conclusions expressed herein are those of the authors and do not represent the views of the U.S. Census Bureau. The Census Bureau has ensured appropriate access and use of confidential data and has reviewed these results for disclosure avoidance protection (Project 7519874: CBDRB-FY23-CES014-009, CBDRB-FY23-0492). This research was funded by Opportunity Insights, the Stone Program at Harvard Kennedy School, the National Science Foundation (DGE-2140743), Harvard University, and Cornell University. We are deeply grateful to Raj Chetty, Claudia Goldin, Nathan Hendren, Larry Katz, and Jesse Shapiro for their detailed comments and invaluable advice. We also thank Sule Alan, Isaiah Andrews, Zach Bleemer, Steve Coate, Kaveh Danesh, Will Dobbie, Brad Foster, John Friedman, Ed Glaeser, Jeff Gortmaker, Philipp Kircher, Martin Koenen, João Loula, Zhuan Pei, Danielle Sandler, Ben Scuderi, Matthew Staiger, Elie Tamer, Winnie van Dijk, and Joseph Winkelmann for their insightful feedback and suggestions. Additionally, we extend our thanks to seminar participants at Cornell, Dartmouth, Georgetown, Harvard, Northwestern, the University of Chicago, the University of California Berkeley, the University of California San Diego, and Yale.

†Email: bgoldman@cornell.edu

I Introduction

Americans rarely marry outside their race or class groups. Among Black Americans born in the 1980s, only 2.1% are married to a White spouse. Likewise, only 3.1% of individuals from low-income families (bottom parental income quartile) are married to someone from a high-income family (top quartile). The division of families along race and class lines is central to many issues of longstanding societal concern, including intergenerational mobility, discrimination, and the transmission of values and practices from parents to children (Bisin and Verdier, 2001; Bobo, 2001; Chadwick and Solon, 2002; Fryer, 2007; Mare, 2000; Olivetti et al., 2024).

Why is marriage across race and class lines so rare? One possibility is that individuals have a *preference* for partners from their own race or class group.[1] Alternatively, people may be open to marrying across group lines, but a lack of *exposure* to other groups prevents such marriages from forming. If exposure is a key driver of marital homophily, policies that increase cross-group interactions—such as inclusionary zoning, school redistricting, and housing choice vouchers—could foster more intergroup marriages, even if these policies are not explicitly designed to influence marriage patterns. These downstream effects on intergroup marriage could, in turn, influence broader dynamics of inequality and intergenerational mobility.

In this paper, we estimate the effect of exposure on marital homophily, focusing specifically on the impact of residential segregation—one of many forms of exposure. Our analysis proceeds in three steps. First, we document overall levels of marital homophily by race and class, providing the first population-wide estimates of marital homophily by parental income (or "class") for recent cohorts in the U.S. We also present novel evidence on how distance shapes marriage outcomes, since the effect of reducing segregation depends on whether proximity influences marital choices. Second, we estimate, in partial equilibrium, whether increased exposure to opposite-sex members of other groups leads individuals to marry across race and class lines. We instrument for exposure using sex ratio variation in childhood neighborhoods, driven by whether other race and class groups had more opposite-sex children of similar age. Finally, we develop and estimate a spatial model of the marriage market to predict the effect of reducing segregation in general equilibrium, accounting for spillovers across neighborhoods. Our findings suggest that exposure plays a significant role in marital homophily by class, but has little effect on marital homophily by race.

We draw on administrative data that links the universe of federal tax records to decennial censuses. Our objective is to measure marital sorting based on parental income, which is a key determinant of the overall level of intergenerational mobility (Buckles et al., 2023; Chadwick and

[1] We use the term "preferences" to describe the propensity for within group marriage conditional on exposure. Preferences need not be innate. They can reflect preferences of family or community, and they can arise endogenously from life experiences and social norms that can, in principle, be affected by exposure (Alan et al., 2023; Allport, Clark and Pettigrew, 1954; Barlow et al., 2012; Corno, La Ferrara and Burns, 2022; Paolini, Harwood and Rubin, 2010).

Solon, 2002; Olivetti et al., 2024). We use the term "class" as shorthand for parental income and do not view this measure as encompassing a broader concept of social class. We refer to parents in the top and bottom quartiles of the household income distribution as "high-income" and "low-income," and use "White" and "Black" to denote non-Hispanic White and non-Hispanic Black individuals, respectively.[2] We focus on White-Black interracial marriage due to its low prevalence and the significant role that marriage rates and spousal income play in economic disparities between Black and White Americans (Binder et al., 2022; Chetty et al., 2020). Our baseline sample consists of individuals born between 1982 and 1989, measuring marital status at age 30, when 35% of individuals are married. We provide additional estimates for a subset of cohorts at later ages.

Rates of interclass and interracial marriage are low.[3] As noted above, only 3.1% of people from a low-income family have a spouse from a high-income family, and only 2.1% of Black individuals have a White spouse. We quantify marital homophily by comparing these intergroup marriage rates to two benchmarks. In the first, all individuals marry at the population marriage rate and select spouses randomly. In the second, spouses are randomly reassigned within the married population, keeping group-specific marriage rates fixed. Observed rates of interclass and interracial marriage are only one-third to one-half of these benchmarks, meaning there would be two to three times more intergroup marriages if relationships formed without regard to race and class.

The extent to which residential segregation suppresses intergroup marriage depends on how often people marry neighbors. Nearby residents shape marital choices by influencing social interactions through schools, workplaces, and recreation, as well as by determining physical proximity to potential partners. For instance, online dating technologies, now the primary way couples meet, often require users to specify a search radius, with many restricting searches to their own neighborhoods (Kirkham, 2019). Consistent with this, we find that 68% of married couples lived within 50 census tracts of each other five years before marriage, even though, on average, only 0.07% of the U.S. population resides within this radius.[4] Within this range, marriage probabilities decline sharply with distance; marrying someone from the same childhood census tract is ten times more likely than marrying someone from the 50th closest tract, typically just a few miles away.

Given that neighbors constitute an important segment of the marriage market, we next explore how likely individuals are to live near people from different racial and class backgrounds.

[2] Although race and class are related, they are far from perfectly correlated. For example, the group of individuals from low-income families is racially diverse: 38.3% are White, while 23.9% are Black.

[3] Measuring class-based homophily requires longitudinal data tracking individuals from childhood to adulthood (e.g., Althoff, Gray and Reichardt, 2024; Charles, Hurst and Killewald, 2013; Ermisch, Francesconi and Siedler, 2006; Olivetti et al., 2024), so much of the available evidence comes from 19th- and early 20th-century Census microdata or from other countries (e.g., Fagereng, Guiso and Pistaferri, 2022). In contrast, interracial marriage is easier to measure in recent data. Although White-Black marriage has increased in recent decades, it remains rare (Deal, 2024*b*; Fryer, 2007; Qian and Lichter, 2007, 2011).

[4] Census tracts are small geographic units with about 4,000 residents; there were 74,134 tracts in the U.S. in 2010.

The U.S. is known to be segregated by race and class (e.g., Hall, Iceland and Yi, 2019; Logan et al., 2020; Massey and Denton, 1988; Reardon and Bischoff, 2011). We measure segregation, or neighborhood exposure to other groups, based on the composition of the nearest 50 census tracts. The neighborhoods where individuals live as young adults (ages 18 to 27) are highly segregated by race and class. On average, Americans live near about twice as many people from their own race or class group as they do from other groups.

People tend to marry neighbors, and neighbors are typically from the same race and class group. This naturally raises the question: would intergroup marriage be more common if neighborhoods were more diverse? To explore this, we begin by documenting the association between intergroup marriage rates and exposure to other groups in young adulthood. For individuals from low-income families, a 10-percentage-point (pp) increase in the share of neighbors from high-income families is associated with a 2.2pp (71.0%) increase in the probability of marrying a high-income spouse. Among Black individuals, a 10pp increase in the share of White neighbors corresponds to a 0.9pp (42.9%) increase in the probability of having a White spouse. Thus, intergroup marriage is substantially more common in integrated neighborhoods than in segregated ones, though this relationship may reflect either a causal effect of exposure or selection in the types of individuals who choose to live in more integrated neighborhoods.

In the second part of the paper, we estimate the causal effect of exposure in young adulthood by constructing an instrument based on race- and class-specific sex ratios in childhood neighborhoods.[5] We define the instrument based on childhood neighborhoods to avoid relying on variation in neighborhood composition that arises endogenously from individuals' choices about where to live as adults. The intuition behind this instrument can be illustrated with a simple hypothetical: suppose a male child is born in a place where, by chance, nearby high-income families have more female than male children in the years around his birth. Because people often live near where they grew up, as an adult he will likely enter a marriage market with a higher proportion of females from high-income families. Hence, random variation in whether certain groups had more male or female children in a given year can generate lasting shocks to marriage market composition.

The instrument generates differences in market "tightness," defined as the difference between the proportions of opposite- and same-sex neighbors from a given race or class group. For instance, transitioning from a market where 10% of both male and female neighbors are Black to one where 5% of male neighbors and 15% of female neighbors are Black represents a 10pp increase in Black

[5]The first wave of studies using sex ratios relied on observational variation in adult neighborhoods (Cox, 1940; Easterlin, 1961; Groves and Ogburn, 1928; South and Lloyd, 1992). More recent research has shifted toward leveraging policy changes (Goni, 2022) or quasi-experimental designs that influence sex ratios (Abramitzky, Delavande and Vasconcelos, 2011; Angrist, 2002; Bazzi et al., 2023; Brainerd, 2017; Charles and Luoh, 2010; Hoxby, 2000; Kirkebøen, Leuven and Mogstad, 2021). Our approach—exploiting random variation in sex ratios at birth to ensure exogenous variation—is unique, with the exception of a concurrent study by Zaremba (2024), which also uses birth sex ratio variation to examine how marriage market bargaining power affects maternal health outcomes for Black women.

market tightness for men. We then estimate how the higher availability of Black females increases the probability of marrying a Black partner, capturing the partial-equilibrium effect of exposure to opposite-sex members of other groups. This approach relies on three empirical facts: 1) 99% of legal marriages in the U.S. are opposite-sex marriages (Kreider et al., 2023); 2) it is common to marry a neighbor close in age; and 3) people often live near where they grew up as young adults (Foster, 2017; Sprung-Keyser, Hendren and Porter, 2022).[6] The first fact motivates our definition of tightness as opposite sex share relative to own sex share. The second fact defines cells of nearby neighborhoods and birth cohorts that are important for marriage prospects but small enough to generate sex ratio variation. The third ensures a first stage by linking the demographic composition of child and adult neighborhoods.

Our instrumental variable (IV) estimates show that increased market tightness—the difference in the fraction of opposite- versus same-sex neighbors from other class groups between ages 18 and 27—substantially increases interclass marriage. For individuals from low-income families, a 10pp increase in market tightness among individuals from high-income families raises the probability of an interclass marriage by 1.2pp (38.7%). This change amounts to more than one-third of the gap between the observed intergroup marriage rate (3.1%) and the rate expected if partners were chosen without regard to class (6.3%). The strong response to market tightness across class lines suggests that exposure is an important factor in low interclass marriage rates.

The IV estimates for race reveal a notably different result. For White individuals, a 10pp increase in market tightness among Black individuals results in only a 0.0002pp (0.04%) increase in the probability of interracial marriage. For Black individuals, a 10pp increase in market tightness among White individuals leads to just a 0.05pp (2.4%) increase in the probability of interracial marriage. Increased market tightness does not lead to higher rates of interracial marriage, suggesting that exposure is unlikely to be an important factor in the low rates of interracial marriage.

We use multiple strategies to assess the validity of our IV estimates. First, the distribution of our instrument across neighborhoods matches the distribution expected under the assumption that sex is an IID Bernoulli random variable. This pattern would not hold if parents sorted into neighborhoods based on their child's sex, supporting the assumption that the instrument's variation arises solely from random sex assignment at birth. Additionally, our estimates are robust to controlling for census tract fixed effects, which capture variation across cohorts within the same tract, and family fixed effects, which capture variation across cohorts between siblings.

In the final part of the paper, we integrate our IV estimates with a spatial model of the marriage market to predict how marriage outcomes respond to changes in segregation, addressing two challenges. First, we translate our IV estimates—based on variation in market tightness, the share

[6]The Supreme Court ruling legalizing same-sex marriage, Obergefell v. Hodges, occurred in 2015, during our sample period. Consequently, marriages recorded in the tax data during this time are almost exclusively opposite-sex.

of a given group that is male versus female—into estimates of segregation, the overall share of a given group. Second, we extend our estimates of the effect of market tightness on marriage outcomes within a single neighborhood (partial equilibrium) to predict changes in overall marriage outcomes (general equilibrium). Because neighborhoods are not isolated marriage markets, changes in marriage outcomes in one neighborhood can spill over to other areas.

We start from the Choo and Siow (2006) transferable utility matching model and extend it in two ways. First, we introduce neighborhoods into the type space and add a distance cost to the utility function, allowing individuals to prefer spouses based on proximity. Second, we relax the logit assumption by allowing individuals to sort into neighborhoods based on unobserved characteristics.[7] Both intuition and evidence—such as the difference between OLS and IV estimates of exposure effects on interracial marriage—underscore the importance of accounting for correlated unobserved traits. Including these traits enables the model to replicate the strong observed relationship between intergroup marriage and segregation, even without a causal exposure effect. Simulations show that our instrument effectively disentangles selection from the causal effect of exposure, offering a framework applicable to other matching problems with correlated unobserved traits. We estimate the model using the simulated method of moments (McFadden, 1989), identifying preferences for spouse characteristics and proximity that best align simulated moments with the observed national marriage outcomes, the association between marriage outcomes and neighborhood exposure, and the IV effects of market tightness.

Small reductions in segregation, targeting the most segregated neighborhoods, can significantly impact both interclass and interracial marriage. These moves resemble those studied in programs like Moving to Opportunity, HOPE VI, and the Gautreaux Project (Bergman et al., 2024; Chyn, Collinson and Sandler, 2023; Ludwig et al., 2013; Popkin et al., 2004). We find that a 10pp increase in cross-class exposure from such policies raises the probability of marrying across class lines by 3pp. A comparable increase in cross-race exposure increases interracial marriage by 1pp, closely aligning with quasi-experimental estimates from the Gautreaux Project (Chyn, Collinson and Sandler, 2023). Having validated the model against recent quasi-experimental findings, we next use it to estimate the total effect of segregation by simulating a "no segregation" counterfactual where individuals are evenly distributed by race and class across neighborhoods.

Residential segregation—one of many forms of exposure—plays a substantial role in marital homophily by class but a smaller role by race. In our no-segregation counterfactual, interclass marriage rates rise from 3.1% to 4.6%, closing 35.2% of the gap between observed interclass marriage rates and the no-homophily benchmark, where all class groups marry at the same rates and choose

[7]This approach builds on recent advances in modeling unobserved heterogeneity (Ciscato, 2024; Galichon and Salanié, 2022; Graham, 2013; Gualdani and Sinha, 2023; Fox, 2010; Sinha, 2018) by incorporating a second unobservable component, with identification supported by an instrument unrelated to the unobservable type.

spouses randomly. Meanwhile, rates of White-Black interracial marriage rise from 2.1% to 3.0%, closing only 4.3% of the gap between observed interracial marriage rates and the no-homophily benchmark. These results suggest that lack of exposure across class lines partly explains low interclass marriage rates, making policies that promote cross-class interaction a promising tool to increase interclass marriage rates. On the other hand, residential segregation alone cannot explain persistently low interracial marriage rates, suggesting that other forms of exposure or deeper insights into preferences are needed to fully understand these patterns.

Our model helps reconcile findings in prior work that may initially appear contradictory. In addition to the Gautreaux Project, small increases in cross-race exposure have been shown to increase rates of interracial marriage in various contexts (e.g., Binder et al., 2022; Corno, La Ferrara and Burns, 2022; Gordon and Reber, 2018; Merlino, Steinhardt and Wren-Lewis, 2019). Conversely, other studies have documented strong preferences against interracial marriage or found that larger changes in segregation fail to significantly increase interracial marriage rates (Deal, 2024*a*; Fisman et al., 2008; Hitsch, Hortaçsu and Ariely, 2010; Wong, 2003). Our parsimonious model captures both phenomena and aligns with empirical findings on small versus large changes in segregation. By incorporating heterogeneous preferences, it shows how small increases in cross-race exposure can meaningfully raise interracial marriage rates, even with strong *average* preferences against interracial marriage, while larger changes in exposure have more modest effects.

Our study builds on and contributes to an extensive literature in economics on marital homophily, or "assortative mating" (Becker, 1973). We construct novel measures of marital homophily based on parental income and the role of distance in partner choice. Prior research has primarily focused on the *consequences* of marital homophily for economic inequality, highlighting marriage's significant role in shaping race- and class-based economic disparities.[8] In contrast, we investigate the *causes* of marital homophily, focusing on whether limited exposure to other groups constrains intergroup marriage. This approach builds on the literature examining the role of "meeting opportunities" in marriage outcomes (e.g., Belot and Francesconi, 2013; Kalmijn and Flap, 2001), with notable recent advancements by Ciscato (2024). Our results are unique in their use of population-wide data to estimate the effects of exogenous changes in exposure. The spatial model and quasi-experimental variation used here could also be applied to study the role of exposure in homophily for traits beyond race and class.

This study also contributes to the literature on the determinants of inequality and intergenerational mobility, specifically focusing on the role of social interactions and segregation. A growing body of work documents the extent of intergroup social interactions and their role in shaping in-

[8]This literature includes studies on the impact of marital homophily within (Ciscato and Weber, 2020; Clark and Cummins, 2022; Eika, Mogstad and Zafar, 2019; Schwartz, 2010) and across generations (Bailey and Lin, 2024; Binder et al., 2022; Buckles et al., 2023; Chadwick and Solon, 2002; Keller and Shiue, 2023; Ermisch, Francesconi and Siedler, 2006; Eriksson et al., 2023; Espín-Sánchez, Ferrie and Vickers, 2023; Mare, 2000; Olivetti et al., 2024).

equality and economic opportunity (Chetty et al., 2022*a*; Fischer and Mattson, 2009; Putnam, 2016; Smith, McPherson and Smith-Lovin, 2014). We contribute by examining marriage as a particular form of intergroup interaction—a relationship of unique importance due to its impact on individual well-being and future generations. We examine how segregation influences intergroup marriage, identifying an additional channel through which segregation may contribute to inequality (Ananat, 2011; Andrews et al., 2017; Cutler and Glaeser, 1997; Diamond, 2016; Derenoncourt, 2022; Fernández and Rogerson, 2001; Kremer, 1997; Massey and Denton, 1993).

The paper is organized as follows: Section II describes the data; Section III presents descriptive evidence on marital homophily, neighborhood segregation, and the role of distance in partner choice; and Section IV estimates the causal effect of exposure to opposite-sex members of other groups on intergroup marriage. In Section V, we build a spatial model of the marriage market to predict the effect of reducing segregation in general equilibrium. Section VI concludes.

II Data

We rely on three primary sources: the 2000 and 2010 decennial censuses, federal tax returns from 1979, 1984, 1989, 1994, 1995, 1998-2019, and the American Community Survey (ACS) from 2005-2019. The datasets are linked using a unique person identifier called a Protected Identification Key (PIK). PIKs are assigned using information including Social Security Numbers (SSNs), names, addresses and dates of birth. The record linkage process is described in detail in Wagner and Lane (2014). We limit our analysis to individuals who are assigned a PIK. The data are stripped of personally identifiable information. We share our analysis sample with Chetty et al. (2024), which extends the data in Chetty et al. (2020).

II.A Sample Definitions

Our sample starts with the set of individuals who were ever claimed as a dependent on a tax return. We match the individual and their claimer(s) to the 2020 Numident file to assign a year of birth. The claimer is a potential parent if they were between the ages 15 and 50 in the year the child was born. The first potential parent(s) who claim the individual are assigned as the invariant parent(s), regardless of any subsequent changes in marital status or dependent claiming.[9] This procedure relies on matching to the Numident file, a dataset covering all people with SSNs, and will therefore exclude any individuals who are unauthorized immigrants to the U.S. or who were

[9] A person who claims a dependent on a tax return does not have to be that individual's biological parent, but they do have to be financially supporting the individual. We restrict those ages to 15-50 to avoid links to grandparents or siblings. Chetty et al. (2020) show that for children claimed in the year 2000, 93% live with those same parents in the 2000 decennial census.

claimed by unauthorized immigrants.

We measure parental household income during childhood in our sample of linked children and parents. Following Chetty et al. (2024), we construct mean parent income between ages 13 and 17 based on the parent(s)' tax returns in those years. We use Adjusted Gross Income in the years where parents file a tax return and impute household income using W-2 income for non-filers. The W-2 data begin in 2005 and we code non-filers as having zero income if the W-2 data are unavailable.[10] We restrict to families with positive income since having negative income is typically due to capital losses, which is a sign that the family is wealthy (Chetty et al., 2014). The remaining individuals form our parent and child linked data.

We limit our parent- and child-linked data to individuals born between 1978 and 1999. Claiming rates begin to drop at age 16 (Chetty et al., 2014)—the 1978 birth cohort turns 16 in 1994, which is our first year of dependent claiming data. Our primary analysis sample consists of individuals born between 1982 and 1989. We use parental income data to measure the class of both individuals and their spouse. Starting with the 1982 cohort allows us to reliably measure spouse class, as 86% of married individuals have a spouse who is less than five years older than themselves, ensuring that most spouses of the 1982–1989 cohorts are included in our linked analysis file. Our baseline estimates use marriage outcomes at age 30, which requires us to end our sample with the 1989 birth cohort since our last year of tax data is 2019.

II.B Variable Definitions

We measure all monetary variables in 2015 dollars, adjusting for inflation using the consumer price index (CPI-U).

Class. Our measure of class is based on parental income during childhood. We define parental income using total pre-tax household income, which we label as *household income* for simplicity. This measure excludes income from cohabiting partners or other household members, apart from the primary tax filer's spouse. In years when a parent files a tax return, household income is defined as the sum of Adjusted Gross Income, Social Security payments, and tax-exempt interest payments reported on their 1040 tax return. For non-filers, we define household income as W-2 income when available. If no W-2 data is available, household income for non-filers is set to zero. We calculate parental income as the mean household income over the five years when the child is between the ages of 13 and 17. We construct percentile ranks by ranking parents relative to other parents with children in the same birth cohort. For much of our analysis, we divide the sample into quartiles and refer to parents in the bottom and top quartiles of the parental income distribution as "low-income" and "high-income" parents, respectively.

[10] Chetty et al. (2020) show that the median income reported in the ACS for non-filers is $5,000.

Spouse Class. For individuals who are married, spouse class is defined analogously to own class using the same measure of parent income during childhood. Spouse class is missing if the spouse cannot be linked to parents. This group includes spouses born prior to 1978 and those who did not spend their childhood in the U.S., as well as people who fail to meet any of the criteria used in the construction of the linked parent and child data described above. We are able to assign spouse class for 85% of married individuals in our primary analysis sample. When measuring marriage outcomes, we assign those with missing spouse class to an "out-of-sample" class category.

Race. We assign each individual to a race and ethnicity group using information from the U.S. Census Bureau and the ACS. We take the responses in priority order from the 2010 short form, the 2000 short form, and the 2005-2019 ACS surveys. Following Chetty et al. (2020), we focus on the following race and ethnicity groups: non-Hispanic White, non-Hispanic Black, non-Hispanic Asian, non-Hispanic American Indian and Alaskan Native (AIAN), Hispanic and non-Hispanic "other." The last group contains Native Hawaiian and Other Pacific Islanders, the "some other race" category, and people who list two or more race groups. These categorizations cover 95% of our core sample and the remainder are coded as having a missing race.

Spouse Race. We assign race and ethnicity to spouses analogously to own race. Because race is an attribute of the individual rather than of their parents, we are able to define race even for spouses that we cannot match to parents. We are able to assign spouse race for 95% of married individuals in our primary analysis sample. When measuring marriage outcomes, we assign those with missing spouse race to an "out-of-sample" race category.

Marriage. Our primary definition comes from tax records at age 30. A person who filed a tax return with the status "married filing jointly" or "married filing separately" is considered married. Anyone who filed a tax return with a different status or does not file a return in the year they are 30 is coded as single. We construct marriage outcomes for each individual in our primary analysis sample and associate married individuals with their spouse's PIK. In addition to our baseline measure, we are able to measure marriage outcomes at age 37 for the 1982 birth cohort, the first in our sample.

Cohabitation. Cohabitation is defined using the set of individuals who received the ACS in the calendar year in which they turn 30. The survey is at the household level and individuals are cohabitating with a partner if they are living with a spouse or un-married partner in the survey.

Childhood Locations. We measure childhood location between child ages 0 and 18 using data on parent locations. For years prior to 2003, we assign parent location from the 1040 form in years where a tax return is filed. Starting in 2003, we use addresses on W-2's and other information returns for non-filers. All addresses are geocoded and assigned to standard Census geographic units (e.g., block, tract, and county) by Census staff in a procedure described by Brummet et al. (2014).

In years where no address is assigned, we impute using the chronologically closest non-missing parental address, choosing randomly in the case of ties. Our primary analysis uses childhood location at age 18, which we define using parental location data for all children in our sample, regardless of whether the child is alive and living with their parents at age 18.

Adulthood Locations. After age 18, we define location using an individual's own location rather than their parent(s)' location. If the individual does not file, but an information return is available, that address is used. If not, the location is imputed using the chronologically closest address including the parental address data. At age 27, we are able to assign an address to 99.8% of the sample.

Income in Adulthood. We measure an individual's income using their total pre-tax income when they are 30 years old, which we measure at the household and individual level. If an individual files a tax return, we define household income as the sum of Adjusted Gross Income, social security payments and tax-exempt interest payments on the 1040 tax return. We define individual income as wage income reported on the W-2 plus self-employment and other non-wage income reported on 1040 tax returns. For non-filers, we define both individual and household income as wages from the W-2, or as 0 if no W-2 is filed. We assign percentile ranks relative to other individuals who appear in our child and parent linked data and who are born in the same year.

Sex and Birth Cohort. Sex and birth cohort are obtained from the 2020 Numident file.

Nearest Neighbor Census Tracts. For each census tract, we ordinally rank the nearest 50 census tracts using the geographic distance between census tract centroids.

II.C Summary Statistics

Tables I and II report summary statistics separately by class and race. There are 31.1 million individuals in our analysis sample, all of whom are assigned class by construction, and 95% of whom have non-missing race information. We define class based on parental income quartiles so that 25% of the sample is in each class group.

There are substantial differences in the socioeconomic status of our race and class groups. Individuals with parents in the bottom quartile have a median parent household income of $\$15,010$, compared to $\$136,400$ for individuals with parents in the top quartile. Only 39.2% of individuals with bottom-quartile parents had married parents in the year they were claimed versus 91.9% for individuals with parents in the top quartile. White individuals have median parental earnings of $\$72,280$ and 79.9% have married parents when they are claimed. This compares to $\$29,610$ and 29.1% for Black individuals.

Consistent with prior work, we find large disparities in adult household income by race and class (e.g., Chetty et al., 2020; Mazumder, 2005; Solon, 1992). Individuals with bottom-quartile

parents have a mean household income percentile rank of 37.5 at age 30 compared to 63.1 for those with top-quartile parents. The mean percentile rank for White individuals is 55.4 compared to 35.6 for Black individuals.

Disparities in household income by race and class are closely tied to differences in marriage rates and spousal incomes (Bayer and Charles, 2018; Binder et al., 2022; Chadwick and Solon, 2002). At age 30, 23.9% of individuals with bottom-quartile parents are married, compared to 43.7% for those with top-quartile parents. Among married individuals, those from families in the bottom quartile marry a spouse with an average individual earnings percentile of 52.3 versus 67.1 for those from top-quartile families. At age 30, 43.3% of White individuals and 11.7% of Black individuals are married. Among those who are married, spouses of White individuals have an average individual earnings percentile of 61.2 compared to 53.4 for spouses of Black individuals. White individuals and those from high-income families are more likely to be married and have higher earning spouses than those who come from low-income families or are Black.

Our empirical strategy relies on being able to measure location for a large fraction of the sample and individuals tending to live, as adults, near where they grew up. We are able to measure location for 86.9% of our sample at age 27 and 99.8% of the sample after imputing location with the nearest non-missing year with a valid address. At age 27, 67.2% of individuals in our primary sample still live in their childhood commuting zone, while 56.0% live in their childhood county, and 51.0% live within 50 census tracts of their childhood home. Individuals from high-income families move farthest from home, but even among this group, 61.8% still live in their childhood commuting zone, and 45.3% live within 50 census tracts of their childhood home.

III Descriptive Facts on Marriage, Homophily, and Neighborhoods

III.A Marital Homophily at the National Level

We begin by examining overall levels of marital homophily by class. Throughout, we focus on rates of intergroup marriage *unconditional* on marriage to capture the overall share of a group in an intergroup marriage and avoid conditioning on the selected sample of individuals who choose to marry. The decision to marry or not can partly reflect attitudes toward intergroup marriage, especially for minority groups, as most potential partners are outside their group. We show in Figure I Panel *A* that at age 30, 19.0% of individuals from high-income (top quartile) families have a spouse who also comes from a high-income family. In contrast, only 3.1% of individuals from low-income (bottom quartile) families have a spouse from a high-income family. Conversely, 5.0% of individuals from low-income families have a spouse from a low-income family, compared to just 3.1% of individuals from high-income families. Assessing whether these differences represent

significant levels of marital homophily requires further context.

To quantify the extent of homophily, we compare the observed rates of interclass marriage to two benchmarks where individuals select spouses without regard to class. These benchmarks, shown in Figure I Panel *A*, and summarized for all four class groups in Table III, provide a natural baseline for comparison. The first benchmark, labeled the "fully random benchmark," calculates marriage outcomes assuming that everyone marries at the same rate and selects spouses regardless of class background. The second, labeled the "random spouse benchmark," fixes the class-specific marriage rates but computes outcomes under the assumption that individuals choose spouses from the existing pool of married individuals without regard to class.

Under the fully random benchmark, 7.4% of individuals from either low- or high-income families would marry someone from the opposite class group. The random spouse benchmark predicts that 6.3% of individuals from these groups would intermarry. These results indicate a significant degree of marital homophily—interclass marriage would be 2–3 times more prevalent if individuals selected spouses without regard to class.

Figure I Panel *B* presents the corresponding statistics for interracial marriage, while Table IV provides detailed marriage outcomes and benchmarks for all groups. At age 30, only 2.1% of Black individuals have a White spouse, compared to 38.2% of White individuals. Group differences in marriage rates play a more substantial role for race than for class. Under the random spouse benchmark, 8.6% of Black individuals would have a White spouse, compared to 21.8% under the fully random benchmark. Even so, removing homophily while preserving racial differences in marriage rates would quadruple the rate of interracial marriage.

We next evaluate whether the low rates of interracial and interclass marriage are sensitive to our measure of marriage. In recent decades, Americans have increasingly delayed marriage and opted for cohabitation as an alternative (Cherlin, 2010; Kennedy and Bumpass, 2008; Lundberg, Pollak and Stearns, 2016; Manning and Cohen, 2012). Measures of homophily may therefore vary depending on the age at which marriage is measured and whether cohabiting partnerships are included, particularly if individuals who marry later or cohabit are more likely to be in interclass or interracial relationships.

To address these considerations, Tables A.1 and A.2 examine individuals surveyed in the ACS at age 30, including both married couples and cohabiting partners. Additionally, Tables A.3 and A.4 analyze marriage outcomes for the 1982 cohort at age 37, rather than age 30, using 2019 tax data. In both cases, we find higher overall rates of partnership or marriage compared to our primary estimates. However, the proportional differences between observed rates of interclass and interracial marriage and their benchmarks remain similar. For example, in Table A.2, the random spouse benchmark implies a share of Black individuals with a White spouse that is 4.6 times higher than the factual pairings, compared to 4.1 times higher in our primary estimates.

Rates of interclass and interracial marriage are low for both sexes. Tables A.5 and A.6 present marriage outcomes by class and race, disaggregated by sex. Among individuals from low-income families, 2.9% of men and 3.3% of women have a spouse from a high-income family. Among Black individuals, 2.6% of men and 1.7% of women have a White spouse. For the remainder of our analyses, we focus on the baseline estimates at age 30, pooling both sexes.

III.B Neighbors as Possible Marriage Partners

The impact of segregation on interclass and interracial marriage hinges on whether proximity influences marital choices. Neighborhoods play a critical role in shaping professional and social networks and determining school peers, thereby creating pathways through which nearby residents can influence who individuals date and eventually marry (Athey et al., 2021; Chetty et al., 2022*b*; McPherson, Smith-Lovin and Cook, 2001). For instance, online dating platforms, which have gained widespread popularity in recent years, often require users to restrict their search within a specified distance radius, establishing a direct relationship between geographic proximity and potential partners (Rosenfeld and Thomas, 2012; Rosenfeld, Thomas and Hausen, 2019). A recent survey finds that 35% of online daters are willing to date only within their neighborhood (Kirkham, 2019). In this section, we examine the extent to which people tend to marry someone who lived near them in the past.

We begin by defining a neighborhood as an individual's home census tract and the 50 nearest census tracts. Census tracts are designed to represent similar population sizes, so this definition spans a larger physical area in rural towns and a smaller area in cities. In 2010, there were 74,134 census tracts in the U.S. Figure A.1 illustrates what this definition looks like in areas with different population densities. For example, the nearest 50 census tracts to Harvard University exclude nearly all of Boston. In contrast, in a more rural area like Gilford, New Hampshire, the nearest 50 census tracts encompass much of central New Hampshire. On average, only 0.07% of the U.S. population lives within 50 census tracts of a given person. If people selected spouses without regard to distance or if marriage markets were significantly larger than 50 census tracts, only 0.07% of marriages would occur between individuals who lived within this radius of each other.

Even though the nearest 50 census tracts encompass a small share of the overall population, we find that a significant fraction of marriages occur between individuals who had lived within this radius prior to marriage. Figure II Panel *A* shows the fraction of couples who lived within 50 census tracts of each other by the year relative to their marriage. Five years prior to marriage, 68% of couples had already lived within 50 census tracts of each other at some point in their lives. This is partly explained by couples who had moved in together by that time. Excluding cohabiting couples, 58% of couples still had lived within 50 census tracts of each other—but not at the same

address—five years before their marriage.

Even after excluding couples at the same address, these rates can still be influenced by couples who choose to live near each other—though not at the same address—in the years leading up to their marriage. Ten years prior to marriage, when 90% of eventually married couples had not yet met (Rosenfeld, 2017), 45% had already lived within 50 census tracts of each other (excluding those at the same address), suggesting that many couples were living near each other before meeting. Table V presents these statistics across various geographic definitions, ages, and years relative to marriage. Remarkably, 23% of marriages are between two people who were born within 50 census tracts of each other. Across different ages, the share of couples living within 50 census tracts is similar to the share living within 10 miles of each other or within the same county.

Within 50 census tracts, individuals are significantly more likely to marry someone from the nearer tracts. Figure II Panel *B* illustrates marriage probabilities relative to someone in one's own census tract, based on neighborhood at age 18. Individuals are nearly 20 times more likely to marry someone from their own tract than someone who lives 50 tracts away, even though this distance is often just a few miles. The probability of marrying someone within one's own tract or the nearest nine tracts is equal to the probability of marrying someone located between 10 and 50 tracts away.

III.C Segregation by Race and Class in Childhood and Adulthood

Having established that people tend to marry someone from their neighborhood, we now turn to documenting neighborhood segregation by race and class to examine whether the pool of potential partners predominantly consists of individuals from similar racial and class backgrounds. We measure neighborhood composition within the nearest 50 census tracts across nine birth cohorts (the individual's own cohort and the four immediately older and younger cohorts). Within these 50 tracts, individuals are weighted based on their geographic and age distance. Details on the construction of these weights are provided in Appendix A.

For geographic distance, we weight individuals based on their relative likelihood of marrying someone at a given distance, as shown in Figure II Panel *B*. For example, a person living 25 tracts away is assigned a weight of 0.11, since the probability of marrying someone 25 tracts away is approximately $\frac{1}{0.11} \approx 9.1$ times lower than marrying someone from one's own tract. In practice, these weights are adjusted by terciles of population density, as marriages are more spatially concentrated in rural areas. The weights by population density are displayed in Figure A.2 Panel *A*.

For age, we use the weights shown in Figure A.2 Panel *B*, allowing the weights to vary by sex since men are more likely to marry someone younger, while women are more likely to marry someone older. The weight for the modal age difference is fixed at one. For men, the cohort one year younger receives a weight of 0.97, reflecting that men are nearly as likely to marry someone

one year younger as someone from their own cohort—the modal age difference for men. For women, the cohort one year younger has a weight of 0.52, indicating that women are approximately $\frac{1}{0.52} \approx 1.9$ times more likely to marry someone one year older than someone one year younger—the modal age difference for women.[11]

For the remainder of this paper, we use the term "neighbors" to refer to individuals living within 50 census tracts, weighted according to the geographic and age-based weights described above. For example, the "fraction of neighbors who are Black" refers to the proportion of individuals within the nearest 50 tracts and nine birth cohorts who are Black, weighted by their distance and relative age. In practice, this weighting concentrates most of the emphasis on individuals living within 10 tracts who are 0 to 3 years younger for men or 0 to 3 years older for women.

Figure III shows the distribution of the fraction of neighbors belonging to specific race or class groups across individuals in our sample. Since we measure marriage at age 30, we focus on neighborhoods during young adulthood, defined as the ten-year period from ages 18 to 27, with each age receiving equal weight. Panels *A* and *B* depict segregation by class, while Panels *C* and *D* show segregation by race.

During young adulthood, the median individual from a high-income family lives in a neighborhood where 34% of their neighbors also come from high-income families. In contrast, the median individual from a low-income family has a neighborhood composition in which only 17% of neighbors are from high-income families. This means the median person from a high-income family has twice as many high-income neighbors as the median person from a low-income family. Similarly, individuals from low-income families have 1.6 times more neighbors from low-income families than those from high-income families.

We observe similar levels of segregation by race. The median White person during young adulthood lives in a neighborhood where 78% of their neighbors are White, whereas the median Black person lives in a neighborhood where only 45% of neighbors are White. White individuals have 1.7 times more White neighbors than Black individuals. Conversely, the median Black person resides in an environment where 38.2% of their neighbors are Black—5.6 times higher than that for the median White person. Thus, during young adulthood, individuals tend to reside in environments where members of their own race and class are overrepresented, resulting in limited exposure to individuals from other groups.

III.D Association Between Segregation and Marriage Outcomes

Having established that people tend to marry neighbors and that neighborhoods are segregated, we now examine the relationship between segregation and marriage outcomes. Although aver-

[11] The modal age difference can differ between men and women because spouses may fall outside the primary sample.

age levels of segregation by race and class are high, Figure III highlights substantial variation in individuals' exposure to other racial and class groups. Before presenting our quasi-experimental estimates, we use this variation to explore whether individuals with greater exposure to neighbors from different race or class backgrounds are more likely to enter interracial or interclass marriages.

We begin by examining individuals from low-income families. In Figure IV Panel *A*, we present a binned scatter plot of interclass marriage versus exposure. The sample is divided into ventiles based on the fraction of neighbors between ages 18 and 27 who come from high-income families. For each ventile, we plot the average percentage of individuals with a spouse from a high-income family against the average fraction of neighbors from high-income families. The dashed line, labeled the "random sampling benchmark," has a slope equal to the marriage rate (with an intercept of zero) among individuals from low-income families. This represents what the data would look like if people selected spouses from their pool of neighbors regardless of class background.

The plot reveals a clear upward-sloping relationship. Among individuals from low-income families, a 10pp increase in the fraction of neighbors from high-income families leads to a 2.2pp (71.0%) increase in the probability of having a spouse from a high-income family. Only 0.3% of individuals in the bottom ventile of exposure have a spouse from a high-income family, suggesting that individuals who do not have neighbors from high-income families during young adulthood are unlikely to end up in an interclass marriage. The slope of 0.22 closely aligns with the random sampling benchmark of 0.24, suggesting that the association between marriage outcomes and exposure approximates what would be expected if spouses were chosen randomly from their pool of neighbors.

In Figure IV Panel *B*, we create a similar plot for individuals from high-income families, focusing on exposure to and marriage with individuals from low-income families. The relationship is again upward sloping, with a slope of 0.14, slightly flatter than the 0.22 slope observed for individuals from low-income families. Nevertheless, individuals from high-income families are significantly more likely to have a spouse from a low-income family if their neighborhood in young adulthood contains a higher fraction of neighbors from low-income families.

Individuals in more racially diverse neighborhoods are more likely to enter into interracial marriages, although the relationship between intergroup marriage and cross-group exposure is substantially flatter for race than for class. Figure IV Panels *C* and *D* present our binned scatter plots by race. For White individuals, the slope is 0.02, meaning a 10pp increase in the share of neighbors who are Black is associated with a 0.2pp (40.0%) increase in the probability of having a Black spouse. For Black individuals, the slope is 0.09, indicating a 0.9pp (42.9%) increase in the probability of having a White spouse when the fraction of neighbors who are White increases by 10pp. This relationship is primarily driven by Black individuals living in majority-White neigh-

borhoods, where the association is significantly stronger than in neighborhoods with fewer White residents.

To summarize these relationships, we combine the panels from Figure IV in Figure V. In Panel *A*, we combine samples of individuals from high- and low-income families to create a single binned scatter plot. Ventiles are constructed based on the fraction of neighbors from the "other" class group; for individuals from high-income families, the ventiles are based on the fraction of neighbors from low-income families, and vice versa. The series in Panel *A* plots the likelihood of having a spouse from the other class group against neighborhood exposure to that group. Panel *B* presents an analogous plot for race, using a combined sample of White and Black individuals. We use this combined sample to construct our quasi-experimental estimates of neighborhood exposure.

IV Quasi-Experimental Effects of Neighborhood Exposure

The association between segregation and marriage outcomes documented in Section III.D examines individuals who, as adults, choose to live in more versus less segregated neighborhoods. The observation that individuals are more likely to have an interracial or interclass marriage when they reside in neighborhoods with more members of other race or class groups could indicate a causal effect of exposure or selection. For instance, people might choose to live in neighborhoods with more members of a specific race or class group if they have traits that make them more appealing to members of that group or if they prefer to marry someone from that group. We address this challenge by developing an instrument for exposure to opposite-sex members of other race and class groups. Our instrument allows us to estimate the causal effect of exposure on interracial and interclass marriage.

IV.A Sex Ratios in Childhood Neighborhoods

We use variation in sex ratios in childhood neighborhoods as an instrument for exposure to opposite-sex members of other race and class groups. Our approach is unique in leveraging the random assignment of sex at birth, though it is related to other strategies that aim to isolate exogenous variation in sex ratios (Abramitzky, Delavande and Vasconcelos, 2011; Angrist, 2002; Bazzi et al., 2023; Brainerd, 2017; Charles and Luoh, 2010).[12]

As discussed in Section III, individuals living in nearby neighborhoods and of similar age constitute a significant component of the marriage market. Since people often reside near where they grew up, chance variation in sex ratios—such as Black families having more opposite-sex

[12]Zaremba (2024) also uses random sex assignment at birth but uses this variation as an instrument for intra-household bargaining to examine impacts on health outcomes. In contrast, we focus on the direct impact of this variation on marriage market outcomes.

than same-sex children in the years around an individual's birth—can have a lasting impact on the composition of the marriage market, increasing the number of potential Black partners within a specific time frame.

Our approach relies on three key facts. The first is that, as of 2020, 99% of legal marriages in the U.S. were opposite-sex marriages (Kreider et al., 2023). We measure marriages from 2012 to 2019, a period during which same-sex marriage was illegal in many states until the Obergefell v. Hodges Supreme Court decision in 2015. Because legal marriage—which we measure using tax return data—was predominantly opposite-sex during our sample period, we focus on sex ratio variation and define market tightness from the perspective of opposite-sex relationships.

The second fact, as shown in Section III.B, is that it is common to marry a nearby neighbor of similar age. As a result, neighborhood sex ratios in nearby birth cohorts influence the demographic composition of a key segment of the marriage market while generating sufficient variation for precise IV estimates.

The third fact is that individuals tend to live, as adults, near their childhood neighborhoods (Foster, 2017; Sprung-Keyser, Hendren and Porter, 2022). Consequently, the demographic composition of childhood neighborhoods shapes the demographic composition of adult marriage markets. Our instrument estimates a local average treatment effect (LATE) for individuals who stay near their childhood homes. This excludes, for instance, those who move far away for college and do not return. We argue that the "compliers" are an important group, as most Americans live near their childhood homes. At age 27, 67.2% of our sample remains in their last childhood commuting zone. This rises to 70.1% for those from low-income families but also comprises the majority (61.8%) of Americans from high-income families. Within commuting zones, more than half of Americans (51.0%) live within 50 census tracts of their childhood census tract.

By design, our approach does not change what we term market "thickness," or *overall* exposure to a given group, but instead alters the share of that group that is male versus female. We quantify this variation as market "tightness," defined as the difference in the fraction of opposite-sex versus same-sex neighbors from a given race or class group.[13] A White male born in a year and place where nearby Black families had more female than male children faces a "tighter" interracial marriage market. Segregation, by contrast, is about the overall fraction of neighbors from a given race or class group. In Section V, we incorporate our IV estimates of market tightness into a matching model to evaluate how changes in *overall* exposure—or market thickness—impact marriage outcomes. In this section, we focus on constructing the causal effect of market tightness on interclass and interracial marriage.

A helpful feature of our approach is that the data-generating process of the instrument—sex

[13] We also hold the *overall* sex ratio in one's neighborhood constant and focus on *group-specific* sex ratios. This ensures that our results exclude any effects of having more female classmates (Anelli and Peri, 2019; Hoxby, 2000).

assignment via a Bernoulli trial—is already known. This allows us to non-parametrically test whether the observed distribution of the instrument aligns with the distribution expected if the variation were driven entirely by random sex assignment at birth. Any deviation from Bernoulli variation could suggest that parents sort into neighborhoods based on the sex of their child, potentially introducing selection bias.

In Figure VI, we compare the observed distribution of the instrument to the one generated under random Bernoulli sex assignment. For each of our four key groups—individuals from low- and high-income families, and White and Black individuals—we draw from a binomial distribution at the census tract and group level. This process randomly assigns sex to each individual in the data and calculates the resulting sex ratio for each tract and group. The success probability of the binomial distribution is set to the national fraction of children who are female, and the simulation is repeated 1,000 times. We estimate each percentile of the distribution by averaging that percentile across simulations. The observed distribution of the instrument closely matches the distribution expected if the variation were solely driven by random sex assignment.[14] Consistent with this result, Figures A.3 and A.4 demonstrate that the instrument is uncorrelated across groups within a tract and unrelated to tract-by-birth-cohort covariates, such as mean parental income.

IV.B Market Tightness, Interracial and Interclass Marriage

Our baseline estimates define childhood neighborhoods using parent locations when the child is age 18. If parent location at age 18 is missing because parents did not receive tax or information returns in that year, we impute location using the closest age under 18 with available data. A childhood neighborhood is assigned even if the child has died before age 18 or is no longer living with their parents. Using this approach, we assign a childhood neighborhood to more than 99% of the sample.

We define market tightness, our endogenous treatment variable, as:

$$T_{i,g} = \frac{N^a_{i,-s,g}}{N^a_{i,-s}} - \frac{N^a_{i,s,g}}{N^a_{i,s}},$$

where $T_{i,g}$ represents the difference, for individual i and race or class group g, in the share of opposite-sex and own-sex neighbors who belong to group g in the individual's adult census tracts from ages 18 to 27. Neighbors include individuals within 50 census tracts and four birth cohorts. The term $N^a_{i,s,g}$ is the weighted count of neighbors of individual i between ages 18 and 27 who are of sex s and belong to group g. The term $N^a_{i,s}$ represents the weighted number of neighbors of

[14] Figure VI shows that the instrument even reflects the integer breaks caused by repeated integer values in some cells.

individual i who are of sex s across all groups. The subscript $-s$ indicates counts for the opposite sex, while s refers to counts for the same sex. The weights, which account for census tract distance and relative age, are described in detail in Section III.C.

Our instrument,

$$Z_{i,g} = \frac{N^c_{i,-s,g}}{N^c_{i,g}},$$

is the fraction of group g children in person i's childhood neighborhood that belong to the opposite sex. The variable $N^c_{i,-s,g}$ is the weighted count of the number of opposite-sex members of group g in person i's childhood neighborhood. The variable $N^c_{i,g}$ is the weighted count, including both sexes, of group g people in i's childhood neighborhood.

We use the following specification for our first stage and reduced form:

$$T_{i,g} = \beta^f Z_{i,g} + \delta^f_{o,l,s} + \Gamma^f X_{o,l,s,g} + \varepsilon^f_{i,g} \tag{1}$$

$$Y_{i,g} = \beta^r Z_{i,g} + \delta^r_{o,l,s} + \Gamma^r X_{o,l,s,g} + \varepsilon^r_{i,g}. \tag{2}$$

We instrument for market tightness, $T_{i,g}$, in the adulthood marriage market for group g with sex ratios among group g in person i's childhood neighborhood, $Z_{i,g}$. Our first stage coefficient is β^f. The o subscript indicates own group: in specifications focused on class, o indicates own-parent quartile, and for race, o indicates own race group. We include own group by childhood census tract (l) by sex fixed effects, $\delta_{o,l,s}$, ensuring that comparisons are only made within neighborhoods or groups. The within-childhood-tract variation compares two children of the same sex and group who faced different sex ratios among group g because they were born in different years.

If $T_{i,g}$ and $Z_{i,g}$ were measured contemporaneously (rather than one in adulthood and one in childhood), they would be related to each other by Bayes' rule. We show in Appendix B that, conditional on the marginal probabilities that appear in Bayes' rule, $T_{i,g}$ is a linear function of $Z_{i,g}$. Because of this relationship, we control for the marginal probability expressions in $X_{o,c,s,g}$, which includes own-group and sex fixed effects interacted with the overall share of (weighted) people who belong to group g and the overall share who belong to sex s in person i's childhood neighborhood.

The outcome variable, $Y_{i,g}$, is an indicator equal to one if person i is married to a member of group g at age 30, and otherwise equals zero. We also estimate effects on the overall likelihood of being married. Specifications by class use a sample of individuals from families in the bottom and top quartiles of parent household income, where g represents the "other" class group. For example, $Y_{i,g} = 1$ if an individual from a top-quartile family is married to a spouse from a bottom-quartile family (and vice versa for individuals from bottom-quartile families). For specifications by race,

$Y_{i,g} = 1$ if White individuals are married to Black spouses or if Black individuals are married to White spouses.

The distribution of childhood neighborhood sex ratio variation, based on our definition using the nearest 50 census tracts and four birth cohorts, is shown in Figure VII Panels *A* and *C*. In Panel *A*, we analyze a sample of individuals from families in the bottom and top parent income quartiles, focusing on sex ratios within the other class group. For individuals from the top quartile, the instrument represents the fraction of bottom-quartile individuals in their childhood neighborhood who are the opposite sex (and vice versa for individuals from bottom-quartile families). The histogram depicts the residual distribution of the instrument defined in equation 1. Nearly all of the variation falls within [-3, 3]. The average $N^c_{i,g}$ across individuals, where g is the opposite class group, is approximately 2,250. This implies that the instrument can shift the environment from one with 1,058 opposite-sex and 1,193 own-sex individuals to one with 1,193 opposite-sex and 1,058 own-sex individuals, representing a "shortage" versus a "surplus" of 135 people in the other class group.

In Panel *C*, the instrument has slightly more variation for race compared to class. This occurs because the typical number of residents from the other race group—e.g., Black neighbors for White individuals and White neighbors for Black individuals—is smaller than the typical number from the other class group. Most of the variation in this case is contained within [-5, 5]. The average $N_{i,g}$, where g is the other race group, is approximately 2,080. The instrument can adjust the environment from one with 936 opposite-sex and 1,144 own-sex individuals to one with 1,144 opposite-sex and 936 own-sex individuals, representing a shortage versus a surplus of 208 people in the other race group.

We overlay the first stage relationship between childhood neighborhood sex ratios and adult marriage market tightness in Figure VII Panels *A* and *C*. We follow Cattaneo et al. (2024) and modify equation 1 to regress $T_{i,g}$ on indicators for each ventile bin of $Z_{i,g}$. We plot the coefficients on each ventile bin against the mean value of $Z_{i,g}$ in that ventile bin and normalize the mean of the coefficients and the mean of $Z_{i,g}$ to zero. This method allows us to non-parametrically evaluate the first stage relationships before proceeding with the linear specification in equation 1. We report the coefficient from the linear model in equation 1 in Panels *A* and *C*.

We see a strong first stage relationship between sex ratios in childhood neighborhoods and marriage market tightness in adulthood. The first stage for class is steeper than race, which reflects the fact that the overall share of neighbors from the other class group is larger than the overall share of neighbors from the other race group. In either case, individuals born in places and years where a larger fraction of children from other race or class groups were of the opposite sex tend to live as adults in neighborhoods where more of their opposite-sex neighbors come from other race or class groups relative to their own-sex neighbors.

We plot the reduced form relationship between interclass and interracial marriage and market tightness in Panels *B* and *D* of Figure VII. Using methods analogous to Panels *A* and *C*, we plot the reduced form relationship using the Cattaneo et al. (2024) method and report the coefficient on β^r from equation 2.

In Panel *B*, we see a clear and significant upward sloping relationship between interclass marriage and exposure across class lines. Individuals growing up in places and years where children of the other-class group were more likely to be opposite sex are more likely to be married to a spouse from the other class group as adults. We report IV coefficients for the pooled sample as well as separate estimates for individuals from low- and high-income families in Figure VIII. The pooled sample coefficient of 0.13 implies that going from a marriage market where, for example, 20% of the own- and opposite-sex neighbors are from the other class group to one where 15% of the own-sex and 25% of the opposite-sex neighbors are from the other class group would generate a 1.3pp (41.9%) increase in the likelihood of having a spouse from the other class group. The IV estimates are similar for individuals from low- and high-income families.

Unlike class, Panel *D* shows that exposure across racial lines has no detectable impact on interracial marriage. White and Black individuals who grew up in years and places where children of the other race were more likely to be opposite-sex are no more likely to enter an interracial marriage as adults. Figure VIII presents IV estimates separately for White, Black, and pooled samples. While we cannot rule out small effects of exposure on interracial marriage, we can decisively reject exposure effects of the same magnitude as those observed for class.

While our main analysis focuses on White-Black interracial marriages, Figure A.5 presents IV coefficients for interracial marriages between White individuals and members of other racial groups. The results suggest that White-Black pairings are distinct: exposure has significantly larger effects on interracial marriages involving other racial combinations, with magnitudes resembling those observed for interclass marriages. This distinction underscores the unique dynamics of White-Black interracial marriage and suggests that other interracial pairings may be more responsive to exposure.

Our IV estimates underscore a central result: exposure appears to play a much larger role in suppressing interclass marriage rates than White-Black interracial marriage rates. We revisit this claim in Section V, where we directly quantify the impact of increasing exposure through reductions in residential segregation on interclass and interracial marriage.

In addition to demonstrating that the distribution of our instrument aligns with random Bernoulli sex assignment, Table VI shows that our IV estimates remain consistent when using within-family variation in market tightness, ruling out any role of parental sorting. In column 3, we exploit variation in market tightness between siblings, arising from both their birth years and their childhood neighborhoods. In column 4, we interact family fixed effects with childhood census tracts to iso-

late variation exclusively from birth timing among siblings who were assigned the same childhood neighborhood. In both cases, the estimates closely align with our baseline results in column 2.

Market tightness influences interclass marriage by shaping whom people marry rather than affecting their overall likelihood of being married. Column 5 of Table VI shows that market tightness has a minimal and statistically insignificant impact on the likelihood of being married. Consistent with this evidence, Figure A.6 demonstrates that market tightness operates in a class-specific manner. For instance, market tightness among neighbors from bottom parent income quartile families has the largest impact on the likelihood of marrying a spouse from a bottom quartile family. In contrast, market tightness among other quartiles has small, negative, and insignificant effects on the likelihood of marrying a spouse from a bottom quartile family.

Our baseline estimates define childhood neighborhoods using parent locations when the child is age 18. Choosing this later age improves the precision of the IV estimates by bringing the measurement of childhood sex ratios chronologically closer to our adult exposure estimates, with even a year of overlap since adult exposure is measured between ages 18 and 27. To ensure that defining childhood neighborhoods at age 18 is not driving our results, we re-estimate our IV model using each age from 0 to 18 to define childhood neighborhoods and measure the instrument. Figure A.7 shows that our baseline finding—a large exposure effect by class and a small or zero effect by race—holds even when using neighborhoods where individuals were born.

Variation in sex ratios in the neighborhoods individuals live in as adults, rather than where they grew up, can bias IV estimates if individuals sort into neighborhoods based on sex (e.g., certain areas may have jobs with a high proportion of male workers). After age 18, neighborhoods are assigned using the individual's own data rather than their parents'. As shown in Figure A.7, the IV estimates increase rapidly when variation from adult neighborhoods is introduced, highlighting the importance of isolating sex ratios during childhood to ensure unbiased estimates.

How can we be certain that sex ratios in childhood neighborhoods influence marriage outcomes through their effect on adulthood marriage markets? It is possible that growing up with opposite-sex members of a specific race or class group could affect marriage outcomes through other mechanisms, such as learning how to interact with opposite-sex members of that group. If this were the case, we might expect that having more opposite-sex individuals from other class groups in one's own birth cohort and neighborhood would increase the likelihood of interclass marriage with individuals outside of that cohort or neighborhood range.

In Figure IX, we evaluate the impact of cohort-specific sex ratios on cohort-specific marriage outcomes. In Panel *A*, we examine how cohort-specific sex ratios within one's own class influence the probability of marrying an own-class spouse born in the prior cohort. In Panel *C*, we repeat this analysis with other-class sex ratios and interclass marriage outcomes. Similarly, in Panels *B* and *D*, we analyze the impact of sex ratios on marriage to a spouse born in the cohort after one's

own.

Across all four panels, we find that the sex ratio with the largest impact on marriage outcomes corresponds to the cohort in which the outcomes are measured. Having more opposite-sex individuals from the other class group in one's own cohort does not significantly affect the likelihood of interclass marriage in preceding or following birth cohorts. Furthermore, Figure A.8 shows that our IV estimates predominantly reflect marriages between individuals who shared the same childhood neighborhood, reinforcing the idea that exposure, rather than learning, drives these results.[15]

The childhood sex ratio instrument produces changes in adult marriage market tightness across different race and class groups. While tighter marriage markets across class lines lead to an increase in interclass marriage, tighter marriage markets across race lines show no detectable effect on interracial marriage. Tightening a marriage market does not alter the overall share of each race and class group in the market but changes the fraction belonging to each sex. Reducing segregation, however, necessitates altering market thickness—the overall share of each race and class group in a neighborhood.

IV.C From Market Tightness to Market Thickness

There are two important challenges when translating our IV estimates of the effects of market tightness into the effects of reducing residential segregation, or market thickness. The first, briefly discussed in Section IV.A, is the distinction between market tightness and market thickness. Our IV estimates alter market tightness by substituting a male for a female (or vice versa) in a given neighborhood. In contrast, reducing segregation changes market thickness by moving individuals across neighborhood boundaries. Reducing segregation holds sex ratios constant while altering the overall share of a given race or class group in a neighborhood.

Additionally, there may be spillover effects across neighborhood boundaries. Any counterfactual change in marriage outcomes for residents of one neighborhood will inevitably affect residents of others. Our IV estimates focus on the effect of market tightness on marriage outcomes within a single neighborhood, without accounting for potential spillovers across neighborhood boundaries.

We illustrate this distinction with a simple three-neighborhood example in Figure X. In this example, there are two types of individuals: those with low-income parents and those with high-income parents. We examine the marriage market impact of moving a resident with low-income parents from neighborhood C to neighborhood A. The solid lines represent initial marriage outcomes, and the dashed lines show counterfactual marriage outcomes after the move.

[15] Importantly, if the instrument's effect arises entirely from exposure, we would not expect Figure A.8 to show that 100% of the effect arises from marriages within childhood neighborhoods. As demonstrated in Section V, spillovers across neighborhoods indicate that increased exposure in one's own neighborhood can also affect the likelihood of marrying someone from a different neighborhood, driven solely by changes in exposure.

In Panel *A*, preferences for marrying within class outweigh the costs of searching outside one's neighborhood, so the individual labeled $\star$ marries person 1. In Panel *B*, the costs of searching across neighborhood boundaries outweigh preferences for marrying within class, and $\star$ marries person 2. The reduction in segregation results in one new interclass marriage in Panel *A* and two in Panel *B*. However, if we consider only the partial equilibrium effects in the treated neighborhood, A, the impact on interclass marriage appears identical in both panels. This case underscores the importance of estimating the total, general equilibrium, effect of reducing segregation to fully capture how marriage outcomes respond to such changes, which is the focus of Section V.

V A Spatial Model of the Marriage Market

We build and estimate a spatial model of the marriage market to translate our partial equilibrium IV estimates of market tightness into general equilibrium effects of reducing segregation. Our analysis focuses on two types of segregation reductions.

First, we simulate the impact of programs that reduce segregation by relocating a small share of a city's population, targeting the most segregated neighborhoods. These programs either help individuals from high-poverty areas access low-poverty neighborhoods—such as Moving to Opportunity or the Gautreaux Project (e.g., Bergman et al., 2024; Chyn, Collinson and Sandler, 2023; Ludwig et al., 2013)—or bring middle-income residents to formerly high-poverty neighborhoods, as in the Hope VI program (e.g., Haltiwanger et al., 2020; Popkin et al., 2004). Our estimates shed light on the expected marriage market impacts of such programs while also allowing us to validate our model by comparing its predictions to quasi-experimental estimates from the Gautreaux Project (Chyn, Collinson and Sandler, 2023). Since the model relies on assumptions about the structure of preferences and unobserved heterogeneity, ensuring that its predictions align with recent quasi-experimental findings is a critical first step before applying it to larger counterfactuals or other settings.

Second, we use the model to estimate interclass and interracial marriage rates in a hypothetical scenario without segregation. This exercise quantifies the total impact of segregation on marital homophily and provides insights into the effects of larger-scale reductions in segregation, such as those achieved by loosening land-use regulations (e.g., Glaeser and Gyourko, 2002; Rothwell and Massey, 2009, 2010). This counterfactual corresponds to the effect of eliminating segregation while holding fixed downstream changes to preferences or endogenous resorting.

V.A Model Setup and Estimation

We build on the Choo and Siow (2006) transferable utility model of the marriage market. Each individual belongs to a "type," defined by their characteristics, and has preferences regarding the type of spouse they wish to marry. In our setting, these characteristics include race, class, and the neighborhood where the individual lives as an adult. The market clears through a set of equilibrium transfers that balance the types who wish to marry each other. To adapt the Choo and Siow (2006) model to our context, we introduce two key modifications that we anticipate will also be valuable in other spatial transferable utility matching frameworks.

First, we explicitly model the role of geography in the marriage market by incorporating census tract into the type space and allowing individuals to have preferences over the distance to their eventual spouse. Prior applications often assume a single large marriage market or use coarse regional groupings, such as states (e.g., Chiappori, Oreffice and Quintana-Domeque, 2012, 2018; Mourifié and Siow, 2021; Siow, 2015), which may be appropriate depending on the setting or parameters of interest. In our context, the large market assumption is rejected because individuals disproportionately marry their neighbors, and hyperlocal changes in neighborhood composition meaningfully impact marriage outcomes. Expanding the type space to include geography introduces computational challenges, which we address when discussing estimation.

Our second departure is to allow for the possibility that the type space includes a trait unobserved by the researcher—referred to as the "unobserved trait"—that correlates with individuals' residential choices. For example, Black individuals living in majority-White neighborhoods may possess characteristics, unobserved by the researcher, that make them more likely to prefer or be preferred by a potential White spouse. Without accounting for the unobserved trait, the model would implicitly assume that the relationship between neighborhood exposure and marriage outcomes reflects a causal effect of exposure. However, while the association between cross-race exposure and interracial marriage is positive (Figure V), the IV estimate of the causal effect of cross-race exposure on interracial marriage is near zero (Figure VIII), indicating that this assumption does not hold in our setting.

The quasi-experimental estimates based on exogenous variation in sex ratios at birth—variation that is orthogonal to sorting but generates differences in exposure—can be used to disentangle the effects of sorting from the causal effects of exposure. We illustrate the logic of this argument in Appendix E through a simple simulation with two types, shown in Figure A.9. The intuition is that the observed association between intergroup marriage and cross-group exposure can arise entirely from selection—individuals living in more diverse environments may already be more likely to marry across group lines regardless of neighborhood exposure—or from the causal effect of exposure, or a combination of both.

If the observed association is primarily driven by selection, policies aimed at reducing segregation will have little to no effect on intergroup marriage. Conversely, if the association reflects a causal effect of exposure, these policies could meaningfully increase intergroup marriage. Fortunately, the IV estimates based on sex ratio variation allow us to disentangle these two mechanisms. If selection alone drives the observed pattern, the IV coefficients will tend towards zero.

Following Choo and Siow (2006), we define the marriage market as having two sides, which we refer to as "men" and "women." Each individual belongs to one of T possible types; the market contains m_j type j men and w_k type k women. Individuals have preferences over potential spouses, which consist of two components: a "systematic" component shared by all type k women over type j men (and analogously for type j men over type k women) and an idiosyncratic component that varies at the individual level by spouse type. Market clearing transfers, τ_{jk}, are determined such that the number of type k women who prefer to marry a type j man equals the number of type j men who prefer to marry a type k woman.[16] The types in the model, men indexed by j and women by k, are defined at the intersection of

$$j,k \in \{\text{Parent Income Quartile} \times \text{Race} \times \text{Adult Census Tract} \times \text{Unobserved Trait}\}.$$

We model the unobserved trait as a binary characteristic.

Preferences for men,

$$U^m_{k,i} = \alpha^m_{j(i),k} - \tau_{j(i),k} + \varepsilon_{k,i}, \tag{3}$$

and women,

$$U^w_{j,i} = \alpha^w_{j,k(i)} + \tau_{j,k(i)} + \varepsilon_{j,i}, \tag{4}$$

are a function of the systematic component, the transfer, and the idiosyncratic component. The idiosyncratic component follows a type one extreme value distribution,

$$\varepsilon_{t,i} \sim \text{Gumbel}(0,1). \tag{5}$$

The term $U^m_{k,i}$ is the preference that person i, who is a type $j(i)$ man, has over marrying a type k woman and $U^w_{j,i}$ is the preference that person i, who is a type $k(i)$ woman, has over marrying a type j man.

[16] These transfers need not be monetary but may include implicit, non-monetary exchanges between partners.

Marital surplus,

$$\gamma_{j,k} \equiv (\alpha^m_{j,k} + \alpha^w_{j,k}) - (\alpha^m_{j,0} + \alpha^w_{0,k})$$

is the value of a j,k marriage net of the utility that j and k would attain if they were single. Choo and Siow (2006) show that the equilibrium number of matches between types j and k,

$$\mu_{j,k} = \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}}, \tag{6}$$

is a function of the marital surplus and the number of types j and k who remain single, $\mu_{j,0}$ and $\mu_{0,k}$. The equilibrium also must obey adding up: the total number of type t in the market must be equal to the sum of the number of t who marry and the number who remain single:

$$m_j = \mu_{j,0} + \sum_{k=1}^{T} \mu_{j,k}$$
$$w_k = \mu_{0,k} + \sum_{j=1}^{T} \mu_{j,k}.$$

Substituting the expression for the equilibrium number of matches, $\mu_{j,k}$, into the adding up constraints yields a concise definition of the model equilibrium,

$$m_j = \mu_{j,0} + \sum_{k=1}^{T} \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}} \tag{7}$$

$$w_k = \mu_{0,k} + \sum_{j=1}^{T} \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}}. \tag{8}$$

This is a system of $2T$ equations and $2T$ endogenous variables. Before solving the model, our first step is to select a sample of neighborhoods. Solving the model for the entire U.S. would involve over five trillion endogenous variables, making it computationally infeasible. Instead, we focus on a sample of 150 census tracts and construct the sample within these tracts to be nationally representative. A map of the selected tracts, drawn from Boston, MA, is shown in XI Panel *A*. Since the sample is reweighted to be nationally representative, the results are not sensitive to the specific tracts selected. Details of this procedure are provided in Appendix C.I.

Our objective is to estimate the set of model parameters that generates equilibrium matches, μ, which most closely replicate the observed data moments. Using these parameters, we can then compute new equilibrium matches after reducing segregation. This involves adjusting the m_j and w_k to counterfactual levels that represent a more even distribution of individuals across neighborhoods by race and class. Below, we outline this procedure at a high level, referencing

the relevant appendix sections for detailed explanations of our parameterization and estimation methods.

We estimate the model using simulated method of moments (SMM) (Laffont, Ossard and Vuong, 1995; McFadden, 1989; Pakes and Pollard, 1989). In a typical application of the Choo and Siow (2006) model, the types, m_j and w_k, are known, and Equation 6 can be inverted to solve for $\gamma_{j,k}$. However, due to the unobserved trait in our setting, m_j and w_k are not directly observed.

The first set of parameters governing our estimation determines the prevalence of the unobserved trait within each race and class group and its distribution across neighborhoods. Conditional on these parameters, m_j and w_k become known. We allow each race and class group to vary in the share of individuals possessing the unobserved trait and in the relationship between the prevalence of the unobserved trait and the average income levels in each neighborhood. This method is described in Appendix C.II.

The second set of parameters are the inputs to marital surplus, $\gamma_{j,k}$. There are four key components: (1) geographic distance between j and k, including a premium for living within 50 tracts and a smooth decay between tract 0 (same tract) and 50 tracts away; (2) class, the value of a marriage between individuals from parent quartiles $c(j)$ and $c(k)$; (3) race, the value of a marriage between individuals from race groups $r(j)$ and $r(k)$; and (4) the value of the unobserved trait, which varies by race and class group. Details on these parameters, as well as those governing the unobserved trait, are provided in Appendix D.I.

Once both sets of parameters are known, $m_j, w_k, \gamma_{j,k}$ are fixed, and the system defined in Equations 7 and 8 can be used to solve for μ.[17] Using μ, we compute the simulated moments analogous to those covered in Sections III and IV, including marriage outcomes for each race and class group (e.g., Figure I), the observed relationship between intergroup marriage and exposure (e.g., Figure IV), and the causal effect of market tightness on intergroup marriage (e.g., Figure VIII). We then iteratively update both sets of parameters using our SMM objective function to minimize the distance between the simulated moments and the observed data moments. Figure A.10 shows that the model provides a close match for the data moments. Details on the SMM objective function and the included moments are provided in Appendix D.

[17] With a large number of types, solving this system becomes computationally expensive. To address this, we implement a change of variables, detailed in Appendix D.III, that reformulates the system as a set of quadratic equations. This redefined system can be solved using fixed-point iterations, which are significantly more computationally efficient than non-linear solvers relying on the $2T \times 2T$ Jacobian matrix. We solve the equations using the method described in Anderson (1965), implemented with JAX, a tool for high-performance numerical differentiation on graphical processing units (Bradbury et al., 2018).

V.B The Impact of Reducing Segregation on Marriage Outcomes

In our first counterfactual, we estimate the effect of moving a small number of people from highly segregated neighborhoods on intergroup marriage. We begin by reducing segregation by class. In Figure XI, Panels *B* and *C*, we show the fraction of residents in our selected sample, by census tract, who grew up in households with bottom and top quartile incomes. We then move one resident of each sex from a low-income family in the blue neighborhood (Panel *D*) to the red neighborhood, and one resident of each sex from a high-income family in the red neighborhood to the blue neighborhood.[18] This type of simulation closely mirrors actual policies aimed at reducing segregation, such as those that help a small number of residents move to higher-income neighborhoods, including the expansion of housing vouchers. The impacts of such policies have been studied in the context of the Moving to Opportunity study and the Gautreaux Project (e.g., Bergman et al., 2024; Chyn, Collinson and Sandler, 2023; Ludwig et al., 2013).

Small reductions in class-based segregation, focused on the most segregated neighborhoods, have large effects on interclass marriage per person moved. The model-predicted effects of this simulation on interclass marriage by neighborhood are shown in Figure XII. Neighborhoods are colored based on quantiles of the change in interclass marriage among their residents, with darker colors indicating the areas experiencing the largest increases. To quantify the overall impact, we express the results of the simulation in the same units as the slope shown in Figure V, Panel *A*: the change in interclass marriage divided by the change in neighborhood exposure to the other class group. The model predicts that moving a small number of individuals from the most segregated neighborhoods would produce a slope of 0.3, implying that a 10pp increase in exposure to the other class group would result in a 3pp increase in interclass marriage rates. This estimate is precise, with a standard error of 0.08.[19] Notably, one-third of the total effects arise from spillovers, meaning new interclass relationships that do not involve residents of either of the treated neighborhoods where individuals were moved.

We repeat this counterfactual for race-based segregation by moving a Black individual of each sex and a White individual of each sex between the same set of neighborhoods used in the prior example—these neighborhoods are also heavily segregated by race. The individuals who move are representative of the national class distribution within their respective race groups. This simulation is designed to mirror the Gautreaux Project, which provided housing vouchers in a quasi-random manner to Black families in Chicago, enabling them to move to predominantly White suburbs.

[18] We move a racially representative resident of each class based on the national share of race groups in bottom and top quartile families. This is feasible because the matches and populations in the model are continuous rather than discrete.

[19] The standard errors are based on a county-level bootstrap of the underlying data moments, as described in detail in Appendix D.IV.

Chyn, Collinson and Sandler (2023) estimates the causal effect of such moves on interracial marriage.

We find that these moves have significant effects on interracial marriage, though smaller than for class-based segregation. These modest reductions in racial segregation suggest that a 10pp increase in exposure to the other race group results in a 1.0pp increase in White-Black interracial marriage. This aligns closely with the estimates from Chyn, Collinson and Sandler (2023), which indicate a 1.1pp increase in interracial marriage following similar moves, suggesting that the model-predicted effects are consistent with the outcomes observed in a natural experiment that reduced segregation.

For both race and class, these small moves generate substantial increases in intergroup marriage per person moved. To contextualize these effects, we compare them to the observed relationship between intergroup marriage and segregation across commuting zones, shown in Figure XIII. For both race and class, the implied effects of these small moves are significantly larger than what one would infer from the commuting zone-level relationship. Furthermore, the finding that these small changes increase interracial marriage is striking, given the earlier result in Section IV—incorporated into the model estimation—that small changes in exposure to opposite-sex members of other race groups do not increase interracial marriage. This naturally raises an important question: are these effects of small reductions in segregation sustainable, or will larger reductions yield diminishing effects on intergroup marriage?

To approximate the effects of larger reductions in segregation, we estimate intergroup marriage rates in a counterfactual scenario without segregation. In this scenario, individuals are redistributed evenly by race and class across neighborhoods, and the model is used to solve for the resulting intergroup marriage rates. This exercise isolates the total effect of segregation assuming other factors are held constant—preferences remain unchanged, and individuals do not resort in response to the reduction in segregation. The results are shown in Figure XIV.

Absent segregation, there would be a significant reduction in marital homophily by class, as shown in in Figure XIV Panel *A*. Interclass marriage rates for individuals from low-income families would increase from 3.1% to 4.6%. Eliminating segregation closes 35.2% (standard error = 5.4%) of the gap between the observed interclass marriage rates and the rates that would occur if individuals married at the same rate but chose their spouse without regard to class. Comparing instead to the rates that would occur under current marriage rates but with random spouse selection, removing segregation would close 46.8% (standard error = 7.1%) of the gap. Thus, more than one-third of marital homophily by class can be attributed to residential segregation—just one of many forms of exposure.

Without segregation, there would be only a modest reduction in marital homophily by race, as shown in Figure XIV Panel *B*. The share of Black individuals married to a White spouse would

increase from 2.1% to 3.0%. While this might initially appear to be a substantial change, it is not directly comparable to the change by class due to differences in group sizes between race and class groups. Removing segregation closes only 4.3% (standard error = 0.6%) of the gap between observed interracial marriage rates and the rates predicted under the fully random benchmark, and 13.2% (standard error = 1.9%) of the gap under the random spouse benchmark. Thus, residential segregation explains only a small portion of the divergence between observed interracial marriage rates and the no homophily benchmark.

The differences between the model-predicted effects of small versus large changes in segregation help reconcile findings in the existing literature that might initially appear contradictory. Small changes in exposure can have meaningful effects even in the presence of strong average preferences for within-race marriage (Corno, La Ferrara and Burns, 2022; Fisman et al., 2008; Gordon and Reber, 2018; Wong, 2003). This is because, with heterogeneous preferences, there will always be a small segment of the population open to interracial marriage, even when average preferences favor within-race marriage. Moving a small number of individuals between the most segregated neighborhoods allows these marriages to form. However, the gains from such moves quickly diminish as marriages form among the limited subset of individuals who prefer an interracial marriage. This parsimonious model of the marriage market replicates quasi-experimental estimates of the effects of small reductions in segregation (Chyn, Collinson and Sandler, 2023) while demonstrating that larger changes are unlikely to produce similar results, consistent with historical experiences of larger reductions in segregation (Deal, 2024*a*).

VI Conclusion

Marital homophily can exacerbate economic inequality between racial and class groups, reinforcing disparities both within and across generations (Binder et al., 2022; Chadwick and Solon, 2002; Mare, 1991, 2000). Low rates of intergroup marriage may stem from preferences for marrying within one's own group, limited opportunities for social interaction with individuals from different backgrounds—shaped by factors like segregation—or a combination of these factors. Our findings show that the mechanisms underlying interclass and interracial marriage differ substantially. Residential segregation, a single form of exposure, accounts for at least one-third of marital homophily by class, but accounts for less than 5% of marital homophily by race.

Reducing residential segregation and increasing exposure across class lines can increase rates of interclass marriage, with potential implications for fostering intergenerational mobility. These findings suggest several promising directions for future work. First, estimating the impact of interclass marriage on the outcomes of children in the next generation could illuminate whether such marriages effectively promote intergenerational mobility. This analysis will soon become feasible

as the children of the individuals in our sample reach adulthood. Second, exploring whether alternative forms of exposure beyond residential segregation are equally effective—and potentially more scalable—could identify new strategies for enhancing cross-class interactions. Third, analyzing the extent to which these efforts can reduce class-based polarization, which has grown significantly in recent decades, could offer valuable insights into fostering greater social cohesion and reducing class divides (Chetty et al., 2022*b*; Chinoy and Koenen, 2024; Doob, 2015; Fischer and Mattson, 2009; Putnam, 2016).

Reducing marital homophily by race is likely to require efforts that go beyond addressing racial segregation. A valuable direction for future research would be to better understand the roots of preferences for within-group versus across-group marriages. In our analysis, we hold preferences fixed when predicting the effects of desegregation policies. However, these preferences may not be entirely exogenous and could be influenced by experiences and social norms (e.g., Corno, La Ferrara and Burns, 2022). For instance, living in a more racially diverse community as a child might make an individual more open to interracial marriage later in life (Merlino, Steinhardt and Wren-Lewis, 2019). Investigating how these preferences are shaped and evolve is a natural next step for future research.

References

Abramitzky, Ran, Adeline Delavande, and Luis Vasconcelos. 2011. "Marrying Up: The Role of Sex Ratio in Assortative Matching." *American Economic Journal: Applied Economics*, 3(3): 124–157.

Alan, Sule, Enes Duysak, Elif Kubilay, and Ipek Mumcu. 2023. "Social Exclusion and Ethnic Segregation in Schools: The Role of Teachers' Ethnic Prejudice." *Review of Economics and Statistics*, 105(5): 1039–1054.

Allport, Gordon Willard, Kenneth Clark, and Thomas Pettigrew. 1954. "The Nature of Prejudice."

Althoff, Lukas, Harriet Brookes Gray, and Hugo Reichardt. 2024. "Intergenerational Mobility and Assortative Mating." Working Paper.

Ananat, Elizabeth Oltmans. 2011. "The Wrong Side(s) of the Tracks: The Causal Effects of Racial Segregation on Urban Poverty and Inequality." *American Economic Journal: Applied Economics*, 3(2): 34–66.

Anderson, Donald G. 1965. "Iterative Procedures for Nonlinear Integral Equations." *Journal of the ACM (JACM)*, 12(4): 547–560.

Andrews, Rodney, Marcus Casey, Bradley L Hardy, and Trevon D Logan. 2017. "Location Matters: Historical Racial Segregation and Intergenerational Mobility." *Economics Letters*, 158: 67–72.

Anelli, Massimo, and Giovanni Peri. 2019. "The Effects of High School Peers' Gender on College Major, College Performance and Income." *The Economic Journal*, 129(618): 553–602.

Angrist, Josh. 2002. "How Do Sex Ratios Affect Marriage and Labor Markets? Evidence from America's Second Generation." *The Quarterly Journal of Economics*, 117(3): 997–1038.

Athey, Susan, Billy Ferguson, Matthew Gentzkow, and Tobias Schmidt. 2021. "Estimating Experienced Racial Segregation in US Cities using Large-scale GPS Data." *Proceedings of the National Academy of Sciences*, 118(46): 1–9.

Bailey, Martha J., and Peter Z. Lin. 2024. "Marital Matching and Women's Intergenerational Mobility in the Late 19th and Early 20th Century US." *The Economic History of American Inequality: New Evidence and Perspectives*. University of Chicago Press.

Barlow, Fiona Kate, Stefania Paolini, Anne Pedersen, Matthew J Hornsey, Helena RM Radke, Jake Harwood, Mark Rubin, and Chris G Sibley. 2012. "The Contact Caveat: Negative Contact Predicts Increased Prejudice More than Positive Contact Predicts Reduced Prejudice." *Personality and Social Psychology bulletin*, 38(12): 1629–1643.

Bayer, Patrick, and Kerwin Kofi Charles. 2018. "Divergent Paths: A New Perspective on Earnings Differences Between Black and White Men since 1940." *The Quarterly Journal of Economics*, 133(3): 1459–1501.

Bazzi, Samuel, Abel Brodeur, Martin Fiszbein, and Joanne Haddad. 2023. "Frontier History and Gender Norms in the United States." National Bureau of Economic Research Working Paper 31079.

Becker, Gary S. 1973. "A Theory of Marriage: Part I." *Journal of Political Economy*, 81(4): 813–846.

Belot, Michèle, and Marco Francesconi. 2013. "Dating Preferences and Meeting Opportunities in Mate Choice Decisions." *Journal of Human Resources*, 48(2): 474–508.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F Katz, and Christopher Palmer. 2024. "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice." *American Economic Review*, 114(5): 1281–1337.

Binder, Ariel, Caroline Walker, Jonathan Eggleston, and Marta Murray-Close. 2022. "Race and Mobility in US Marriage Markets: Quantifying the Role of Segregation."

Bisin, Alberto, and Thierry Verdier. 2001. "The Economics of Cultural Transmission and the Dynamics of Preferences." *Journal of Economic Theory*, 97(2): 298–319.

Bobo, Lawrence. 2001. "Racial Attitudes and Relations at the Close of the Twentieth Century." *America Becoming: Racial Trends and their Consequences*, 1: 264–301.

Bradbury, James, Roy Frostig, Peter Hawkins, Matthew James Johnson, Chris Leary, Dougal Maclaurin, George Necula, Adam Paszke, Jake VanderPlas, Skye Wanderman-Milne, and Qiao Zhang. 2018. "JAX: Composable Transformations of Python+NumPy programs."

Brainerd, Elizabeth. 2017. "The lasting effect of sex ratio imbalance on marriage and family: Evidence from World War II in Russia." *Review of Economics and Statistics*, 99(2): 229–242.

Brummet, Quentin, et al. 2014. "Comparison of Survey, Federal, and Commercial Address Data Quality." *Center for Administrative Records Research and Applications Working Paper (2014-06)*.

Buckles, Kasey, Joseph Price, Zachary Ward, and Haley EB Wilbert. 2023. "Family Trees and Falling Apples: Historical Intergenerational Mobility Estimates for Women and Men." National Bureau of Economic Research Working Paper No. w31918.

Cattaneo, Matias D, Richard K Crump, Max H Farrell, and Yingjie Feng. 2019. "Binscatter regressions." *arXiv preprint arXiv:1902.09615*.

Cattaneo, Matias D, Richard K Crump, Max H Farrell, and Yingjie Feng. 2024. "Binscatter Regressions." *The Stata Journal, Forthcoming*.

Chadwick, Laura, and Gary Solon. 2002. "Intergenerational Income Mobility Among Daughters." *American Economic Review*, 92(1): 335–344.

Charles, Kerwin Kofi, and Ming Ching Luoh. 2010. "Male Incarceration, the Marriage Market, and Female Outcomes." *The Review of Economics and Statistics*, 92(3): 614–627.

Charles, Kerwin Kofi, Erik Hurst, and Alexandra Killewald. 2013. "Marital Sorting and Parental Wealth." *Demography*, 50(1): 51–70.

Cherlin, A.J. 2010. *The Marriage-Go-Round: The State of Marriage and the Family in America Today.* Knopf Doubleday Publishing Group.

Chetty, Raj, John N Friedman, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2018. "The Opportunity Atlas: Mapping the Childhood Roots of Social Mobility." National Bureau of Economic Research Working Paper No. 25147.

Chetty, Raj, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. 2022*a*. "Social Capital I: Measurement and Associations with Economic Mobility." *Nature*, 608(7921): 108–121.

Chetty, Raj, Matthew O Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B Fluegge, Sara Gong, Federico Gonzalez, Armelle Grondin, Matthew Jacob, et al. 2022*b*. "Social Capital II: Determinants of Economic Connectedness." *Nature*, 608(7921): 122–134.

Chetty, Raj, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2020. "Race and Economic Opportunity in the United States: An Intergenerational Perspective." *The Quarterly Journal of Economics*, 135(2): 711–783.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *The Quarterly Journal of Economics*, 129(4): 1553–1623.

Chetty, Raj, Will S Dobbie, Benjamin Goldman, Sonya Porter, and Crystal Yang. 2024. "Changing Opportunity: Sociological Mechanisms Underlying Growing Class Gaps and Shrinking Race Gaps in Economic Mobility." National Bureau of Economic Research Working Paper No. 32697.

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2012. "Fatter Attraction: Anthropometric and Socioeconomic Matching on the Marriage Market." *Journal of Political Economy*, 120(4): 659–695.

Chiappori, Pierre-André, Sonia Oreffice, and Climent Quintana-Domeque. 2018. "Bidimensional Matching with Heterogeneous Preferences: Education and Smoking in the Marriage Market." *Journal of the European Economic Association*, 16(1): 161–198.

Chinoy, Sahil, and Martin Koenen. 2024. "Political Sorting in the U.S. Labor Market: Evidence and Explanations." Working Paper.

Choo, Eugene, and Aloysius Siow. 2006. "Who Marries Whom and Why." *Journal of Political Economy*, 114(1): 175–201.

Chyn, Eric, Robert Collinson, and Danielle Sandler. 2023. "The Long-Run Effects of Residential Racial Desegregation Programs: Evidence from Gautreaux." Working Paper.

Ciscato, Edoardo. 2024. "Assessing Racial and Educational Segmentation in Large Marriage Markets." *The Review of Economic Studies*, rdae115.

Ciscato, Edoardo, and Simon Weber. 2020. "The Role of Evolving Marital Preferences in Growing Income Inequality." *Journal of Population Economics*, 33(1): 307–347.

Clark, Gregory, and Neil Cummins. 2022. "Assortative Mating and the Industrial Revolution: England, 1754-2021." CEPR Discussion Paper No. DP17074.

Corno, Lucia, Eliana La Ferrara, and Justine Burns. 2022. "Interaction, Stereotypes, and Performance: Evidence from South Africa." *American Economic Review*, 112(12): 3848–3875.

Cox, Oliver C. 1940. "Sex Ratio and Marital Status Among Negroes." *American Sociological Review*, 5(6): 937–947.

Cutler, David M, and Edward L Glaeser. 1997. "Are Ghettos Good or Bad?" *The Quarterly Journal of Economics*, 112(3): 827–872.

Deal, Cameron. 2024*a*. "Exposure, Integration, and Interracial Marriage: Evidence from The Great Migration and Residential Segregation." Working Paper.

Deal, Cameron. 2024*b*. "Interracial Couples in the United States 1850–2000: Evidence from Census Data." Working Paper.

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

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980–2000." *American Economic Review*, 106(3): 479–524.

Doob, Christopher. 2015. *Social Inequality and Social Stratification in US Society.* Routledge.

Easterlin, Richard A. 1961. "The American Baby Boom in Historical Perspective." *The American Economic Review*, 51(5): 869–911.

Eika, Lasse, Magne Mogstad, and Basit Zafar. 2019. "Educational Assortative Mating and Household Income Inequality." *Journal of Political Economy*, 127(6): 2795–2835.

Eriksson, Katherine, Gregory T Niemesh, Myera Rashid, and Jacqueline Craig. 2023. "Marriage and the Intergenerational Mobility of Women: Evidence from Marriage Certificates 1850-1920." Working Paper.

Ermisch, John, Marco Francesconi, and Thomas Siedler. 2006. "Intergenerational Mobility and Marital Sorting." *The Economic Journal*, 116(513): 659–679.

Espín-Sánchez, José-Antonio, Joseph P Ferrie, and Christopher Vickers. 2023. "Women and the Econometrics of Family Trees." National Bureau of Economic Research Working Paper No. 31598.

Fagereng, Andreas, Luigi Guiso, and Luigi Pistaferri. 2022. "Assortative Mating and Wealth Inequality." National Bureau of Economic Research Working Paper No. 29903.

Fernández, Raquel, and Richard Rogerson. 2001. "Sorting and Long-Run Inequality." *The Quarterly Journal of Economics*, 116(4): 1305–1341.

Fischer, Claude S, and Greggor Mattson. 2009. "Is America Fragmenting?" *Annual Review of Sociology*, 35: 435–455.

Fisman, Raymond, Sheena S Iyengar, Emir Kamenica, and Itamar Simonson. 2008. "Racial preferences in dating." *The Review of Economic Studies*, 75(1): 117–132.

Foster, Thomas B. 2017. "Decomposing American Immobility: Compositional and Rate components of Interstate, Intrastate, and Intracounty Migration and Mobility Decline." *Demographic Research*, 37: 1515–1548.

Fox, Jeremy T. 2010. "Identification in Matching Games." *Quantitative Economics*, 1(2): 203–254.

Fryer, Roland G. 2007. "Guess Who's Been Coming to Dinner? Trends in Interracial Marriage Over the 20th Century." *Journal of Economic Perspectives*, 21(2): 71–90.

Galichon, Alfred, and Bernard Salanié. 2022. "Cupid's Invisible Hand: Social surplus and Identification in Matching Models." *The Review of Economic Studies*, 89(5): 2600–2629.

Glaeser, Edward L, and Joseph Gyourko. 2002. "The impact of zoning on housing affordability." National Bureau of Economic Research Working Paper No. 8835.

Goni, Marc. 2022. "Assortative Matching at the Top of the Distribution: Evidence from the World's Most Exclusive Marriage Market." *American Economic Journal: Applied Economics*, 14(3): 445–487.

Gordon, Nora, and Sarah Reber. 2018. "The Effects of School Desegregation on Mixed-Race Births." *Journal of Population Economics*, 31: 561–596.

Graham, Bryan S. 2013. "Comparative Static and Computational Methods for an Empirical One-to-One Transferable Utility Matching Model." In *Structural Econometric Models*. Vol. 31, 153–181. Emerald Group Publishing Limited.

Groves, Ernest Rutherford, and William Fielding Ogburn. 1928. *American Marriage and Family Relationships.* Ayer Publishing.

Gualdani, Cristina, and Shruti Sinha. 2023. "Partial Identification in Matching Models for the Marriage Market." *Journal of Political Economy*, 131(5): 1109–1171.

Hall, Matthew, John Iceland, and Youngmin Yi. 2019. "Racial Separation at Home and Work: Segregation in Residential and Workplace Settings." *Population Research and Policy Review*, 38: 671–694.

Haltiwanger, John C, Mark J Kutzbach, Giordano E Palloni, Henry Pollakowski, Matthew Staiger, and Daniel Weinberg. 2020. "The Children of HOPE VI Demolitions: National Evidence on Labor Market Outcomes." National Bureau of Economic Research Working Paper No. 28157.

Hitsch, Günter J, Al Hortaçsu, and Dan Ariely. 2010. "Matching and Sorting in Online Dating." *American Economic Review*, 100(1): 130–163.

Hoxby, Caroline M. 2000. "Peer effects in the classroom: Learning from gender and race variation." National Bureau of Economic Research Working Paper No. 7867.

Kalmijn, Matthijs, and Henk Flap. 2001. "Assortative Meeting and Mating: Unintended Consequences of Organized Settings for Partner Choices." *Social Forces*, 79(4): 1289–1312.

Keller, Wolfgang, and Carol H Shiue. 2023. "Intergenerational Mobility of Daughters and Marital Sorting: New Evidence from Imperial China." National Bureau of Economic Research Working Paper No. 31695.

Kennedy, Sheela, and Larry Bumpass. 2008. "Cohabitation and Children's Living Arrangements: New Estimates from the United States." *Demographic Research*, 19: 1663.

Kirkebøen, Lars, Edwin Leuven, and Magne Mogstad. 2021. "College as a Marriage Market." National Bureau of Economic Research Working Paper No. 28688.

Kirkham, Elyssa. 2019. "2019 Survey on Dating and Distance: How Far are People Willing to Look for Love?" *Ladders*.

Kreider, Rose M, Lydia Anderson, Thomas Gryn, and Chanell Washington. 2023. "Same-Sex Couples Made Up 1.7% of All Coupled Households." America Counts: Stories.

Kremer, Michael. 1997. "How Much Does Sorting Increase Inequality?" *The Quarterly Journal of Economics*, 112(1): 115–139.

Laffont, Jean-Jacques, Herve Ossard, and Quang Vuong. 1995. "Econometrics of First-Price Auctions." *Econometrica*, 953–980.

Logan, John R, Andrew Foster, Hongwei Xu, and Wenquan Zhang. 2020. "Income Segregation: Up or Down, and for Whom?" *Demography*, 57(5): 1951–1974.

Ludwig, Jens, Greg J Duncan, Lisa A Gennetian, Lawrence F Katz, Ronald C Kessler, Jeffrey R Kling, and Lisa Sanbonmatsu. 2013. "Long-Term Neighborhood Effects on

Low-Income Families: Evidence from Moving to Opportunity." *American Economic Review*, 103(3): 226–231.

Lundberg, Shelly, Robert A Pollak, and Jenna Stearns. 2016. "Family Inequality: Diverging Patterns in Marriage, Cohabitation, and Childbearing." *Journal of Economic Perspectives*, 30(2): 79–102.

Manning, Wendy D, and Jessica A Cohen. 2012. "Premarital Cohabitation and Marital Dissolution: An Examination of Recent Marriages." *Journal of Marriage and Family*, 74(2): 377–387.

Mare, Robert. 2000. "Assortative Mating, Intergenerational Mobility, and Educational Inequality." UCLA CCPR Population Working Papers.

Mare, Robert D. 1991. "Five Decades of Educational Assortative Mating." *American Sociological Review*, 15–32.

Massey, Douglas S, and Nancy A Denton. 1988. "The Dimensions of Residential Segregation." *Social Forces*, 67(2): 281–315.

Massey, Douglas S, and Nancy A Denton. 1993. *American Apartheid: Segregation and the Making of the Underclass.* Harvard university press.

Mazumder, Bhashkar. 2005. "Fortunate Sons: New Estimates of Intergenerational Mobility in the United States using Social Security Earnings Data." *Review of Economics and Statistics*, 87(2): 235–255.

McFadden, Daniel. 1989. "A Method of Simulated Moments for Estimation of Discrete Response Models without Numerical Integration." *Econometrica*, 995–1026.

McPherson, Miller, Lynn Smith-Lovin, and James M Cook. 2001. "Birds of a Feather: Homophily in Social Networks." *Annual Review of Sociology*, 415–444.

Merlino, Luca Paolo, Max Friedrich Steinhardt, and Liam Wren-Lewis. 2019. "More than Just Friends? School Peers and Adult Interracial Relationships." *Journal of Labor Economics*, 37(3): 663–713.

Mourifié, Ismael, and Aloysius Siow. 2021. "The Cobb-Douglas Marriage Matching Function: Marriage Matching with Peer and Scale Effects." *Journal of Labor Economics*, 39(S1): S239–S274.

Olivetti, Claudia, M Daniele Paserman, Laura Salisbury, and E Anna Weber. 2024. "The Emergence of the Socioeconomic Gradient in Women's Marriage Outcomes." *Review of Economics and Statistics*, 1–29.

Pakes, Ariel, and David Pollard. 1989. "Simulation and the Asymptotics of Optimization Estimators." *Econometrica*, 1027–1057.

Paolini, Stefania, Jake Harwood, and Mark Rubin. 2010. "Negative Intergroup Contact Makes Group Memberships Salient: Explaining Why Intergroup Conflict Endures." *Personality and Social Psychology Bulletin*, 36(12): 1723–1738.

Popkin, Susan J, Bruce Katz, Mary K. Cunningham, Karen Brown, Jeremy Gustafson, and Margery A. Turner. 2004. "A Decade of HOPE VI: Research Findings and Policy Challenges." The Urban Institute.

Putnam, Robert D. 2016. *Our Kids: The American Dream in Crisis.* Simon and Schuster.

Qian, Zhenchao, and Daniel T Lichter. 2007. "Social Boundaries and Marital Assimilation: Interpreting Trends in Racial and Ethnic Intermarriage." *American Sociological Review*, 72(1): 68–94.

Qian, Zhenchao, and Daniel T Lichter. 2011. "Changing Patterns of Interracial Marriage in a Multiracial Society." *Journal of Marriage and Family*, 73(5): 1065–1084.

Reardon, Sean F, and Kendra Bischoff. 2011. "Income Inequality and Income Segregation." *American Journal of Sociology*, 116(4): 1092–1153.

Rosenfeld, Michael J. 2017. "Marriage, Choice, and Couplehood in the Age of the Internet." *Sociological Science*, 4: 490–510.

Rosenfeld, Michael J, and Reuben J Thomas. 2012. "Searching for a Mate: The Rise of the Internet as a Social Intermediary." *American Sociological Review*, 77(4): 523–547.

Rosenfeld, Michael J, Reuben J Thomas, and Sonia Hausen. 2019. "Disintermediating Your Friends: How Online Dating in the United States Displaces Other Ways of Meeting." *Proceedings of the National Academy of Sciences*, 116(36): 17753–17758.

Rothwell, Jonathan, and Douglas S Massey. 2009. "The Effect of Density Zoning on Racial Segregation in US Urban Areas." *Urban Affairs Review*, 44(6): 779–806.

Rothwell, Jonathan T, and Douglas S Massey. 2010. "Density Zoning and Class Segregation in US Metropolitan Areas." *Social Science Quarterly*, 91(5): 1123–1143.

Schwartz, Christine R. 2010. "Earnings Inequality and the Changing Association Between Spouses' Earnings." *American Journal of Sociology*, 115(5): 1524–1557.

Sinha, Shruti. 2018. "Identification in One-to-One Matching Models with Nonparametric Unobservables." TSE Working Paper No. 18-897.

Siow, Aloysius. 2015. "Testing Becker's Theory of Positive Assortative Matching." *Journal of Labor Economics*, 33(2): 409–441.

Smith, Jeffrey A, Miller McPherson, and Lynn Smith-Lovin. 2014. "Social Distance in the United States: Sex, Race, Religion, Age, and Education Homophily Among Confidants, 1985 to 2004." *American Sociological Review*, 79(3): 432–456.

Solon, Gary. 1992. "Intergenerational Income Mobility in the United States." *The American Economic Review*, 393–408.

South, Scott J, and Kim M Lloyd. 1992. "Marriage Opportunities and Family Formation: Further Implications of Imbalanced Sex Ratios." *Journal of Marriage and the Family*, 440–451.

Sprung-Keyser, Ben, Nathaniel Hendren, and Sonya Porter. 2022. "The Radius of Economic Opportunity: Evidence from Migration and Local Labor Markets." US Census Bureau Center for Economic Studies Working Paper Series No. CES-22-27.

Wagner, Deborah, and Mary Lane. 2014. "The Person Identification Validation system (PVS): Applying the Center for Administrative Records Research and Applications'(CARRA) Record linkage Software." US Census Bureau CARRA Working Paper Series No. 2014-01.

Wong, Linda Y. 2003. "Why Do Only 5.5% of Black Men Marry White Women?: The Impact of the Mating Taboo, Courtship Opportunities, and Individual Endowments." *International Economic Review*, 44(3): 803–826.

Zaremba, Krzysztof. 2024. "The Role of Dating Markets in Shaping Maternal and Neonatal Health: Evidence from Sex Ratios at Birth." Working Paper.

TABLE I: Summary Statistics by Class

	Pooled (1)	Bottom 25% (2)	Quartile 2 (3)	Quartile 3 (4)	Top 25% (5)
A. Parent Characteristics					
Median Parent Household Income ($)	55170	15010	39270	73850	136400
Mean Parent Household Income Percentile Rank	0.500	0.125	0.375	0.625	0.875
Both Parents Present in Household	0.673	0.392	0.574	0.805	0.919
Parents are Same Race	0.918	0.899	0.909	0.920	0.929
B. Child Earnings at age 30					
Median Household Income ($)	34430	19600	29130	40390	56150
Mean Household Income Percentile Rank	0.500	0.375	0.454	0.540	0.631
Median Individual Income ($)	26090	15330	22430	29740	39980
Mean Individual Income Percentile Rank	0.500	0.393	0.460	0.529	0.618
Employed (Individual Income > 0)	0.801	0.720	0.787	0.834	0.863
C. Child Location at Age 27					
Has a Non-Missing Address	0.869	0.787	0.853	0.902	0.934
After Imputation	0.998	0.996	0.998	0.999	1.000
Lives within 10 Census Tracts of Age 18 Address	0.379	0.370	0.393	0.403	0.353
Lives within 50 Census Tracts of Age 18 Address	0.510	0.521	0.534	0.532	0.453
Lives in Age 18 County	0.560	0.594	0.584	0.561	0.501
Lives in Age 18 Commuting Zone	0.672	0.701	0.697	0.676	0.618
D. Child Marriage Outcomes at age 30					
Ever Married	0.451	0.365	0.434	0.502	0.500
Currently Married	0.350	0.239	0.317	0.405	0.437
Mean Spouse Parent Income Rank	0.580	0.454	0.514	0.582	0.680
Mean Spouse Individual Income Rank (Age 30)	0.603	0.523	0.559	0.601	0.671
Has a Spouse With:					
Non-missing Parent Income	0.297	0.180	0.261	0.355	0.392
Same Parent Income Quartile	0.107	0.050	0.070	0.117	0.190
Fully Random Benchmark	0.074	0.074	0.074	0.074	0.074
Random Spouse Benchmark	0.074	0.035	0.061	0.100	0.116
Number of Children (1,000s)	31060	7765	7765	7765	7765
Fraction in each Parent Income Group		0.250	0.250	0.250	0.250

Notes: This table presents summary statistics and marriage outcomes by class (parent income). See Table II for analogous statistics broken down by race. Parent income ranks are defined relative to all parents with children in the same birth cohort. All statistics are based on the primary analysis sample (individuals in the 1982-1989 birth cohorts). Panels A and B present descriptive statistics on parents' and individuals' incomes, respectively. Panel C provides information on the availability of address data and migration of individuals in adulthood. Panel D presents marriage outcomes measured at age 30. All monetary values are in 2015 dollars. Parent household income is defined as the mean when the child is 13-17, assigning zeros for non-filers. Mean spouse parent income rank is calculated only for those who are married to a spouse with non-missing parent income. See Section II for more detail on this procedure. Spouse individual income is calculated in the year the individual is 30, as opposed to when the spouse is 30, if the two are born in different years, and is calculated only for individuals who are married. The fully random benchmark is the overall marriage rate multiplied by the fraction of the population from each class group. The random spouse benchmark is equal to the actual marriage rate for each class group multiplied by the fraction of the married population that comes from each class group. For both benchmarks, we include individuals with missing class, estimating the size of this group based on the share of spouses with missing class (15%). All values in this and all subsequent tables and figures have been rounded as part of the Census disclosure avoidance protocol. All statistics in this and subsequent tables and figures cleared under Census DRB release authorizations CBDRB-FY23-CES014-009, CBDRB-FY23-0492.

TABLE II: Summary Statistics by Race

	Pooled (1)	White (2)	Black (3)	Asian (4)	Hispanic (5)	AIAN (6)	Other (7)
A. Parent Characteristics							
Median Parent Household Income ($)	55170	72280	29610	56480	34610	33400	52120
Mean Parent Household Income Percentile Rank	0.500	0.581	0.332	0.509	0.375	0.362	0.485
Both Parents Present in Household	0.673	0.799	0.291	0.806	0.545	0.554	0.647
Parents are Same Race	0.918	0.960	0.908	0.901	0.775	0.582	0.475
B. Child Earnings at age 30							
Median Household Income ($)	34430	42650	18470	46640	29900	19270	31900
Mean Household Income Percentile Rank	0.500	0.554	0.356	0.578	0.458	0.372	0.480
Median Individual Income ($)	26090	30080	17230	37510	24160	14340	25170
Mean Individual Income Percentile Rank	0.500	0.535	0.409	0.596	0.475	0.386	0.492
Employed (Individual Income > 0)	0.801	0.822	0.791	0.830	0.798	0.720	0.803
C. Child Location at Age 27							
Has a Non-Missing Address	0.869	0.903	0.823	0.884	0.860	0.745	0.856
After Imputation	0.998	0.999	0.999	0.999	0.998	0.987	0.999
Lives within 10 Census Tracts of Age 18 Address	0.379	0.377	0.349	0.400	0.399	0.411	0.345
Lives within 50 Census Tracts of Age 18 Address	0.510	0.508	0.492	0.491	0.531	0.553	0.466
Lives in Age 18 County	0.560	0.519	0.612	0.603	0.662	0.539	0.548
Lives in Age 18 Commuting Zone	0.672	0.640	0.721	0.705	0.754	0.635	0.650
D. Child Marriage Outcomes at age 30							
Ever Married	0.451	0.533	0.197	0.378	0.423	0.393	0.402
Currently Married	0.350	0.433	0.117	0.316	0.289	0.256	0.299
Mean Spouse Parent Income Rank	0.580	0.607	0.428	0.572	0.463	0.458	0.555
Mean Spouse Individual Income Rank (Age 30)	0.603	0.612	0.534	0.686	0.566	0.521	0.589
Has a Spouse With:							
Non-missing Race	0.331	0.421	0.107	0.264	0.246	0.244	0.278
Same Race	0.274	0.382	0.072	0.148	0.144	0.072	0.043
Fully Random Benchmark	0.154	0.218	0.048	0.012	0.050	0.003	0.009
Random Spouse Benchmark	0.210	0.319	0.005	0.010	0.033	0.002	0.006
Number of Children (1,000s)	31060	19080	4176	1071	4410	261	801
Fraction in each Race Group		0.609	0.133	0.034	0.141	0.008	0.026

Notes: This table presents summary statistics analogous to Table I, but broken down by race. All racial groups, except Hispanics, exclude individuals of Hispanic ethnicity. Column 6 includes individuals with two or more racial groups listed. Note that the pooled row includes the missing race group, which accounts for approximately 5% of the sample; as a result, the sample shares listed at the bottom do not sum to 1. The benchmarks are defined as in Table I, where population shares account for spouses with missing race data (5% of spouses). For additional details, see the notes to Table I.

TABLE III: Marriage Outcomes by Class Relative to Random Matching Benchmark

Parent Income Group		Fraction Married (1)	Fraction with Spouse from Parent Income Group at Age 30: Bottom 25% (2)	Quartile 2 (3)	Quartile 3 (4)	Top 25% (5)	Missing Par. Inc. (6)
Bottom 25%	Truth	0.239	0.050	0.052	0.047	0.031	0.059
	Fully Random	0.350	0.074	0.074	0.074	0.074	0.053
	Random Spouse	0.239	0.035	0.046	0.059	0.063	0.036
Quartile 2	Truth	0.317	0.052	0.070	0.079	0.059	0.056
	Fully Random	0.350	0.074	0.074	0.074	0.074	0.053
	Random Spouse	0.317	0.046	0.061	0.078	0.084	0.048
Quartile 3	Truth	0.405	0.048	0.080	0.117	0.111	0.050
	Fully Random	0.350	0.074	0.074	0.074	0.074	0.053
	Random Spouse	0.405	0.059	0.078	0.100	0.108	0.061
Top 25%	Truth	0.437	0.031	0.060	0.111	0.190	0.045
	Fully Random	0.350	0.074	0.074	0.074	0.074	0.053
	Random Spouse	0.437	0.063	0.084	0.108	0.116	0.066

Notes: This table presents observed rates of interclass marriage, measured at age 30, alongside two benchmarks. For each of the four class quartiles, we report the overall marriage rate, the fraction of individuals with a spouse whose parents' income falls within each quartile, and the fraction with a spouse whose parent income data is missing. These spouses cannot be matched to parents using the procedure described in Section II. The fully random benchmark is calculated as the overall marriage rate multiplied by the fraction of the population from each class group. The random spouse benchmark is the actual marriage rate for each class group multiplied by the fraction of the married population belonging to each class group. For both benchmarks, we include individuals with missing class, estimating the size of this group based on the share of spouses with missing class data (15%). When calculating each benchmark, we assume that marriage outcomes do not vary by sex.

TABLE IV: Marriage Outcomes by Race Relative to Random Matching Benchmark

Race Group		Fraction Married (1)	Fraction with Spouse from Race Group at Age 30: White (2)	Black (3)	Asian (4)	Hispanic (5)	AIAN (6)	Other (7)	Missing Race (8)
White	Truth	0.433	0.382	0.005	0.005	0.020	0.002	0.007	0.012
	Fully Random	0.358	0.218	0.048	0.012	0.050	0.003	0.009	0.018
	Random Spouse	0.433	0.319	0.019	0.013	0.049	0.003	0.009	0.021
Black	Truth	0.117	0.021	0.072	0.001	0.008	0.000	0.004	0.009
	Fully Random	0.358	0.218	0.048	0.012	0.050	0.003	0.009	0.018
	Random Spouse	0.117	0.086	0.005	0.004	0.013	0.001	0.002	0.006
Asian	Truth	0.316	0.077	0.006	0.148	0.017	0.001	0.015	0.052
	Fully Random	0.358	0.218	0.048	0.012	0.050	0.003	0.009	0.018
	Random Spouse	0.316	0.233	0.014	0.010	0.036	0.002	0.007	0.016
Hispanic	Truth	0.289	0.083	0.008	0.005	0.144	0.001	0.006	0.043
	Fully Random	0.358	0.218	0.048	0.012	0.050	0.003	0.009	0.018
	Random Spouse	0.289	0.213	0.013	0.009	0.033	0.002	0.006	0.014
AIAN	Truth	0.256	0.132	0.007	0.003	0.020	0.072	0.010	0.012
	Fully Random	0.358	0.218	0.048	0.012	0.050	0.003	0.009	0.018
	Random Spouse	0.256	0.189	0.011	0.008	0.029	0.002	0.005	0.013
Other	Truth	0.299	0.159	0.021	0.021	0.031	0.003	0.043	0.021
	Fully Random	0.358	0.218	0.048	0.012	0.050	0.003	0.009	0.018
	Random Spouse	0.299	0.220	0.013	0.009	0.034	0.002	0.006	0.015

Notes: This table presents observed rates of interracial marriage, measured at age 30, alongside two benchmarks. For each of the six race groups, we report the marriage rate, the fraction of individuals with a spouse from each race group, and the fraction with a spouse whose race is missing. Spouses are assigned race following the procedure in Section II, regardless of whether they can be matched to parents. The rate of missing spouse race is substantially lower than that of missing spouse class for this reason. The fully random benchmark is calculated as the overall marriage rate multiplied by the fraction of the population from each race group. The random spouse benchmark is the actual marriage rate for each race group multiplied by the fraction of the married population belonging to each race group. For both benchmarks, we include individuals with missing race, estimating the size of this group based on the share of spouses with missing race data. Unlike Table II, which uses the full sample, these benchmarks exclude individuals who are themselves missing race ($\sim 5\%$ of the sample). The benchmarks are calculated using only individuals who are white, Black, Asian, Hispanic, AIAN, or in the other race group. When calculating each benchmark, we assume that marriage outcomes do not vary by sex.

TABLE V: Distance from Eventual Spouse

	By Age			By Year Prior to Marriage					Excluding Couples at Same Address: By Age			Excluding Couples at Same Address: By Year Prior to Marriage				
	0	10	18	1	3	5	10	20	0	10	18	1	3	5	10	20
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
A. Percent of Couples that Ever Lived in Nearest Census Tracts																
100 Census Tracts	0.29	0.40	0.51	0.96	0.85	0.73	0.52	0.36	0.29	0.40	0.51	0.73	0.70	0.65	0.51	0.36
50 Census Tracts	0.23	0.34	0.45	0.96	0.81	0.68	0.46	0.30	0.23	0.34	0.45	0.66	0.63	0.58	0.45	0.30
25 Census Tracts	0.18	0.28	0.39	0.95	0.78	0.62	0.40	0.24	0.18	0.28	0.38	0.58	0.55	0.50	0.38	0.24
10 Census Tracts	0.12	0.20	0.29	0.93	0.73	0.54	0.30	0.17	0.12	0.20	0.29	0.46	0.44	0.39	0.29	0.17
B. Percent of Couples that Ever Lived in Radius																
25 Miles	0.38	0.47	0.58	0.97	0.88	0.78	0.60	0.44	0.37	0.47	0.58	0.80	0.78	0.73	0.59	0.44
10 Miles	0.24	0.34	0.44	0.95	0.82	0.68	0.46	0.30	0.24	0.34	0.44	0.67	0.64	0.59	0.45	0.30
5 Miles	0.14	0.23	0.32	0.94	0.75	0.57	0.33	0.19	0.14	0.22	0.31	0.52	0.49	0.44	0.32	0.19
C. Percent of Couples that Ever Lived in Same Geographic Area																
State	0.56	0.65	0.75	0.99	0.94	0.89	0.76	0.62	0.56	0.65	0.75	0.91	0.89	0.86	0.75	0.62
Commuting Zone	0.40	0.49	0.59	0.97	0.89	0.79	0.61	0.46	0.40	0.49	0.59	0.81	0.79	0.74	0.60	0.46
County	0.27	0.37	0.47	0.96	0.83	0.70	0.49	0.33	0.27	0.37	0.47	0.70	0.67	0.62	0.48	0.33
Census Tract	0.03	0.05	0.09	0.89	0.60	0.36	0.11	0.04	0.03	0.05	0.09	0.18	0.17	0.14	0.09	0.04

Notes: This table presents statistics on the geographic proximity of married couples during childhood and adulthood. All statistics are based on a dataset with one row per couple. Ages are evaluated relative to the younger member of the couple if the individuals were not born in the same year. For ages 18 and below, parent addresses are used; for later ages, the individuals' own addresses are used. Missing years are imputed separately for childhood and adulthood locations. For more details on the construction of the address panel, see Section II. Columns 1–3 report the fraction of eventual couples who had ever lived within a defined geographic area by ages 0, 10, and 18, respectively. Columns 4–7 present analogous statistics evaluated at specific points in time relative to the year the couple marries. In Columns 1–7, couples may appear geographically close because they are living together at the same address. Columns 8–14 repeat these calculations, excluding years in which couples resided at the exact same address.

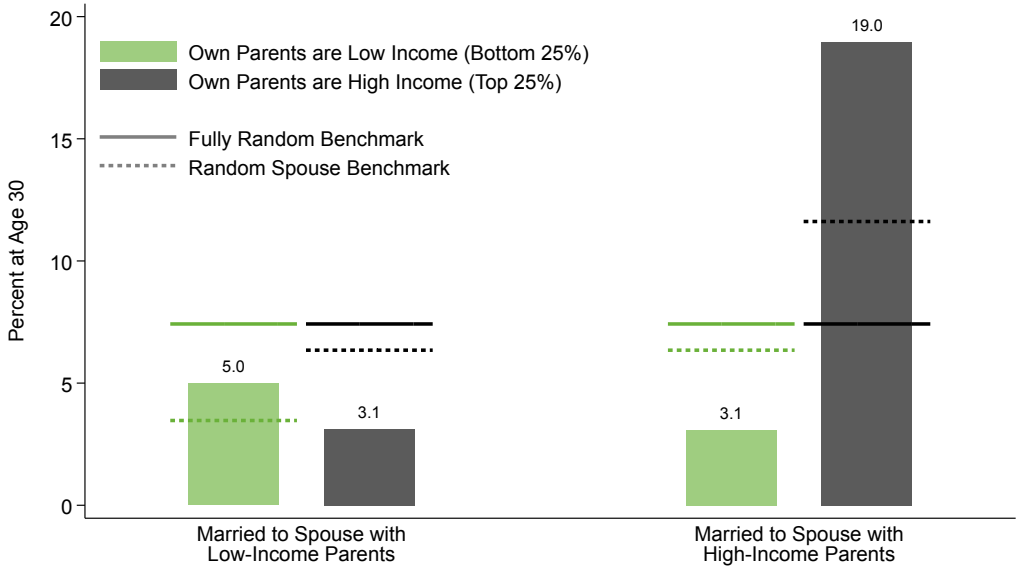
TABLE VI: Effect of Market Tightness Among Other Race or Class Group on Intergroup Marriage and Overall Marriage Rates

		IV Estimates			
	First Stage	Y = Married to Other Group			Y = Married
	(1)	(2)	(3)	(4)	(5)
A. Estimates by Class					
Bottom & Top 25%	0.173	0.132	0.113	0.140	0.027
	(0.006)	(0.027)	(0.038)	(0.034)	(0.083)
Observations (1,000s)	15070	15070	7062	5399	15070
B. Estimates by Race					
White & Black	0.035	0.001	-0.022	-0.001	0.035
	(0.001)	(0.018)	(0.027)	(0.024)	(0.173)
Observations (1,000s)	22660	22660	11750	8987	22660
Baseline Controls	Yes	Yes	Yes	Yes	Yes
Family FEs	No	No	Yes	No	No
Family x Childhood Tract FEs	No	No	No	Yes	No

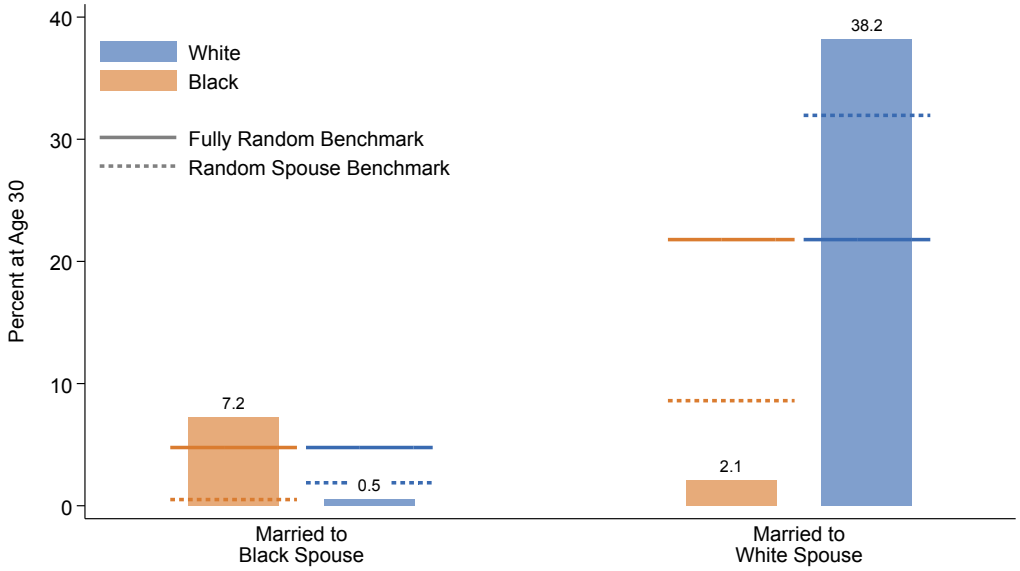
Notes: This table presents regression coefficients for the effect of exposure to different class and race groups on marriage outcomes. In Panel A, the sample consists of individuals from low-income (bottom quartile) and high-income (top quartile) families. In Panel B, the sample consists of white and Black individuals. The baseline controls include class (or race) by sex by age 18 Census tract fixed effects and linear controls for the fraction of neighbors who are own-sex and the fraction of neighbors in the other class (or race) group, interacted with class (or race) and sex fixed effects. In Column (1), we present the first stage coefficients, where the endogenous treatment variable is the market tightness measure defined in Section IV.B, and the instrument is the sex-ratio variation described in Section IV.A. In Column (2), we present the baseline IV coefficients shown in Figure VIII. In Panel A, the outcome is an indicator for having a spouse from the other class group (e.g., a spouse from a low-income family if the individual is from a high-income family, and vice versa). In Panel B, the outcome is an indicator for having a spouse from the other race group (e.g., a white spouse if the individual is Black, and vice versa). In Column (3), we restrict the sample to siblings and include family fixed effects instead of the class (or race) by sex by age 18 Census tract fixed effects. In Column (4), we further restrict the sample to siblings who lived in the same Census tract at age 18. In Column (5), we report outcomes on the extensive margin using an indicator for being married at all. Standard errors are clustered at the level of a person's age 18 county in each specification.

FIGURE I: Marital Homophily by Class and Race

A: Rates of Interclass vs. Within-Class Marriage



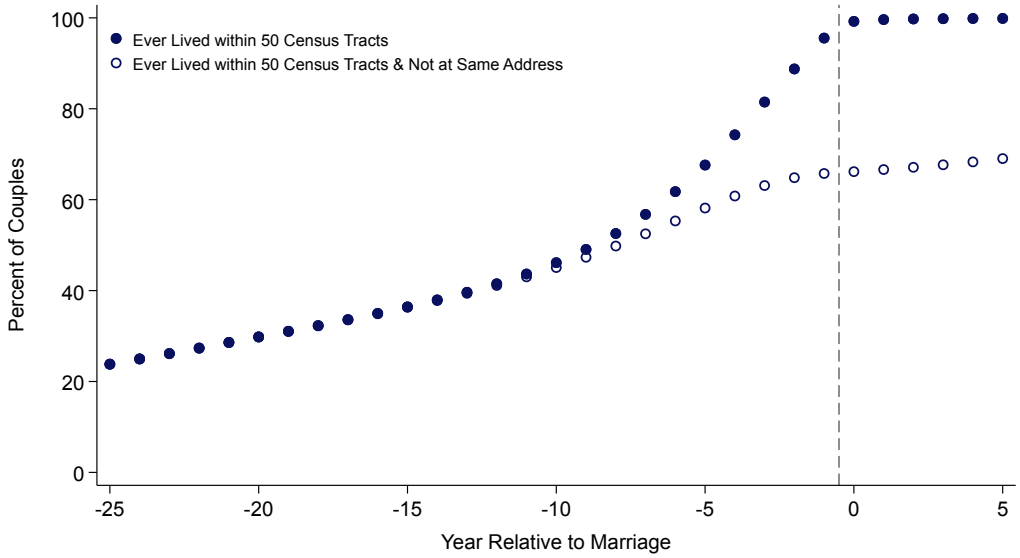
B: Rates of Interracial vs. Within-Race Marriage



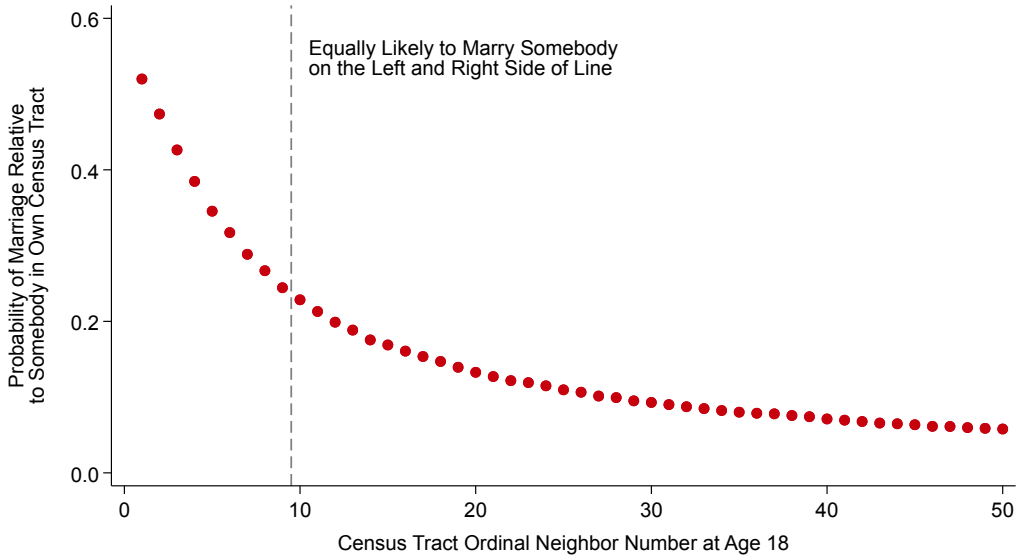
Notes: This figure summarizes key patterns of marital homophily by class (parent income) and race. In Panel *A*, we plot the fraction of individuals married at age 30 to a spouse from a low-income (bottom parent quartile) or high-income (top quartile) family, separately for individuals from low- and high-income families. Unmarried individuals are included in the calculations. In Panel *B*, we construct analogous statistics using a sample of white and Black individuals. We include two benchmarks. The fully random benchmark is calculated as the overall marriage rate multiplied by the fraction of the population from each class or race group. The random spouse benchmark is the actual marriage rate for each class or race group multiplied by the fraction of the married population belonging to each class or race group. For both benchmarks, individuals with missing class or race data are included in the population, with their share estimated based on the proportion of spouses with missing data. Similar statistics for other class and race groups are available in Tables III and IV, respectively.

FIGURE II: The Role of Neighborhoods in Marriage Markets

A: Distance From Spouse by Years from Marriage



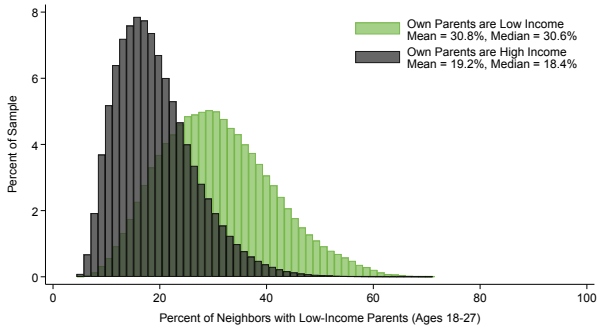
B: Marriage Probability Decay by Distance



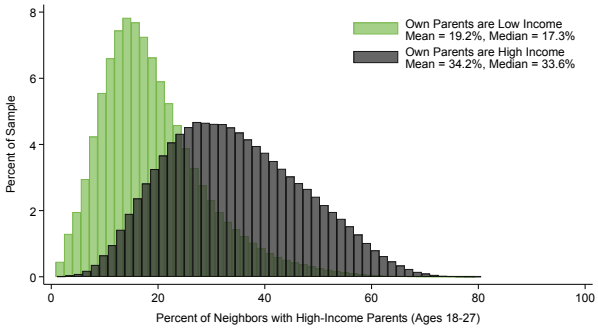
Notes: This figure presents statistics on the local nature of marriage markets. In Panel *A*, we plot the fraction of couples who had ever lived within 50 Census tracts by a given year relative to the year they married. The solid circles represent individuals who lived near each other, including those who lived together at the same address. The hollow circles show the fraction of couples who had ever lived within 50 tracts, excluding years when they resided at the exact same address. For similar statistics at fixed ages and for different geographic areas, see Table V. In Panel *B*, we plot the probability that individuals marry someone from each of the 50 nearest Census tracts to their age 18 address, relative to the probability of marrying someone from their exact age 18 Census tract. The dashed vertical line between 9 and 10 indicates that individuals are equally likely to marry someone from their own tract or the nearest 9 neighbors as they are to marry someone from neighbors 10–50 combined. A version of this plot showing separate probabilities by tercile of population density is available in Figure A.2 Panel *A*. That version provides the decay weights used to weight neighbors, as described in Section III.C.

FIGURE III: Segregation in Adult Neighborhoods
Distribution of Exposure to Race and Class Groups

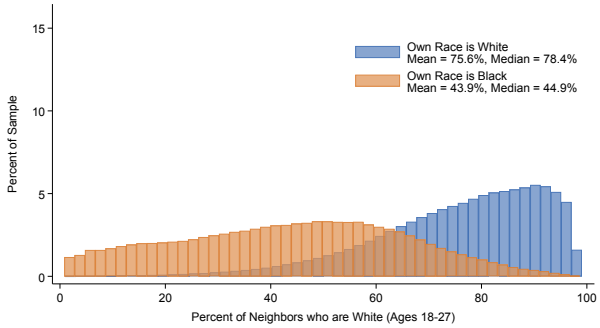
A: Low-Income (Bottom 25%) Parents



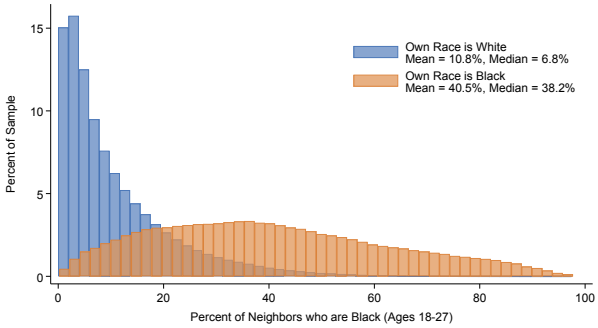
B: High-Income (Top 25%) Parents



C: White

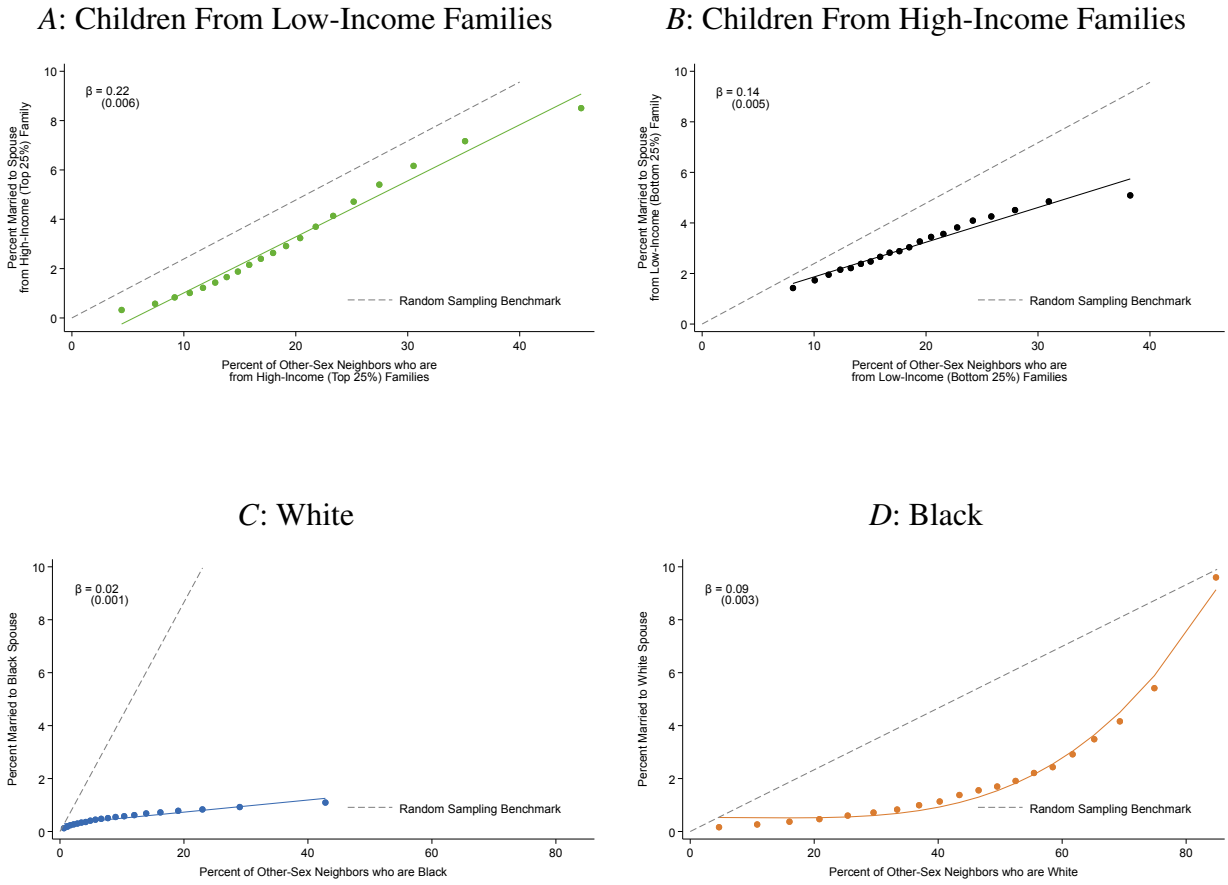


D: Black



Notes: This figure shows the distribution of neighborhood exposure to different groups in adulthood. Panels *A* and *B* show the distribution of exposure for individuals from low- and high-income families, respectively, to opposite-sex peers from low-income families (Panel *A*) and high-income families (Panel *B*). Panels *C* and *D* show the distribution of exposure for white and Black individuals, respectively, to opposite-sex peers who are white (Panel *C*) and Black (Panel *D*). In each panel, exposure is calculated as a time-weighted average of an individual's addresses from ages 18 to 27, with missing data imputed as described in Section II. Exposure is constructed using weights for geographic and age proximity. We consider individuals in the nearest 50 Census tracts, weighted using the decay weights shown in Figure II Panel *B*, and those within the same cohort or up to four cohorts younger or older, weighted using the sex-specific cohort weights shown in Figure A.2*B*. For more details on the construction of the exposure measure, see Section III.C.

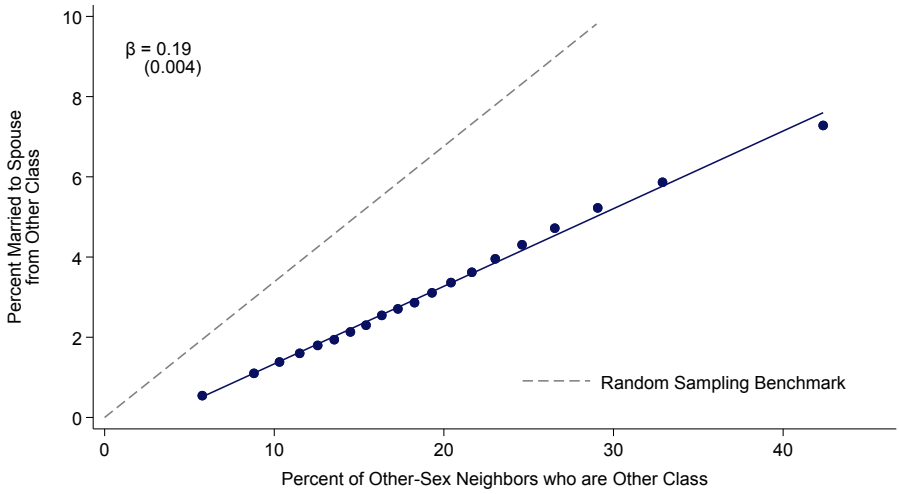
FIGURE IV: Association of Cross-Group Exposure and Cross-Group Marriage



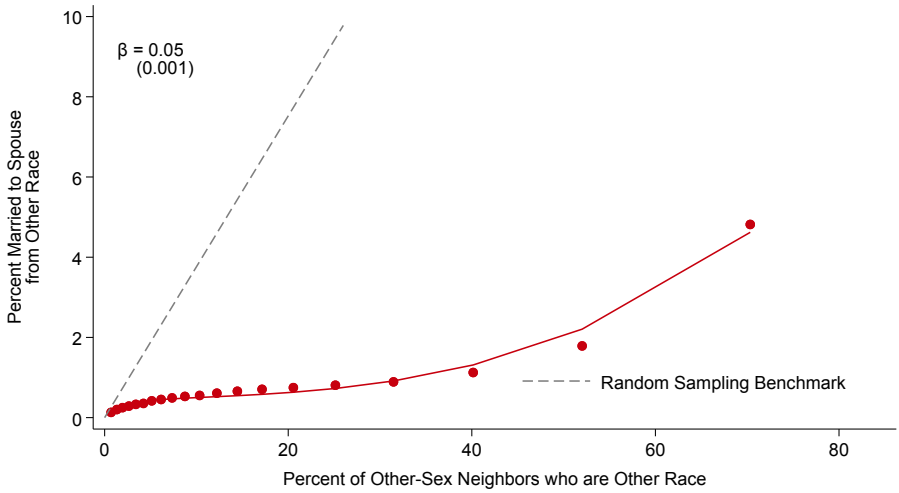
Notes: This figure presents binned scatter plots of the relationship between cross-group exposure and cross-group marriage. In Panel *A*, we show the relationship between the fraction of individuals from low parent income families who are married to a spouse from a high parent income family and the fraction of their opposite-sex peers who come from high-income families. In Panel *B*, we present the analogous relationship for high parent income individuals, using their exposure to opposite-sex low parent income peers and defining the outcome as having a spouse from a low-income family. In Panel *C*, we plot the relationship between the fraction of white individuals married to a Black spouse and their exposure in adulthood to opposite-sex Black peers. Panel *D* shows the analogous relationship for Black individuals, using their exposure to opposite-sex white peers and defining the outcome as having a white spouse. Exposure is measured in adult neighborhoods from ages 18 to 27, as described in Section III.C. The distributions of these exposure measures are shown in Figure III. To construct each plot, we bin the exposure variable into 20 equally sized bins and plot the mean exposure against the mean fraction married to the spouse type within each bin. The slope displayed in each panel corresponds to the OLS slope from regressing the marriage outcome on the exposure variable in the micro-level data, with standard errors clustered at the individual's age 23 county. A random sampling benchmark is included as a dashed line. The benchmark slope equals the group-specific marriage rate (e.g., the marriage rate for low parent income individuals in Panel *C*). The marriage rates used in the benchmarks are reported in Tables I and II for class and race, respectively.

FIGURE V: Impact of Race and Class Exposure on Spouse Characteristics

A: Rates of Interclass Marriage vs. Exposure to Other Class
Sample: Bottom 25% and Top 25%



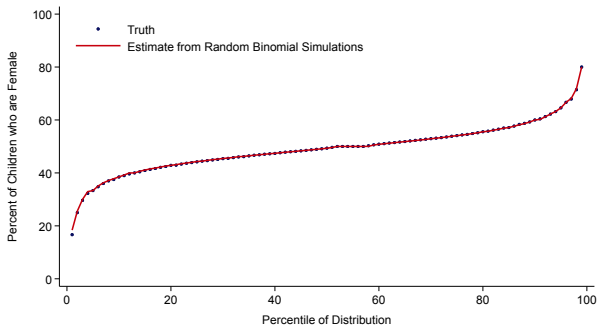
B: Rates of Interracial Marriage vs. Exposure to Other Race
Sample: White and Black



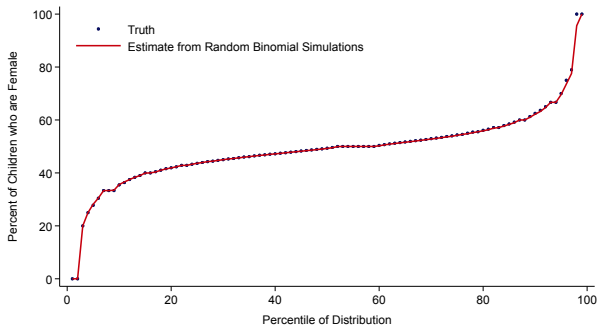
Notes: This figure presents binned scatter plots of the relationship between cross-group marriage and cross-group exposure, separately for class (Panel *A*) and race (Panel *B*). In Panel *A*, we combine the data used to create Figure IV Panels *A* and *B* to construct a single binned scatter plot. The sample includes individuals from low- and high-income families. For individuals from low-income families, exposure is defined as exposure to opposite-sex peers from high-income families, and for individuals from high-income families, exposure is to opposite-sex peers from low-income families. The marriage outcome for each group is defined as having a spouse from the other class group. The binned scatter plot is constructed as in Figure IV; see the figure notes for additional details. In Panel *B*, we combine the data used to create Figure IV Panels *C* and *D*. The sample includes individuals who are white or Black, with exposure and marriage outcomes defined as involving the other race group. The slope displayed in each panel corresponds to the OLS slope from regressing the marriage outcome on the exposure variable in the micro-level data, with standard errors clustered at the individual's age 23 county. A random sampling benchmark is included in each panel, with the slope equal to the marriage rate for the class and race samples, respectively.

FIGURE VI: Distribution of Sex Ratio across Neighborhoods Observed vs. Simulated

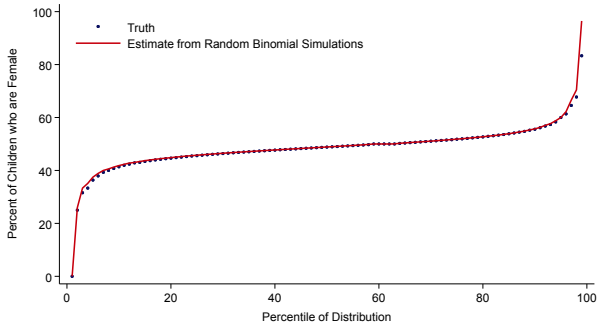
A: Children From Low-Income Families



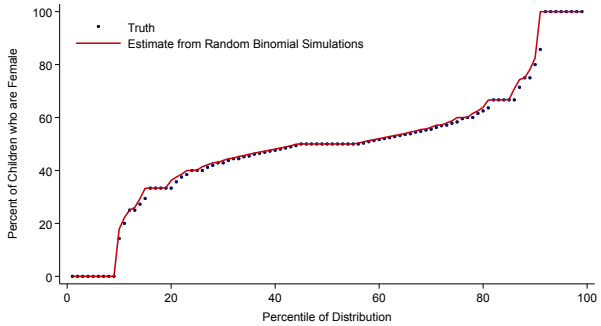
B: Children From High-Income Families



C: White



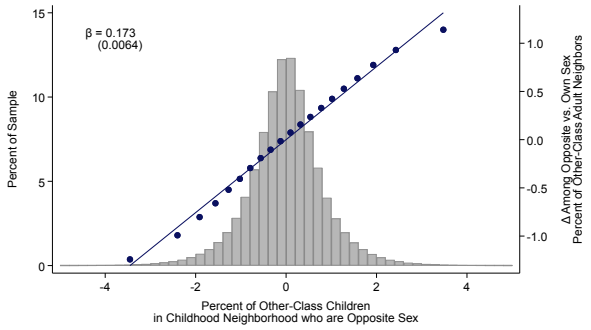
D: Black



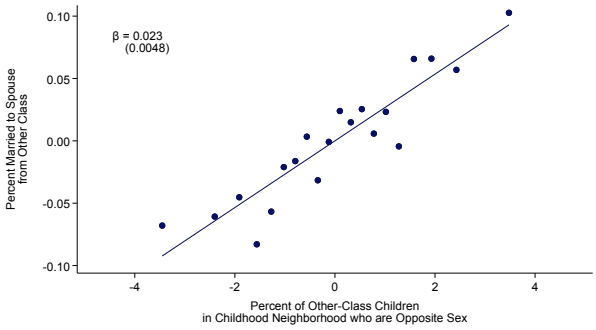
Notes: This figure examines whether observed sex ratio variation is consistent with random variation drawn from a binomial distribution, separately for individuals from low- and high-income families (Panels *A* and *B*) and for white and Black individuals (Panels *C* and *D*). The sex ratio is calculated as the fraction of individuals in each tract (and race or class group) who are female, restricted to people born between 1983 and 1987. We group tracts into percentiles and plot the mean fraction female within each percentile against the percentile number. To compare with random variation, we run binomial simulations. For each simulation, we draw the number of females from a binomial distribution centered on the true sex ratio (approximately 49% female), with the number of draws equal to the total number of individuals in each tract and group. We then construct percentiles of the simulated sex ratios in the same way as the true data. This process is repeated 1,000 times, and we compute the mean sex ratio within each percentile across the simulations. The mean simulated sex ratios are plotted alongside the true data in each figure.

FIGURE VII: Distribution of Instrument, First Stage, and Reduced Form

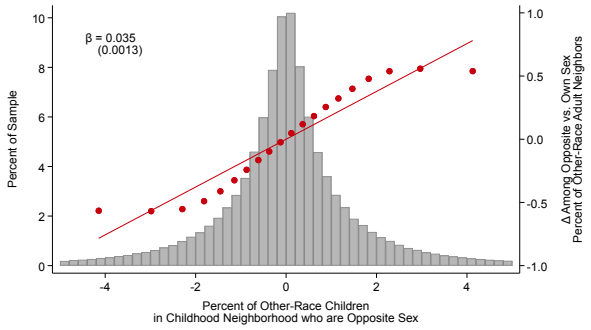
A: First Stage and *Z* Distribution
Sample: Bottom 25% and Top 25%



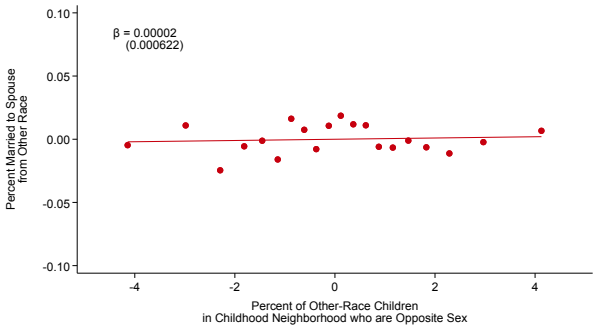
B: Reduced Form, Interclass Marriage
Sample: Bottom 25% and Top 25%



C: First Stage and *Z* Distribution
Sample: White and Black

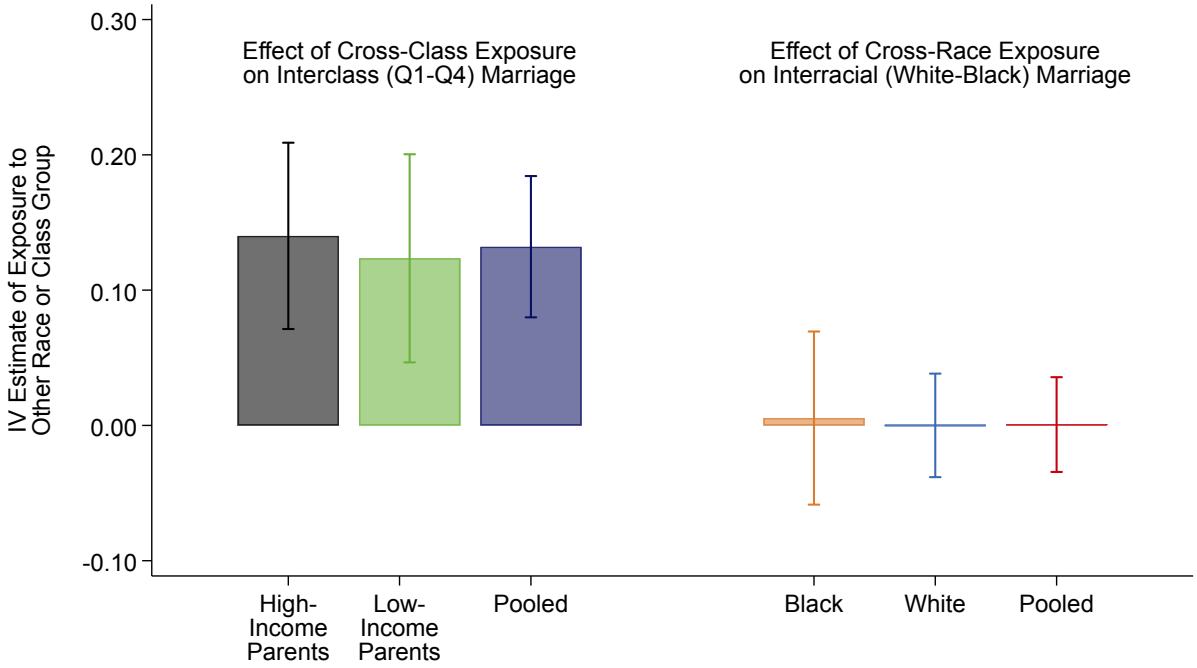


D: Reduced Form, Interracial Marriage
Sample: White and Black



Notes: This figure presents the first stage and reduced form relationships between cross-group exposure and cross-group marriage for class (Panels *A* and *B*) and race (Panels *C* and *D*). The class results are estimated using individuals from families in the bottom or top parent income quartile, while the race results are estimated using white and Black individuals. Panel *A* depicts the variation in the instrument—the fraction of other-class neighbors (e.g., high-income neighbors for individuals from low-income families) in an individual's childhood neighborhood (measured at age 18) who are the opposite sex. The distribution is residualized using class by sex by age 18 Census tract fixed effects and linear controls for the fraction of neighbors who are own-sex and the fraction of neighbors in the other class group, interacted with class and sex fixed effects. The first stage relationship between the fraction of other-sex neighbors in adulthood (measured over ages 18-27) who are other-class and the instrument is shown as a binned scatter plot. To construct the scatter, we follow Cattaneo et al. (2019). First, we divide the instrument into 20 equal-sized bins. Then, we regress the endogenous exposure measure on indicator variables for each bin, including the same controls as in the histogram. The coefficients for the indicator variables are plotted against the mean of the instrument in each bin, centering the plot at $(0,0)$. The first stage coefficient is estimated using the same specification on individual-level data, clustering standard errors at the age 23 county level. In Panel *B*, the reduced form relationship is presented as a binned scatter plot, constructed using the same procedure as the first stage. Panels *C* and *D* present analogous figures for cross-race exposure and interracial marriage.

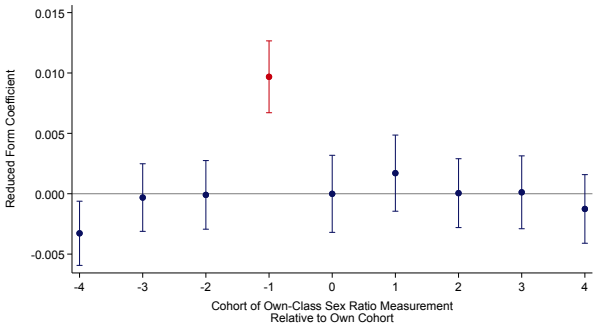
FIGURE VIII: IV Estimates of Cross-Group Exposure on Cross-Group Marriage



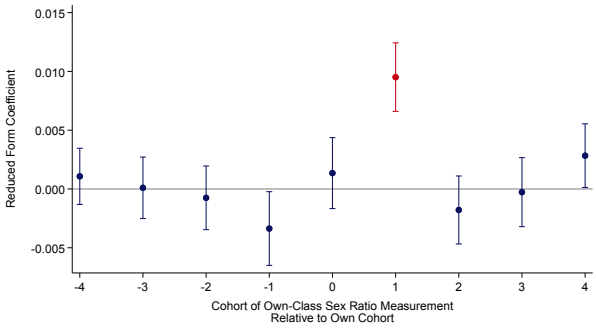
Notes: This figure summarizes the key IV results for interclass and interracial marriage. Results are presented separately for individuals from low- and high-income families, as well as for white and Black individuals, and are pooled for class and race, respectively. For details on the specification procedure, see Section IV.B and the notes to Table VI. Additional robustness checks are provided in Table VI.

FIGURE IX: Outcome Based Placebo Test
Cohort Specific Sex Ratio Variation

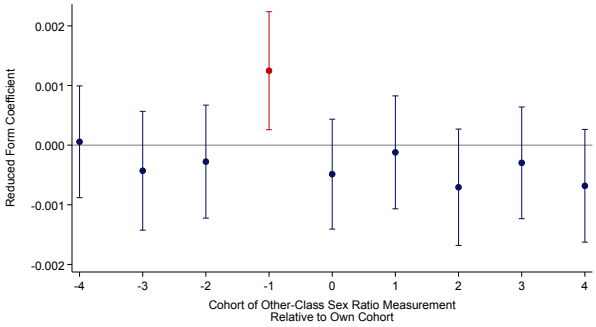
A: Own-Class Marriage to Spouse Born in Relative Cohort -1



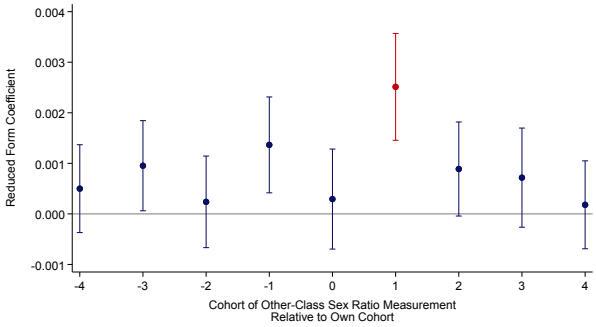
B: Own-Class Marriage to Spouse Born in Relative Cohort +1



C: Interclass Marriage to Spouse Born in Relative Cohort -1



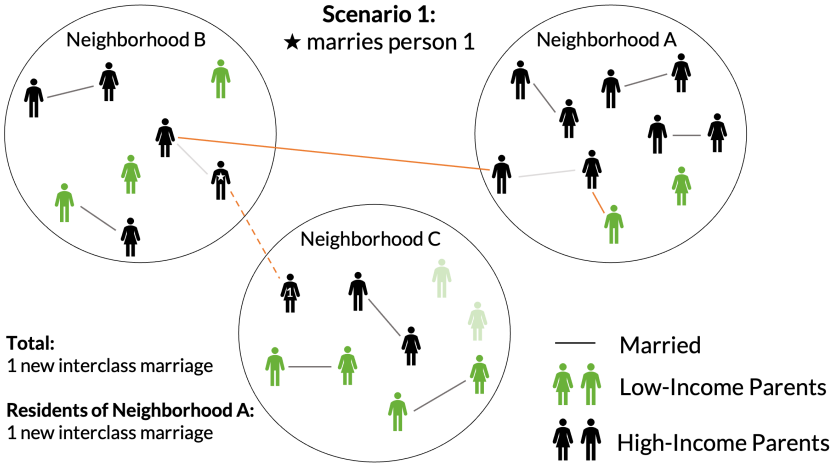
D: Interclass Marriage to Spouse Born in Relative Cohort +1



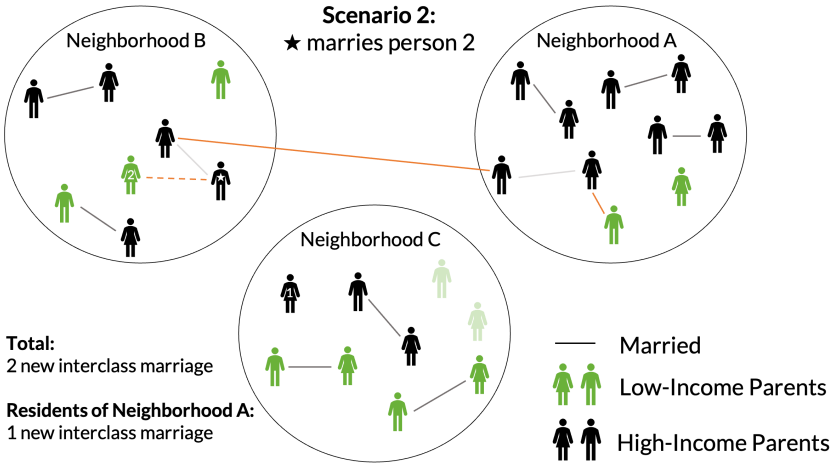
Notes: This figure presents a series of reduced form, outcome-based placebo tests to demonstrate the cohort specificity of the sex ratio variation. In each panel, the outcome variable is defined as having a spouse in a particular birth cohort by class group. In Panels *A* and *B*, the outcome is an own-class spouse, while in Panels *C* and *D*, the outcome is an other-class spouse. In Panels *A* and *C*, the spouse is from the birth cohort one year younger than the individual, whereas in Panels *B* and *D*, the spouse is from the birth cohort one year older than the individual. In each case, the outcome is regressed jointly on the nine sex ratio instruments, defined as the sex ratio in each relative birth cohort. Baseline controls are included. See Section IV.B and the notes to Table VI for further details on the specification.

FIGURE X: Three Neighborhood Example of Spillover Effects

A: ★ Marries Person 1



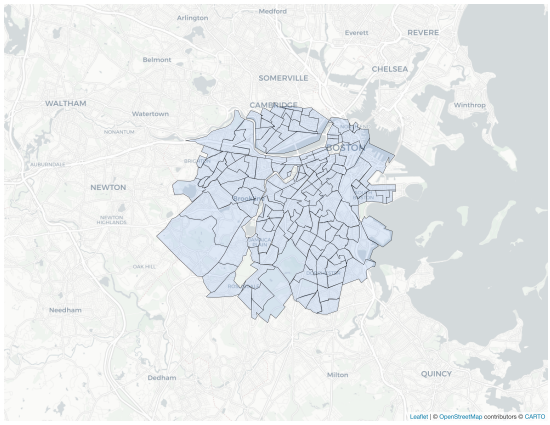
B: ★ Marries Person 2



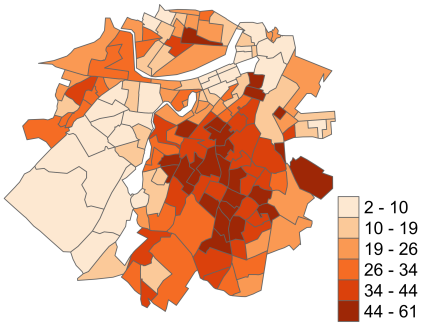
Notes: This figure presents a simplified, three-neighborhood example to illustrate the potential role of spillovers following a desegregation policy. In this example, there are two types of individuals: those from low- and high-income families. The grey lines represent the initial marriages. Each panel depicts a possible counterfactual scenario after relocating a resident from a low-income family from Neighborhood C to Neighborhood A, which serves as the reference neighborhood. In the counterfactual scenarios, the marriages change, and the new pairings are shown with orange lines. In Panel *A*, a resident from a high-income family in Neighborhood B, denoted by the star, marries a high-income resident from Neighborhood C. As a result of the policy, there is one new interclass marriage involving an original resident of Neighborhood A. In Panel *B*, the person denoted by the star marries a resident from a low-income family in Neighborhood C. In this case, there is still one new interclass marriage in Neighborhood A, but there is an additional interclass marriage involving the star individual. In this example, spillovers account for 50% of the new interclass marriages.

FIGURE XI: Model Setup and Demopgrahics of Chosen Neighborhoods

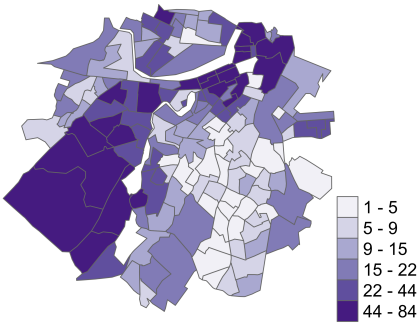
A: Analysis Sample of Census Tracts



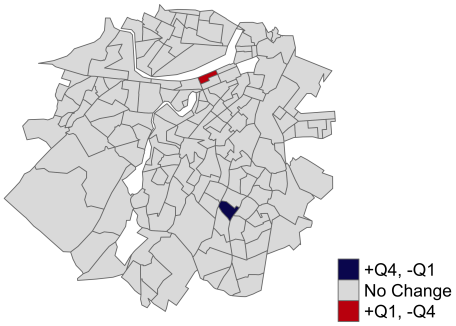
B: Percent from Low-Income Family



C: Percent from High-Income Family

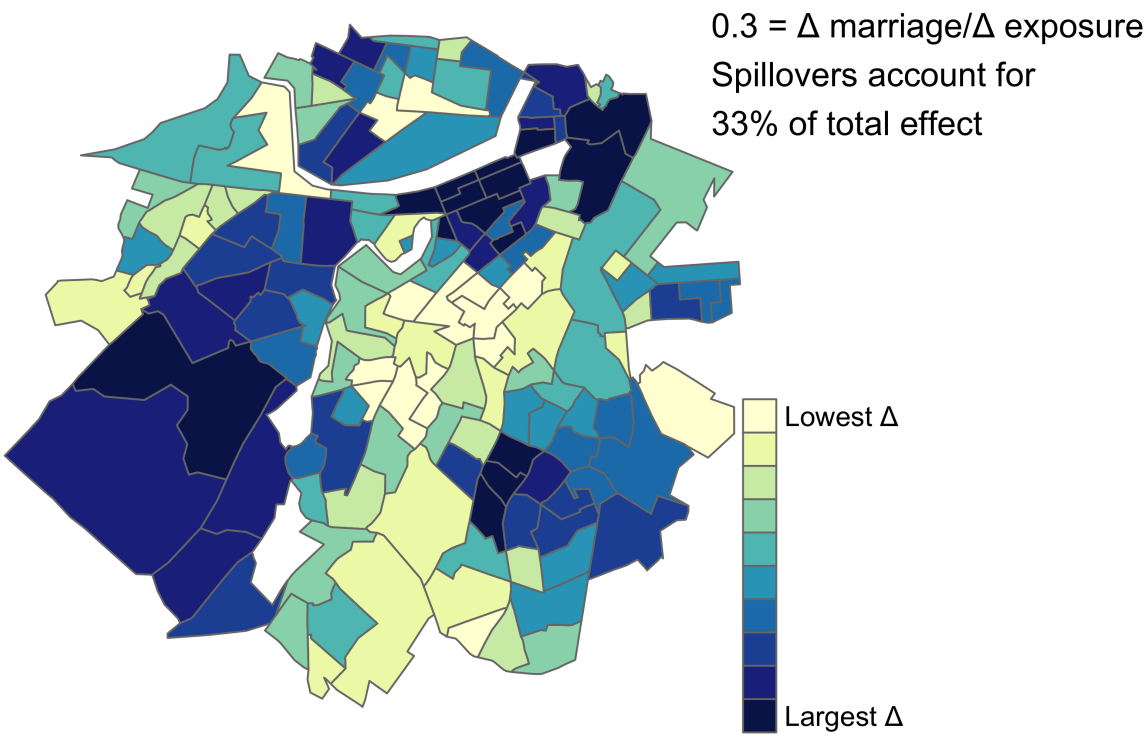


D: Simulated Moving Experiment



Notes: This figure shows the region of Boston used to estimate the model described in Section V and the policy experiment described in Section V.B. In Panel *A* we show the 150 Census tracts used to estimate the model. In Panels *B* and *C* we plot the fraction of residents from low- and high-income families, respectively. Note that we adjust the overall proportion of residents from low- and high-income families across the entire area so that the shares are nationally representative. For more details, see Appendix C. In Panel *D*, we illustrate the policy experiment described in Section V.B. In it, two racially representative individuals from low-income families (one male and one female) are moved from a tract in Dorchester, shown in blue, to a tract in Backbay, shown in red, and are replaced with two analogous individuals from high-income families from Back Bay.

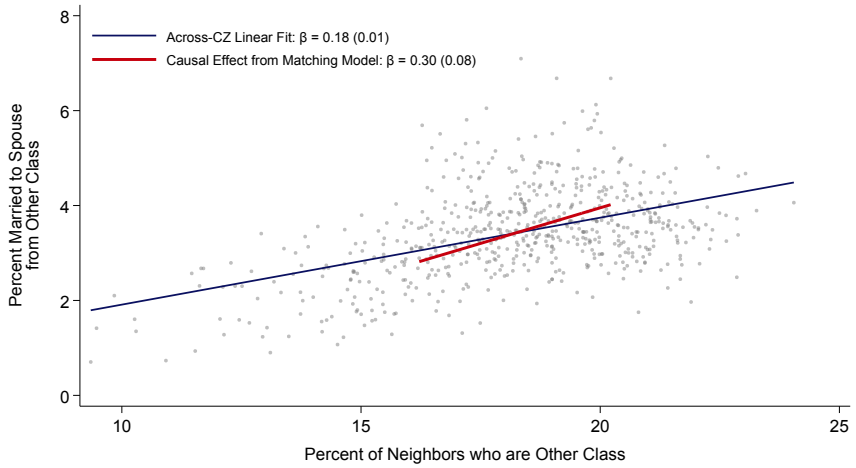
FIGURE XII: Impact of Policy Experiment on Interclass Marriage



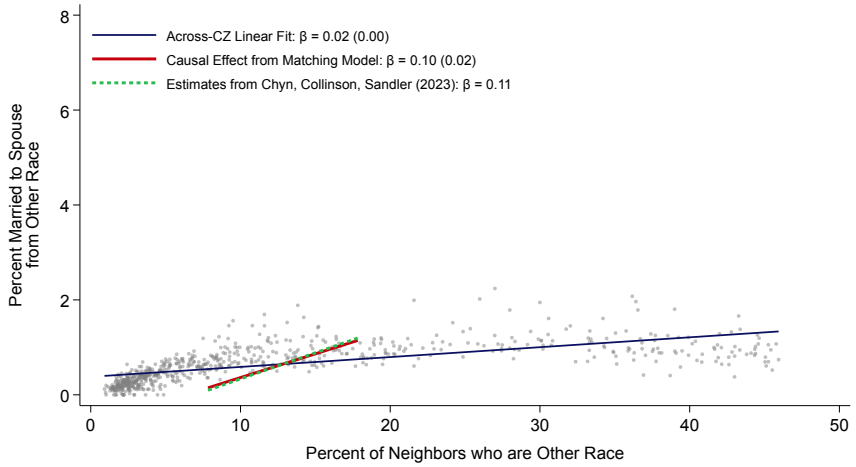
Notes: This figure illustrates the effect of the policy experiment depicted in Figure XI, Panel *D*, on interclass marriage rates. In the experiment, racially representative individuals from low-income families (one male and one female) are relocated from a tract in Dorchester to a tract in Back Bay. They are replaced by racially representative individuals from high-income families in Back Bay (one male and one female). The map colors represent quantiles of the changes in interclass marriage rates resulting from the policy change, as estimated by the model. We report the total change in the share of individuals in interclass marriages divided by the total change in exposure, expressed in the same units as the slopes shown in Figure V. Spillovers are defined as the difference between the total change in the number of interclass marriages across the entire market and the change observed for residents of the two focal tracts. These spillovers reflect new interclass marriages formed between individuals where neither party resides in the focal tracts.

FIGURE XIII: Commuting Zone Cross-Group Marriage vs. Cross-Group Exposure
Across Area Variation vs. Model Predictions from Policy Experiment

A: Rates of Interclass Marriage vs. Exposure to Other Class
Sample: Bottom 25% and Top 25%



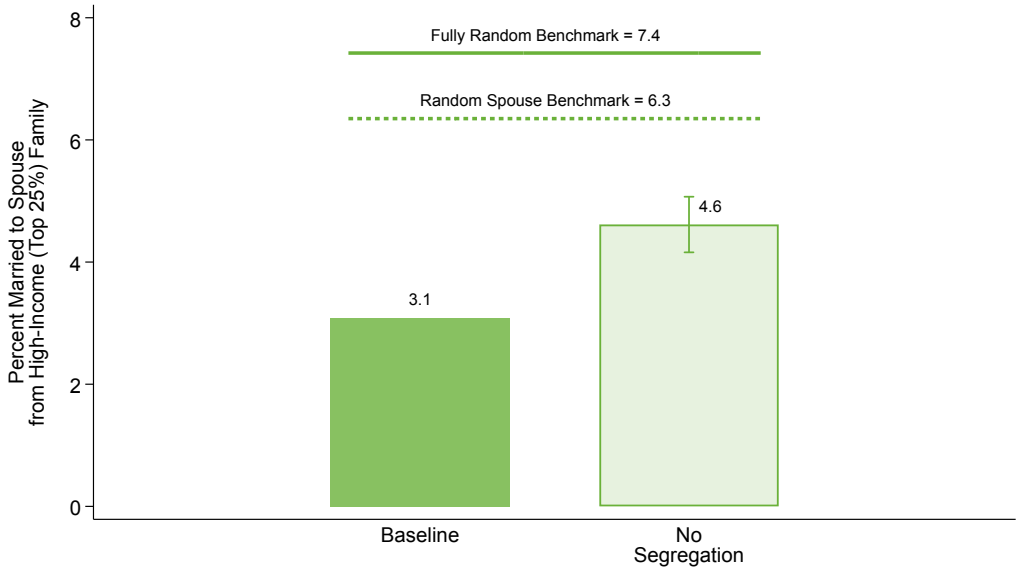
B: Rates of Interracial Marriage vs. Exposure to Other Race
Sample: White and Black



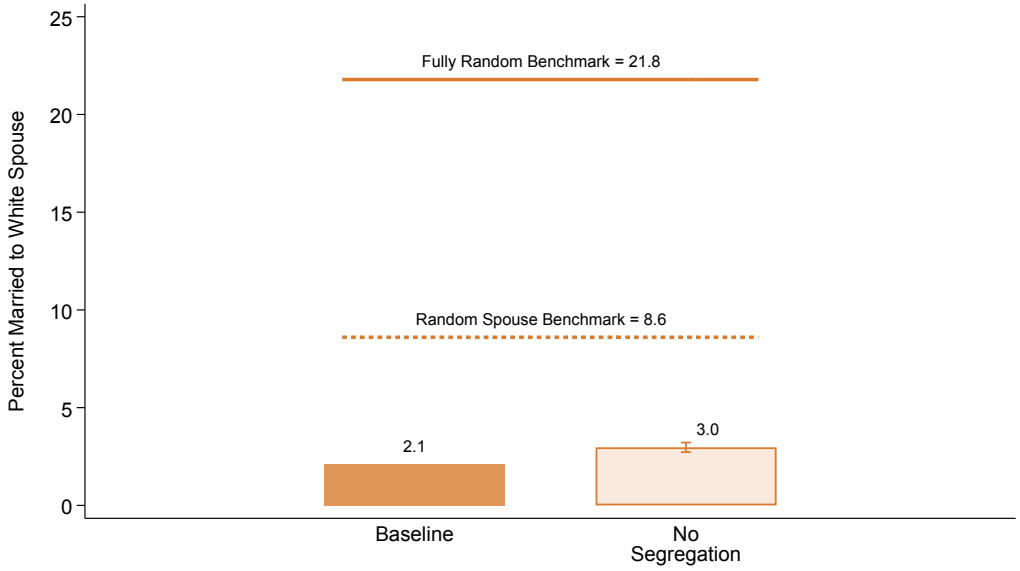
Notes: This figure plots the relationship between interclass (Panel *A*) and interracial (Panel *B*) marriage at the commuting zone level against exposure to neighbors of different class and race groups. This relationship is compared to estimates from the matching model developed in Section V. To construct the CZ measures, we assign individuals to the CZ they lived in at age 27. The exposure measure is the average of the exposure measure shown in Figure III across individuals in the CZ. The solid blue line shows the line of best fit across commuting zones, with even-weighting. The red, dashed lines in each panel show the implied slope from the counterfactual exercise described in SectionV.B and shown in XI Panel *D*. Standard errors are given in parenthesis; see Appendix D.IV for details. We compute the slope as the change in inter-group marriage as a result of the policy relative to the implied change in inter-group exposure. In Panel *B*, we additionally include a comparison to results from the Gatreaux Project (Chyn, Collinson and Sandler, 2023). They measure changes in exposure to white neighbors at the time of placement when the children were approximately 7 (Table 2 Column 4) and in 2019 when they were 26 on average (Table 5 Column 4). They also measure effects on interracial marriage (Table 6 Column 4). We estimate the change in exposure by assuming it changes linearly from the time of relocation until age 26, and then calculate the average predicted exposure over ages 18–27 to align with our baseline measure. The county level rates of interclass and interracial marriage will be available to download here.

FIGURE XIV: Intergroup Marriage Rates Absent Segregation

A: Interclass Marriage
Sample: Children From Low-Income Families



B: Interracial Marriage
Sample: Black



Notes: This figure summarizes national rates of interclass and interracial marriage, alongside counterfactual estimates under a model with no segregation. The solid bars, repeated from Figure I, provide observed rates; see the notes for that figure for additional details. The fully random benchmark is calculated as the overall marriage rate multiplied by the population share of each class or race group. The random spouse benchmark uses the actual marriage rate for each class or race group, multiplied by the fraction of the married population belonging to each class or race group. The hollow bars represent a counterfactual constructed using the model described in Section V. In this counterfactual, segregation is eliminated, as individuals are evenly distributed across neighborhoods by race and class. For further details on the model setup and counterfactual construction, see Section V.B. The standard error bars report 95% confidence intervals based on the procedure detailed in Appendix D.IV.

Supplementary Appendix

A Exposure

As detailed in Section III, we define a neighborhood as an individual's own census tract at a given age and the 50 nearest census tracts. Neighbors are defined as individuals within this radius who are within four birth cohorts older or younger than the individual. To measure exposure, we calculate a weighted average of neighbors, applying separate weights based on ordinal tract distance and relative birth cohort.

The weights used are shown in Figure A.2. The relative cohort weights are sex-specific and are constructed as the share of marriages involving individuals at a particular census tract distance or age gap, scaled by the number of marriages at the most common distance level or age gap nationally. Specifically, as shown in Figure A.2 Panel *A*, separate distance weights are applied for each tercile of the population density distribution, but for simplicity, this distinction is omitted from the notation in this section. Let M^k denote the number of marriages between individuals k tracts apart, and $M^{m,s(i)}$ denote the number of marriages involving individuals in relative birth cohort m for people of sex s. We calculate the combined weights by multiplying the distance and age weights as follows:

$$\omega_{k,m,s} = \frac{M^k}{\max\limits_{k\in[0,50]}\{M^k\}} \frac{M^{m,s}}{\max\limits_{m\in[-4,4]}\{M^{s,m}\}}.$$

We define $l_\alpha(i)$ as the census tract of individual i at age α. Let $n_{g,l_\alpha(i),k,b(i)+m}$ represent the raw count of individuals in group g located in the k-th nearest tract to $l_\alpha(i)$ and born in the birth cohort $b(i)+m$. Similarly, $n_{l_\alpha(i),k,b(i)+m}$ denotes the total number of individuals in the same tract-neighbor and relative cohort cell. By combining these raw counts with the weights, we define the following weighted sums:

$$N_{i,g}^\alpha = \sum_{k=0}^{50}\sum_{m=-4}^{m=4} \omega_{k,m,s(i)} n_{g,l(i,\alpha),k,b(i)+m},$$
$$N_i^\alpha = \sum_{k=0}^{50}\sum_{m=-4}^{m=4} \omega_{k,m,s(i)} n_{l(i,\alpha),k,b(i)+m}.$$

Here, $N_{i,g}^\alpha$ represents the weighted number of neighbors of type g for individual i in their age-α neighborhood, and N_i^α represents the total weighted number of neighbors in that neighborhood.

We define exposure, at a given age, α, as

$$e^{\alpha}_{i,g} = \frac{N^{\alpha}_{i,g}}{N^{\alpha}_{i}}.$$

We define an individual's adult neighborhood as the aggregate of their neighborhoods over ages 18–27, denoted with the superscript a for simplicity. To measure aggregate exposure during this decadal period, we apply equal weighting across these ages, as shown in the exposure distributions in Figure III. This can be expressed as:

$$e^{a}_{i,g} = \frac{1}{10}\sum_{j=18}^{27}\frac{N^{j}_{i,g}}{N^{j}_{i}}.$$

In the exposure distributions, the relative weights $\omega_{k,m,s}$ are sex-specific, but the raw counts, n, include all individuals, regardless of sex. When defining the endogenous treatment variable $T_{i,g}$ in Section IV.B, we account for the difference in exposure to type g between own-sex and opposite-sex neighbors. To incorporate this distinction, we define the sex-specific weighted counts at age α for own-sex and opposite-sex neighbors, respectively, as

$$\begin{aligned}
N^{\alpha}_{i,s(i),g} &= \sum_{k=0}^{50}\sum_{m=-4}^{m=4}\omega_{k,m,s(i)}n_{g,s(i),l(i,\alpha),k,b(i)+m} \\
N^{\alpha}_{i,s(i)} &= \sum_{k=0}^{50}\sum_{m=-4}^{m=4}\omega_{k,m,s(i)}n_{s(i),l(i,\alpha),k,b(i)+m} \\
N^{\alpha}_{i,-s(i),g} &= \sum_{k=0}^{50}\sum_{m=-4}^{m=4}\omega_{k,m,s(i)}n_{g,-s(i),l(i,\alpha),k,b(i)+m} \\
N^{\alpha}_{i,-s(i)} &= \sum_{k=0}^{50}\sum_{m=-4}^{m=4}\omega_{k,m,s(i)}n_{-s(i),l(i,\alpha),k,b(i)+m}.
\end{aligned}$$

Using these definitions, we can write the endogenous exposure measure as the average over ages 18-27:

$$T_{i,g} = \frac{N^{a}_{i,-s,g}}{N^{a}_{i,-s}} - \frac{N^{a}_{i,s,g}}{N^{a}_{i,s}} = \frac{1}{10}\sum_{j=18}^{27}\left(\frac{N^{j}_{i,-s(i),g}}{N^{j}_{i,-s(i)}} - \frac{N^{j}_{i,s(i),g}}{N^{j}_{i,s(i)}}\right).$$

B Relationship Between Treatment and Instrument

Define the fraction of childhood neighbors who are from group g as

$$P^c_{i,g} \equiv \frac{N^c_{i,g}}{N^c_i}$$

and the fraction who are from sex s as

$$P^c_{i,s} \equiv \frac{N^c_{i,s}}{N^c_i}.$$

The childhood neighborhood version of market tightness is

$$T^c_{i,g} = \frac{N^c_{i,-s,g}}{N^c_{i,-s}} - \frac{N^c_{i,s,g}}{N^c_{i,s}}.$$

Then, using Bayes' rule, we can express market tightness as a function of the instrument and the marginal group and sex shares by

$$T^c_{i,g} = Z_{i,g}\frac{P^c_{i,g}}{P^c_{i,s}} - (1 - Z_{i,g})\frac{P^c_{i,g}}{1 - P^c_{i,s}} = Z_{i,g}\left(\frac{P^c_{i,g}}{P^c_{i,s}} + \frac{P^c_{i,g}}{1 - P^c_{i,s}}\right) - \frac{P^c_{i,g}}{1 - P^c_{i,s}}. \tag{B1}$$

The above formulation shows that, conditional on $P^c_{i,g}$ and $P^c_{i,s}$, $T^c_{i,g}$ is a linear function of $Z_{i,g}$. We control for $P^c_{i,g}$ and $P^c_{i,s}$ in our IV estimates to generate a linear relationship between treatment and instrument (see Figure VII). In practice, our treatment variable $T_{i,g}$ will differ from $T^c_{i,g}$ because some individuals migrate out of childhood neighborhoods. As a result, our actual first stage coefficient will be smaller than the average value of $\left(\frac{P^c_{i,g}}{P^c_{i,s}} + \frac{P^c_{i,g}}{1-P^c_{i,s}}\right)$, which is what we would get if everybody remained in their childhood neighborhood.

C Constructing the Sample of Tracts and Types

C.I Tract Sample

The types in the model are defined at the intersection of the following characteristics

$$j,k \in \{\text{Parent Income Quartile} \times \text{Race} \times \text{Adult Census Tract} \times \text{Unobserved Trait}\}.$$

Including Census tracts in the type space adds substantial computational complexity to the model. The dimension of μ would be more than five trillion if the model were solved on the entire U.S.

Instead, we solve the model on a sample of 150 census tracts so that the dimension of μ is approximately 23 million. We choose the sample of tracts shown in Figure XI Panel *A* and construct the sample to be nationally representative using the procedure described below. The nationally representative sample is important because the moments the model is being matched to are estimated nationally. Because the model does not allow preferences to be vary by geographic area, the chosen sample matters only because it determines the topography of segregation, i.e., who lives where.

Our model is estimated outside of the Census data environment. As a result, we have to estimate tract level counts in each parent income quartile and race cell since there are no external data with this information. To do so, we begin with the data at the census tract and race level published in Chetty et al. (2018). We restrict the sample to tracts with at least 100 people living in them in the 2000 Census. Within each tract and race cell, these data also include the fraction of individuals who have parents with below median parental income and the mean parent income percentile.

We predict the share of individuals in a tract and race who are from families in a given parent income quartile by regressing

$$y_{t,q,r} = \alpha_r + \beta_1 p_t + \beta_2 s_t + \beta_3 p_t \times s_t + \varepsilon_{t,q,r}.$$

Where $y_{t,q,r}$ is the share of individual in tract t and race r who came from families in parent income quartile q, α_r are race group fixed effects, p_t is the mean parent income percentile in tract t, and s_t is the share of individuals who have below median income parents. We use this regression to predict $\hat{y}_{t,1,r}$ and $\hat{y}_{t,4,r}$—we attain R-squared values above 0.9 in both cases. We define $\hat{y}_{t,2,r}$ using the fact that $\hat{y}_{t,1,r} + \hat{y}_{t,2,r} = s_t$ and $\hat{y}_{t,3,r}$ using the fact that $\hat{y}_{t,3,r} + \hat{y}_{t,4,r} = 1 - s_t$. The $\hat{y}$ values then get rescaled so that the sample of tracts in Figure XI Panel *A* is nationally representative by race and parent income quartile.[1] Once we have counts in each tract by race by parent income quartile cell, we simulate sex ratio variation using random draws from a binomial distribution.

C.II Unobserved Trait

The next step is to assign the share of each tract, t, parent income quartile, q, and race, r, that has the unobserved trait. We do so using the logistic function

$$u_{t,q,r} = \frac{e^{\Omega_{t,q,r}}}{1 + e^{\Omega_{t,q,r}}}$$

[1]The rescaling is done by first multiplying the counts in each tract by parent income quartile by race cell by one constant per race group and then one constant by parent income quartile to ensure that 25% of individuals belong to each quartile.

to constrain the share to be between zero and one. The $\Omega_{t,q,r}$ is defined as

$$\Omega_{t,q,r} = \omega_q^1 + \omega_r^2 + (\omega_q^3 + \omega_r^4)p_t$$

so that the share with the unobserved trait is allowed to vary by race, class, and the type of neighborhood, as measured by the average income of parents, p_t. The ω parameters are estimated by our SMM procedure.

D Model Estimation

D.I Parameterization of Marital Surplus

We model marital surplus,

$$\gamma_{j,k} = f(j,k;\zeta) + \chi_{c(j),c(k)} + \rho_{r(j),r(k)} + \mathbb{1}\{u(k)=1\}\upsilon_{c(j),r(j)} + \mathbb{1}\{u(j)=1\}\upsilon_{c(k),r(k)}, \quad \text{(D2)}$$

as a function of the distance between the spouses and the preferences that each person has over the unobserved trait, class, and race. The function $f(.)$ is a continuous function of the census tract distance between a type j and a type k. There is a premium for living within 50 census tracts of each other and then a smooth decay between 0 and 50 census tracts. The ζ parameters govern the within 50 tract premium and the smoothness of the decay function.

Match value based on class, $\chi_{c(j),c(k)}$, is a function of the parent income quartile of type j, $c(j)$, and k, $c(k)$ (e.g., there is a surplus for a top-bottom quartile match, bottom-bottom, bottom-second, etc.). The $\rho_{r(j),r(k)}$ is the analogous component for race, where $r(t)$ is a function that gives the race group of type t. If k has the unobserved trait, the surplus includes that value that a person of race $r(j)$ and class $c(j)$ places on the unobserved trait, $\upsilon_{c(j),r(j)}$. If j has the unobserved trait, then surplus includes the type k's value of the unobserved trait, $\upsilon_{c(k),r(k)}$.

The distance cost portion of surplus is defined by

$$f(j,k;\zeta) = \sum_{h\in\{j,k\}} \left[\mathbb{1}\{n(j,k)\leq 50\} \times (\zeta^1_{r(h)} + \zeta^2_{c(h)}) \left(\frac{1}{n(j,k)+1}\right)^{\kappa_{c(h),r(h)}} \right]$$

where $n(j,k)$ is a function that returns the neighbor number of tract k from the perspective of j, e.g., $n(j,k)=10$ if tract k is the 10th closest tract to tract j. The parameters $\zeta^1_{r(h)}$ and $\zeta^2_{c(h)}$ determine the race and class specific premium to marrying a spouse who is living within 50 census tracts.

The

$$\kappa_{c(h),r(h)} = \frac{\phi e^{\zeta^3_{r(h)}+\zeta^4_{c(h)}}}{1+e^{\zeta^3_{r(h)}+\zeta^4_{c(h)}}}$$

parameter controls the steepness of the distance cost function, i.e., within 50 census tracts, how much do people of a given race and class group prefer to marry somebody in a nearer vs. further tract.[2] We estimate the ζ parameters as part of our simulated method of moments procedure.

D.II SMM Estimation

The SMM objective function minimizes the distance between the model moments and data moments subject to a weighting matrix:

$$\hat{\theta}(W) = \arg\min_{\theta} [\hat{\psi}_d - \hat{\psi}_s(\theta)]^\top W [\hat{\psi}_d - \hat{\psi}_s(\theta)]. \tag{D3}$$

We use three different types of moments in ψ. The first set of moments are the national marriage outcomes by race and class. For each parent income quartile, we compute the fraction of individuals married to a spouse from each of the other quartiles.[3] We compute the analogous marriage outcomes by race, but aggregate our data into four race groups for simplicity: non-Hispanic White, non-Hispanic Black, Hispanic, and other. National marriage outcomes by class and race are in Tables III and IV.

Our second set of moments are the descriptive associations between marriage outcomes and segregation that we documented in Section III.D. Similar to the national marriage outcomes, for each parent income quartile, we estimate the the association between the probability of marrying somebody from a given parent income quartile and the fraction of neighbors between ages 18 and 27 who come from that same parent income quartile. We construct the analogous estimates by race. Figure IV shows these moments for the four key race and class groups we focus our analysis on.

Our third and final set of moments are the IV estimates of the causal effect of market tightness, i.e. exposure to opposite-sex vs. own-sex members of other race and class groups, on interclass and interracial marriage. For each parent income quartile, we include the IV effects of exposure on marriage to somebody from each of the other quartiles and construct the analogous estimates by

[2] The ϕ parameter puts a maximum value on how large this exponent can be. We set ϕ to 0.15 but allow ζ to be chosen flexibly.

[3] The sum of these is equal to the overall share of people from that quartile who are married. If the model is able to match each of these moments individually, it will by definition have matched the overall marriage rate as well.

race.[4] We include random sex ratio variation in our sample, which allows us to estimate a model analogue of our IV coefficients. Section IV details the construction of the empirical version of these estimates and the IV coefficients for our four key race and class groups are summarized in Figure VIII. Figure A.10 shows the relationship between the simulated model moments and the data moments used in estimation.

Once μ is known for fixed θ, we can compute $\hat{\psi}_s(\theta)$, the model moments evaluated at θ. We repeat this process using a numerical solver, adjusting θ at each step until the objective function in Equation D3 is minimized. To ensure all three sets of moments contribute to estimation, we use a modified identity matrix for weighting.[5] The optimal weight matrix would primarily emphasize national marriage outcomes, as these are estimated with near-infinite precision. However, this would assign near-zero weight to our quasi-experimental estimates, which are crucial for our counterfactual analyses, as demonstrated in Appendix E.

D.III Computing Model Equilibrium

Once $\gamma_{j,k}$ is known, we can solve the equilibrium matches, μ using the system

$$m_j = \mu_{j,0} + \sum_{k=1}^{T} \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}}$$
$$w_k = \mu_{0,k} + \sum_{j=1}^{T} \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}}.$$

In a sample of 150 tracts, the total number of types in the model, T, is 4800. This is a system of $2T$ equations and $2T$ unknowns. Solving using the analytic Jacobian is slow and computationally expensive because the Jacobian has $4T^2$ elements which is almost 100 million in this setting. Instead, we can turn this into a system of quadratic equations if we change variables such that $\pi = \sqrt{\mu}$. Define $\nu_{m,j} \equiv \sum_{k=1}^{T} \pi_{0,k} \sqrt{e^{\gamma_{j,k}}}$ so that the male equilibrium equations become

$$m_j = \pi_{j,0}^2 + \pi_{j,0} \nu_{m,j}.$$

[4] We exclude the IV effect of exposure to own-race on within race marriage. There is little sex ratio variation within race groups because of high-levels of racial segregation. We show in Figure A.11 that there is approximately 36 times more sex ratio variation across versus within race groups. This is not an issue for estimating IV effects within class—there is three times more own-class versus own-race variation.

[5] We upweight moments corresponding to our four key groups of interest: bottom- and top-parent-income quartile, White, and Black. Moments involving only one of these groups receive 2x weight. Moments where both own and spouse samples correspond to the same key group are assigned 3x weight. Intergroup moments involving two key groups are assigned 4x weight. Adjustments to this weight matrix have little impact on our counterfactual estimates, given their precision.

We can apply the quadratic equation to solve

$$\pi_{j,0} = \frac{-\nu_{m,j} + \sqrt{\nu_{m,j}^2 + 4m_j}}{2}$$

and the analogous equations for the female equilibrium

$$\pi_{0,k} = \frac{-\nu_{w,k} + \sqrt{\nu_{w,k}^2 + 4w_k}}{2}.$$

This system of equations can be solved with a fixed point algorithm. We implement the Anderson (1965) algorithm with JAX, which does fast floating point operations on graphical processing units (Bradbury et al., 2018). After this is solved, we use $\mu = \pi^2$ to obtain μ.

D.IV Estimating the Variance of Counterfactuals

We compute standard errors of the model parameters, θ, using the delta method:

$$\text{Var}(\hat{\theta}) = (G'WG)^{-1}G'WSWG(G'WG)^{-1}.$$

Here, W is the weight matrix described in Appendix D.II, G is the Jacobian that measures the sensitivity of each model moment to each parameter, and S is the variance matrix of data moments, computed through county-level bootstrapping. Once we compute the variance of θ, we use it to calculate standard errors for the counterfactuals described in Section V.B according to:

$$\text{Var}\Big(c(\hat{\theta})\Big) = \nabla c(\hat{\theta})'\text{Var}(\hat{\theta})\nabla c(\hat{\theta}).$$

To simplify the computation of standard errors, we make a few practical assumptions. First, we exclude two parameters from G: the κ parameters for the second and third parental income quartiles described in Appendix D.I. These parameters are set near their maximum, making G singular if included, as small perturbations to these values do not impact the model moments. However, assuming large standard errors for these parameters yields similar results, as they have little effect on the counterfactuals involving the top and bottom quartile. Additionally, we exclude moments related to the share of each race and class group marrying individuals with missing data from our bootstrap. For these moments, we assume that the corresponding portion of S is diagonal, with diagonal elements set to twice the variance of the average variance of national marriage outcomes. Since the parameters are not particularly sensitive to these moments, this assumption has minimal impact on the final standard errors.

E Sorting on Unobservable in Matching Model

In this section, we simulate a simplified version of our spatial matching model to demonstrate how extending Choo and Siow (2006) to include a binary trait unobserved by the researcher (the "unobserved trait") can produce endogenous sorting into neighborhoods. We then show how our IV estimates, based on random variation in sex ratios, allow us to identify the causal effect of neighborhood exposure. First, we demonstrate that the model can produce a relationship between intergroup marriage and cross-group exposure that can arise either from sorting or from the causal effect of exposure. This implies that the observed relationship in Figure V is insufficient to estimate the effect of reducing segregation. Our IV estimates allow us to disentangle the effects of sorting and exposure, which is crucial for understanding the impact of reducing segregation.

The intuition behind how our IV estimates disentangle sorting from exposure can be illustrated with a simple example with only two types, Packers and Vikings. There are nine total neighborhoods. Packers tend to live near Green Bay, while Vikings typically reside near Minneapolis. In Figure A.9 Panel *A*, we show the fraction of neighborhood residents who are Packers and Vikings as one moves from Minneapolis to Green Bay. In Minneapolis, 90% of residents are Vikings, and the proportion of Vikings gradually decreases until reaching Green Bay, where only 10% of residents are Vikings. Conversely, 10% of Minneapolis residents are Packers, compared to 90% in Green Bay. Approximately 35% of Vikings possess a trait, unobserved by the researcher, that makes them more attractive to Packers.

In Panel *B*, we contrast two scenarios: one with endogenous sorting and one without. In the no-sorting case, shown in purple, Vikings with the unobserved trait are evenly distributed across neighborhoods. In the sorting case, shown in yellow, Vikings with the unobserved trait tend to live closer to Green Bay.

Both cases—sorting and no sorting—can produce nearly identical relationships between rates of intergroup marriage and cross-group exposure, as shown in Panel *C*. In the no-sorting case (purple), rates of intergroup marriage are higher in Green Bay due to distance costs (or a preference to marry within neighborhood). Even with a preference for within-group marriage, Vikings in Green Bay, who live near Packers, are more likely to marry a Packer than Vikings in Minneapolis, where Packers are scarce. In the sorting case (yellow), there are no distance costs. Vikings in Green Bay are more likely to marry a Packer simply because they are more likely to possess the unobserved trait, which is attractive to Packers. Distinguishing between these two scenarios—both of which produce similar relationships between intergroup marriage and cross-group exposure—is key to determining the counterfactual effects of reducing segregation.

In Panel *D*, we show that despite producing similar observable relationships between intergroup marriage and cross-group exposure, the two scenarios lead to very different IV estimates.

In the no-sorting case (purple), Vikings are much more likely to marry a Packer as the share of opposite-sex Packers in their neighborhood increases. However, when the gradient in Panel *C* is driven entirely by sorting, the relationship between the share of Vikings married to Packers and the share of nearby opposite-sex Packers disappears.[6]

We summarize the results of this illustration in Panel *E*. Without our IV estimates, the sorting and no-sorting cases are observably identical, with the same overall rate of intergroup marriage and OLS relationship between intergroup marriage and cross-group exposure. However, the two cases produce very different counterfactual effects of reducing segregation. In the no-sorting case (purple), a 10 percentage point increase in cross-group exposure leads to a 3 percentage point increase in intergroup marriage. In the sorting case (yellow), the causal effect of cross-group exposure on intergroup marriage is close to zero. Thus, distinguishing between these two cases is critical for estimating the counterfactual effects of reducing segregation, and our IV estimates enable us to do so.

[6] As the number of neighborhoods in the simulation increases, this reduced-form relationship approaches zero.

TABLE A.1: Cohabitation Outcomes by Class Relative to Random Matching Benchmark

Parent Income Group		Fraction Cohabiting (1)	Fraction with Partner from Parent Income Group at Age 30: Bottom 25% (2)	Quartile 2 (3)	Quartile 3 (4)	Top 25% (5)	Out of Sample (6)
Bottom 25%	Truth	0.427	0.097	0.092	0.078	0.051	0.109
	Fully Random	0.494	0.103	0.103	0.103	0.103	0.082
	Random Partner	0.427	0.077	0.084	0.096	0.099	0.071
Quartile 2	Truth	0.467	0.082	0.105	0.111	0.080	0.088
	Fully Random	0.494	0.103	0.103	0.103	0.103	0.082
	Random Partner	0.467	0.084	0.092	0.104	0.109	0.078
Quartile 3	Truth	0.530	0.065	0.105	0.148	0.141	0.071
	Fully Random	0.494	0.103	0.103	0.103	0.103	0.082
	Random Partner	0.530	0.096	0.104	0.119	0.123	0.088
Top 25%	Truth	0.551	0.042	0.076	0.138	0.235	0.061
	Fully Random	0.494	0.103	0.103	0.103	0.103	0.082
	Random Partner	0.551	0.099	0.109	0.123	0.128	0.092

Notes: This table presents observed rates of interclass cohabitation, measured at age 30, alongside two benchmarks. Cohabitation is defined using data from the ACS. Specifically, we include individuals in the main analysis sample who are surveyed by the ACS in the year they turn 30. An individual is classified as cohabiting if they respond to the survey as the primary respondent and list a married spouse or a non-married partner in their household as an additional respondent, or if they are listed as the spouse or partner of the primary respondent. The fully random benchmark is calculated as the overall cohabitation rate multiplied by the fraction of the population from each class group. The random partner benchmark is the actual cohabitation rate for each class group multiplied by the fraction of the cohabiting population belonging to each class group. These benchmarks are constructed analogously to those in Table III, substituting cohabitation for marriage.

TABLE A.2: Cohabitation Outcomes by Race Relative to Random Matching Benchmark

Race Group		Fraction Cohabiting (1)	Fraction with Partner from Race Group at Age 30: White (2)	Black (3)	Asian (4)	Hispanic (5)	AIAN (6)	Other (7)	Out of Sample (8)
White	Truth	0.568	0.503	0.009	0.008	0.030	0.003	0.010	0.006
	Fully Random	0.498	0.317	0.063	0.018	0.071	0.004	0.013	0.011
	Random Partner	0.568	0.414	0.040	0.015	0.070	0.004	0.012	0.013
Black	Truth	0.276	0.044	0.191	0.003	0.018	0.001	0.009	0.010
	Fully Random	0.498	0.317	0.063	0.018	0.071	0.004	0.013	0.011
	Random Partner	0.276	0.201	0.019	0.007	0.034	0.002	0.006	0.006
Asian	Truth	0.360	0.113	0.008	0.185	0.024	0.001	0.021	0.009
	Fully Random	0.498	0.317	0.063	0.018	0.071	0.004	0.013	0.011
	Random Partner	0.360	0.262	0.025	0.010	0.045	0.002	0.008	0.008
Hispanic	Truth	0.430	0.123	0.013	0.007	0.241	0.002	0.009	0.035
	Fully Random	0.498	0.317	0.063	0.018	0.071	0.004	0.013	0.011
	Random Partner	0.430	0.313	0.030	0.011	0.053	0.003	0.009	0.010
AIAN	Truth	0.426	0.194	0.022	0.008	0.044	0.130	0.015	0.012
	Fully Random	0.498	0.317	0.063	0.018	0.071	0.004	0.013	0.011
	Random Partner	0.426	0.310	0.030	0.011	0.053	0.003	0.009	0.009
Other	Truth	0.428	0.241	0.040	0.030	0.048	0.005	0.058	0.007
	Fully Random	0.498	0.317	0.063	0.018	0.071	0.004	0.013	0.011
	Random Partner	0.428	0.312	0.030	0.011	0.053	0.003	0.009	0.010

Notes: This table presents observed rates of interracial cohabitation, measured at age 30, alongside two benchmarks. Cohabitation is defined using data from the ACS. Specifically, we include individuals in the main analysis sample who are surveyed by the ACS in the year they turn 30. An individual is classified as cohabiting if they respond to the survey as the primary respondent and list a married spouse or a non-married partner in their household as an additional respondent, or if they are listed as the spouse or partner of the primary respondent. The fully random benchmark is calculated as the overall cohabitation rate multiplied by the fraction of the population from each race group. The random partner benchmark is the actual cohabitation rate for each race group multiplied by the fraction of the cohabiting population belonging to each race group. These benchmarks are constructed analogously to those in Table IV, substituting cohabitation for marriage.

TABLE A.3: Marriage Outcomes by Class Relative to Random Matching Benchmark
Sample: 1982 Birth Cohort

Parent Income Group		Fraction Married (1)	Fraction with Spouse from Parent Income Group at Age 37: Bottom 25% (2)	Quartile 2 (3)	Quartile 3 (4)	Top 25% (5)	Out of Sample (6)
Bottom 25%	Truth	0.297	0.053	0.056	0.053	0.039	0.097
	Fully Random	0.451	0.088	0.088	0.088	0.088	0.101
	Random Spouse	0.297	0.038	0.051	0.065	0.076	0.066
Quartile 2	Truth	0.402	0.056	0.078	0.090	0.077	0.100
	Fully Random	0.451	0.088	0.088	0.088	0.088	0.101
	Random Spouse	0.402	0.051	0.069	0.088	0.103	0.090
Quartile 3	Truth	0.511	0.053	0.090	0.131	0.135	0.101
	Fully Random	0.451	0.088	0.088	0.088	0.088	0.101
	Random Spouse	0.511	0.065	0.088	0.113	0.131	0.114
Top 25%	Truth	0.594	0.040	0.077	0.136	0.236	0.105
	Fully Random	0.451	0.088	0.088	0.088	0.088	0.101
	Random Spouse	0.594	0.076	0.103	0.131	0.152	0.133

Notes: This table presents statistics on rates of interclass marriage, measured at age 37. The sample includes individuals from the 1982 birth cohort—the oldest cohort in the main analysis sample—who were 37 in 2019, the final year of available tax data. For details on the construction of these statistics, see the notes to Table III.

TABLE A.4: Marriage Outcomes by Race Relative to Random Matching Benchmark
Sample: 1982 Birth Cohort

Race Group		Fraction Married (1)	Fraction with Spouse from Race Group at Age 37: White (2)	Black (3)	Asian (4)	Hispanic (5)	AIAN (6)	Other (7)	Out of Sample (8)
White	Truth	0.546	0.482	0.006	0.008	0.025	0.002	0.008	0.014
	Fully Random	0.463	0.292	0.060	0.015	0.060	0.004	0.011	0.020
	Random Spouse	0.546	0.406	0.025	0.020	0.057	0.003	0.011	0.024
Black	Truth	0.161	0.028	0.104	0.002	0.010	0.000	0.005	0.012
	Fully Random	0.463	0.292	0.060	0.015	0.060	0.004	0.011	0.020
	Random Spouse	0.161	0.120	0.007	0.006	0.017	0.001	0.003	0.007
Asian	Truth	0.513	0.123	0.009	0.267	0.027	0.001	0.024	0.064
	Fully Random	0.463	0.292	0.060	0.015	0.060	0.004	0.011	0.020
	Random Spouse	0.513	0.382	0.023	0.019	0.053	0.003	0.011	0.022
Hispanic	Truth	0.371	0.112	0.010	0.007	0.187	0.002	0.008	0.046
	Fully Random	0.463	0.292	0.060	0.015	0.060	0.004	0.011	0.020
	Random Spouse	0.371	0.276	0.017	0.014	0.039	0.002	0.008	0.016
AIAN	Truth	0.304	0.154	0.009	0.004	0.025	0.087	0.012	0.013
	Fully Random	0.463	0.292	0.060	0.015	0.060	0.004	0.011	0.020
	Random Spouse	0.304	0.226	0.014	0.011	0.032	0.002	0.006	0.013
Other	Truth	0.408	0.211	0.028	0.035	0.042	0.004	0.062	0.026
	Fully Random	0.463	0.292	0.060	0.015	0.060	0.004	0.011	0.020
	Random Spouse	0.408	0.303	0.019	0.015	0.043	0.002	0.008	0.018

Notes: This table presents statistics on rates of interracial marriage, measured at age 37. The sample includes individuals from the 1982 birth cohort—the oldest cohort in the main analysis sample—who were 37 in 2019, the final year of available tax data. For details on the construction of these statistics, see the notes to Table IV.

TABLE A.5: Marriage Outcomes by Class and Sex

Parent Income Group		Fraction Married (1)	Fraction with Spouse from Parent Income Group at Age 30: Bottom 25% (2)	Quartile 2 (3)	Quartile 3 (4)	Top 25% (5)	Out of Sample (6)
Bottom 25%	Pooled	0.239	0.050	0.052	0.047	0.031	0.059
	Male	0.217	0.048	0.050	0.046	0.029	0.043
	Female	0.262	0.052	0.053	0.048	0.033	0.076
Quartile 2	Pooled	0.317	0.052	0.070	0.079	0.059	0.056
	Male	0.286	0.050	0.067	0.076	0.055	0.038
	Female	0.349	0.055	0.074	0.082	0.063	0.075
Quartile 3	Pooled	0.405	0.048	0.080	0.117	0.111	0.050
	Male	0.363	0.044	0.074	0.110	0.104	0.031
	Female	0.449	0.051	0.086	0.124	0.118	0.070
Top 25%	Pooled	0.437	0.031	0.060	0.111	0.190	0.045
	Male	0.388	0.029	0.054	0.102	0.174	0.028
	Female	0.489	0.034	0.066	0.121	0.205	0.063

Notes: This table presents statistics on rates of interclass marriage separately by sex. For details on construction, see the notes to Table III.

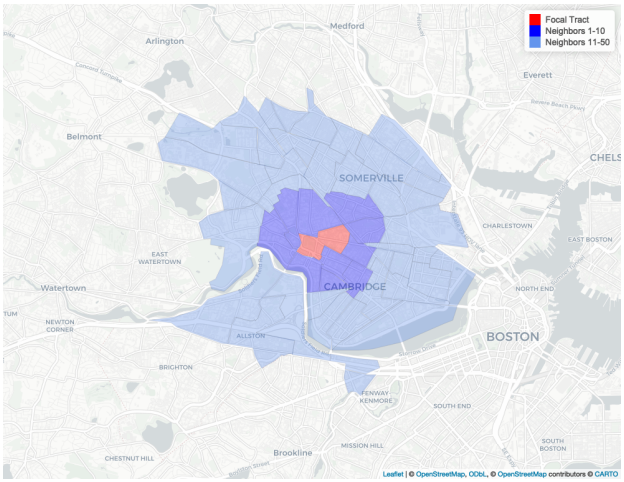
TABLE A.6: Marriage Outcomes by Race and Sex

Race Group		Fraction Married (1)	Fraction with Spouse from Race Group at Age 30: White (2)	Black (3)	Asian (4)	Hispanic (5)	AIAN (6)	Other (7)	Out of Sample (8)
	Pooled	0.433	0.382	0.005	0.005	0.020	0.002	0.007	0.012
White	Male	0.386	0.339	0.003	0.006	0.019	0.002	0.007	0.011
	Female	0.482	0.427	0.007	0.004	0.022	0.002	0.007	0.012
	Pooled	0.117	0.021	0.072	0.001	0.008	0.000	0.004	0.009
Black	Male	0.117	0.026	0.066	0.002	0.010	0.000	0.005	0.008
	Female	0.116	0.017	0.079	0.001	0.006	0.000	0.003	0.010
	Pooled	0.316	0.077	0.006	0.148	0.017	0.001	0.015	0.052
Asian	Male	0.267	0.051	0.003	0.125	0.014	0.001	0.011	0.062
	Female	0.369	0.105	0.008	0.173	0.021	0.001	0.019	0.042
	Pooled	0.289	0.083	0.008	0.005	0.144	0.001	0.006	0.043
Hispanic	Male	0.258	0.073	0.005	0.005	0.131	0.001	0.005	0.038
	Female	0.320	0.093	0.011	0.004	0.156	0.001	0.006	0.048
	Pooled	0.256	0.132	0.007	0.003	0.020	0.072	0.010	0.012
AIAN	Male	0.227	0.119	0.004	0.003	0.017	0.065	0.009	0.010
	Female	0.286	0.146	0.009	0.003	0.023	0.079	0.011	0.014
	Pooled	0.299	0.159	0.021	0.021	0.031	0.003	0.043	0.021
Other	Male	0.269	0.141	0.014	0.022	0.030	0.003	0.039	0.021
	Female	0.328	0.176	0.029	0.020	0.033	0.003	0.046	0.020

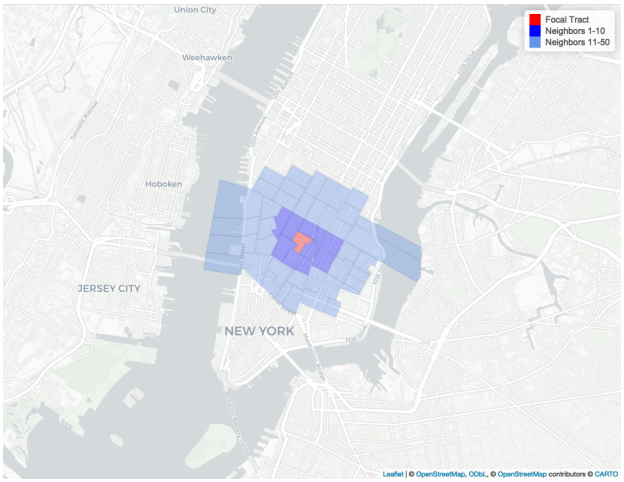
Notes: This table presents statistics on rates of interracial marriage separately by sex. For details on construction, see the notes to Table IV.

FIGURE A.1: Examples of Nearest 50 Census Tracts

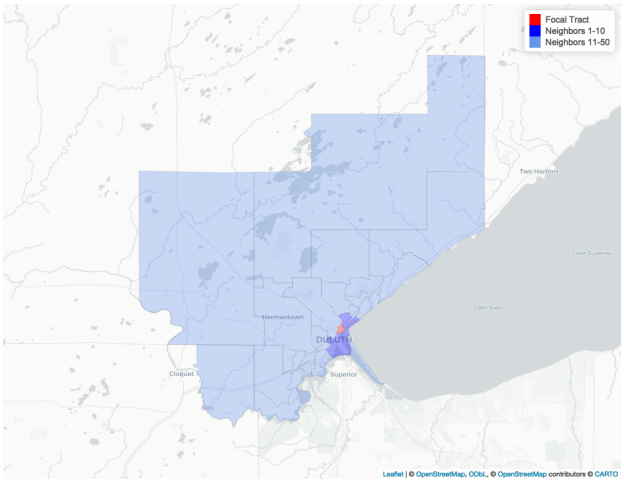
A: Cambridge, MA



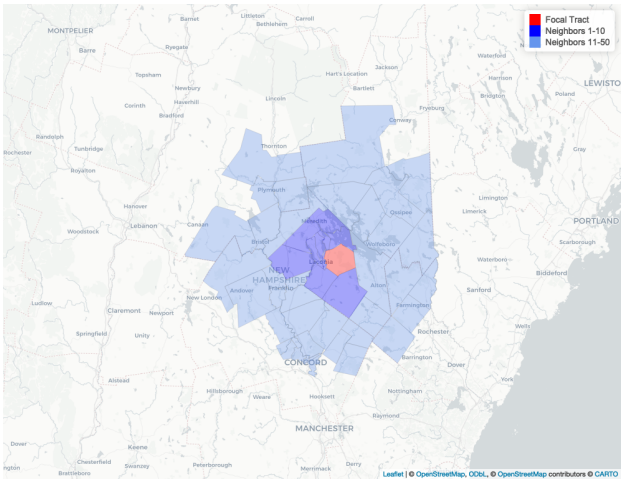
B: New York City, NY



C: Duluth, MN



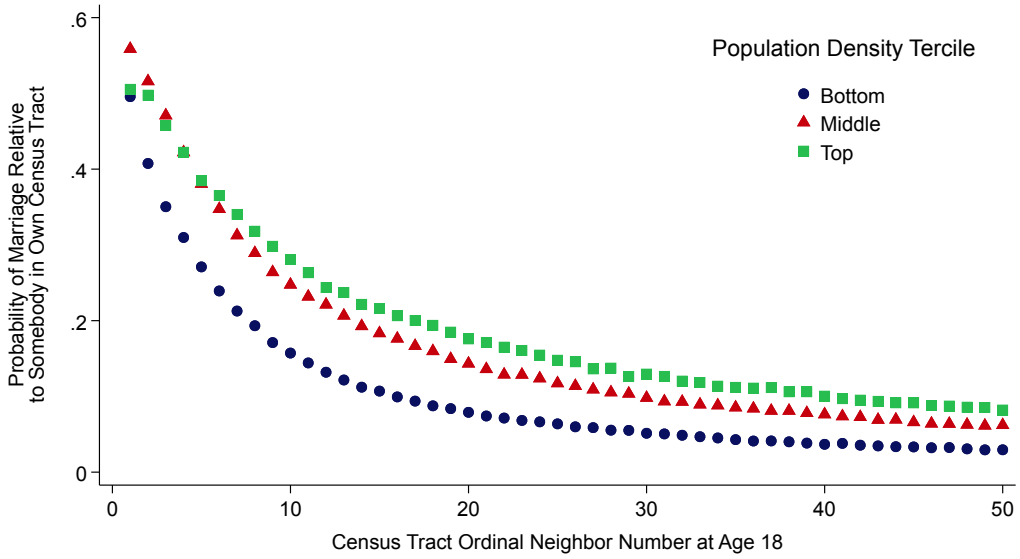
D: Gilford, NH



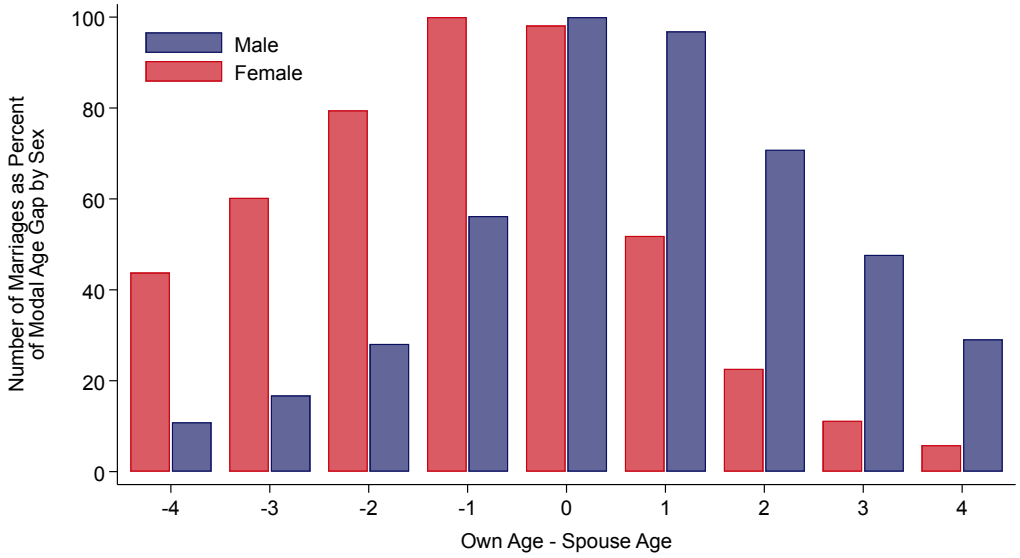
Notes: This figure depicts the nearest 50 Census tract neighbors of a particular focal tract in four U.S. cities. In each case, the neighbors are defined using the distance between Census tract centroids. The focal tract is shown in red, neighbors 1-10 in dark blue, and neighbors 11-50 in light blue. The 2010 FIPS codes for the focal tracts are 25017353700 (Cambridge), 36061005900 (New York), 27137001800 (Duluth), 33001966402 (Gilford).

FIGURE A.2: Distance and Age Weights Used to Construct Exposure Measure

A: Marriage Probability Decay by Distance and Population Density



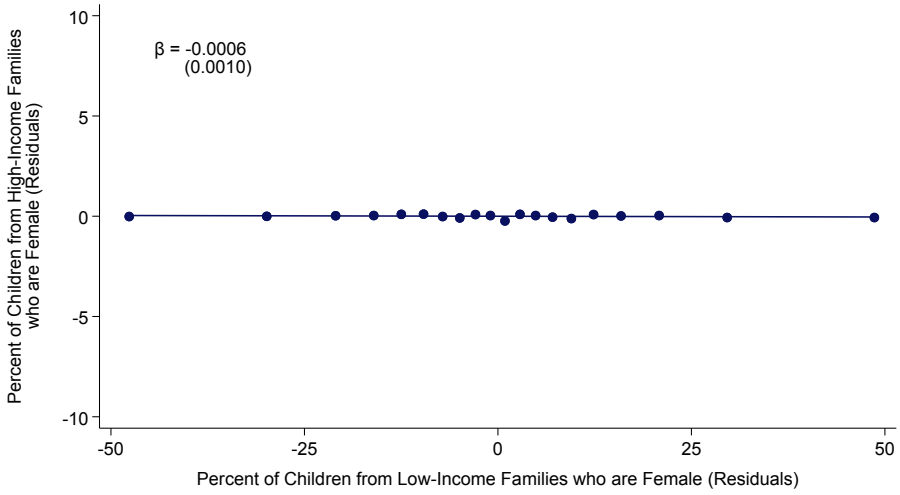
B: Difference in Age From Spouse by Sex



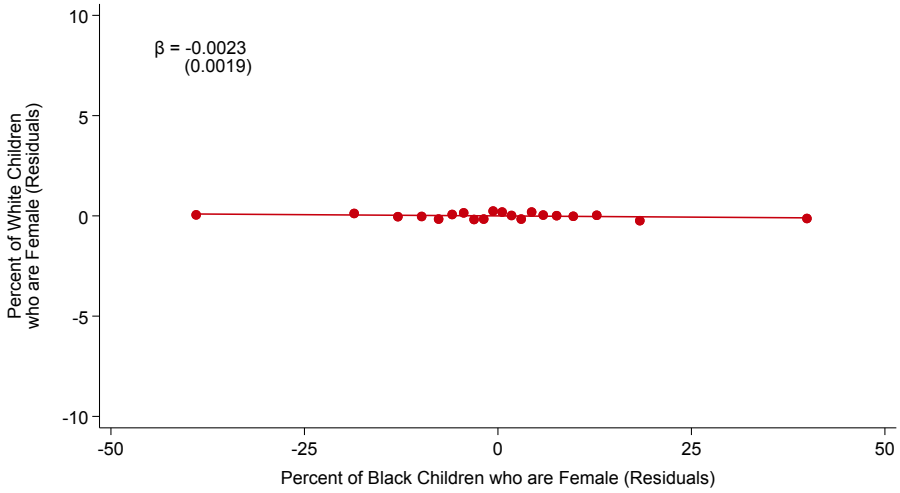
Notes: This figure presents the weights used to construct the exposure measure described in Section III.C. In Panel *A*, we present the distance weights, separately for each tercile of population density, measured in 2000. A pooled version of this figure is shown in Figure II Panel *B*; see figure notes for additional details. Panel *B* presents the frequency of spouse age gaps for married males and females. For each group, the modal age gap is assigned a value of 100. For all other age gaps between -4 and 4, the height of the bar reflects the frequency of the age gap as a percentage of the frequency for the modal age gap.

FIGURE A.3: Within-Neighborhood Correlation of Sex Ratios Across Groups

A: Individuals from Families in the Bottom 25% and Top 25%

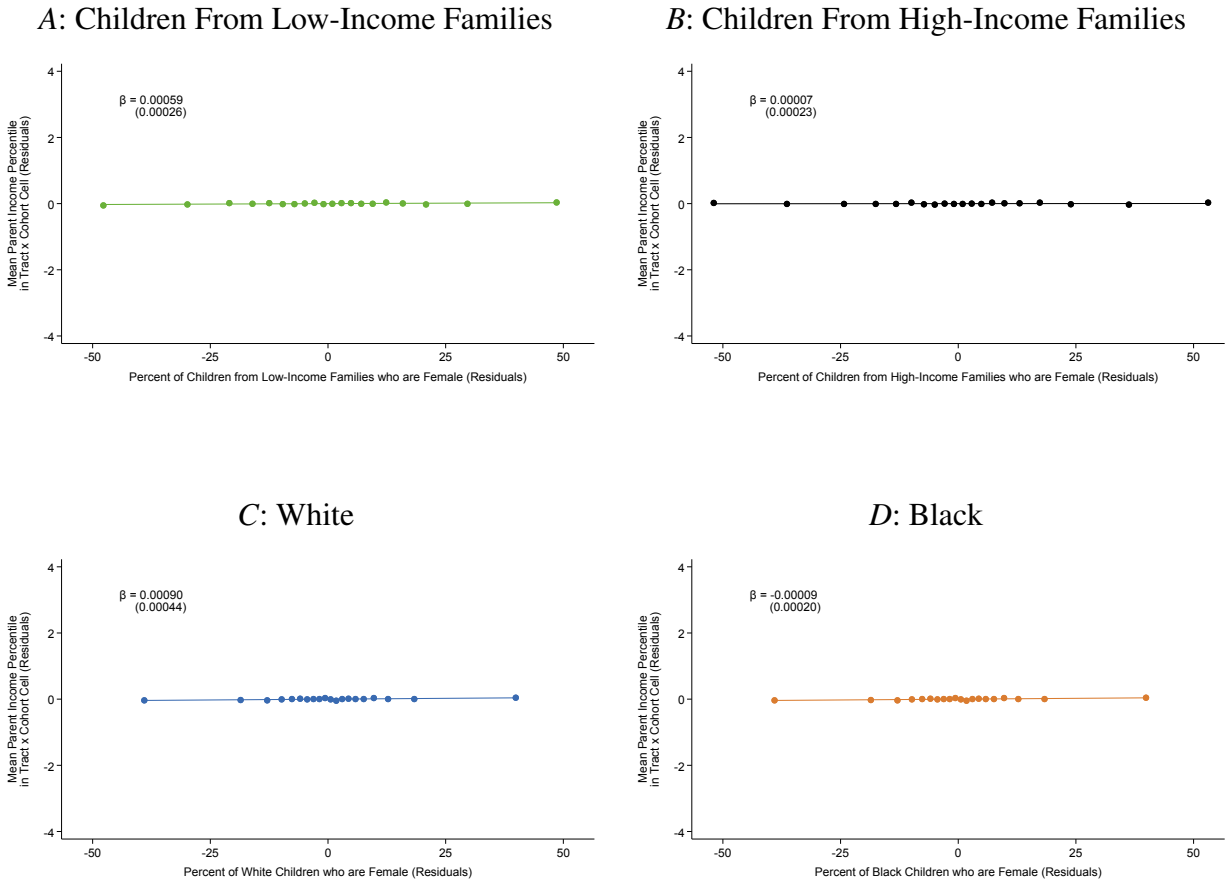


B: White and Black



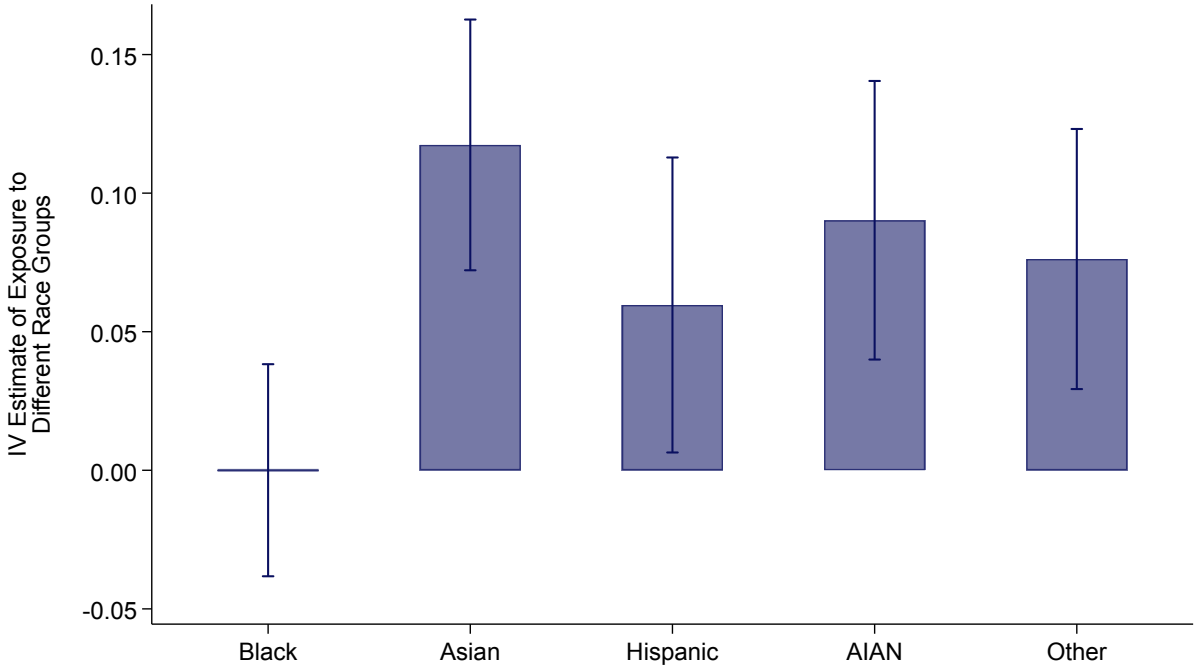
Notes: This figure presents binned scatter plots of the within-neighborhood relationship between the sex-ratios across different class (Panel *A*) and race (Panel *B*) groups. For each Census tract by cohort by income or race group, we calculate the fraction of individuals in the cell who are female. We demean this sex-ratio within Census tract. To generate the plot, we construct 20 bins of the residualized low-income (Black) sex-ratio and plot the mean high-income (white) sex ratio in each bin against the mean low-income (Black) sex-ratio in the bin. The coefficients on the plot are from regressions in a tract by cohort dataset that includes tract fixed effects, with standard errors clustered at the county level.

FIGURE A.4: Tract × Cohort Sex Ratio Balance on Mean Parental Income



Notes: This figure presents binned scatter plots of the within- neighborhood relationship between sex-ratios for specific subgroups and mean parent income rank. We show results for individuals from low-income families (Panel *A*), high income families (Panel *B*), white (Panel *C*), and Black individuals (Panel *D*). In each case, we calculate the fraction of individuals who are female in a Census tract by cohort by family income or race group. We also calculate the mean parent income rank during childhood for that group, using the definition in Section II. To generate the figure, we first construct percentile bins of the number of individuals in the cell. We residualize the sex ratios and the mean parent ranks on Census tract fixed effects and fixed effects for the size bins. To generate the figure, we construct 20 bins of the residualized sex ratios and plot the mean residualized parent income rank against the mean residualized sex ratio in each bin. We report the coefficient from regressions in a tract by cohort dataset of mean parent rank sex ratios controlling for tract and size-bin fixed effects, with standard errors clustered at the county level.

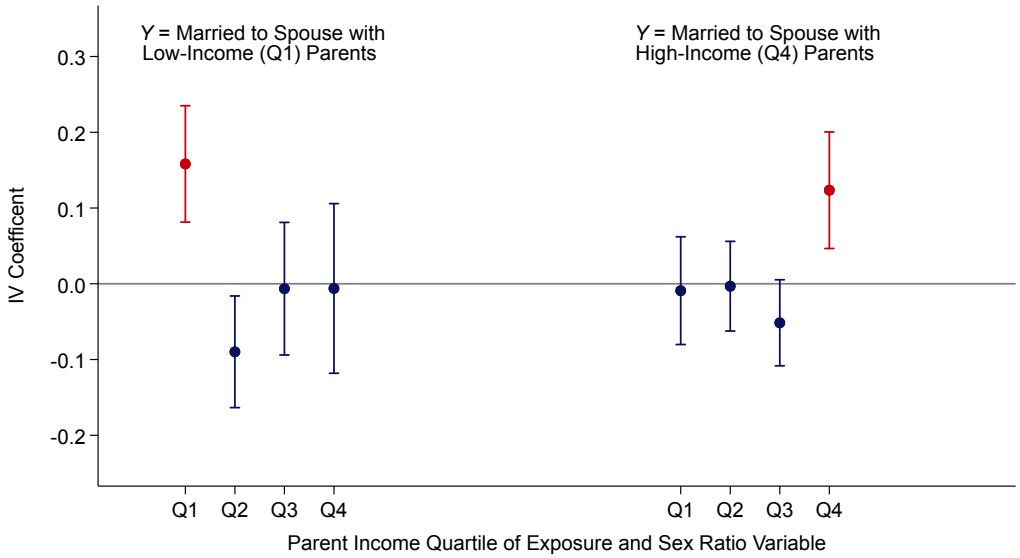
FIGURE A.5: IV Estimates of Cross-Race Exposure for White Individuals



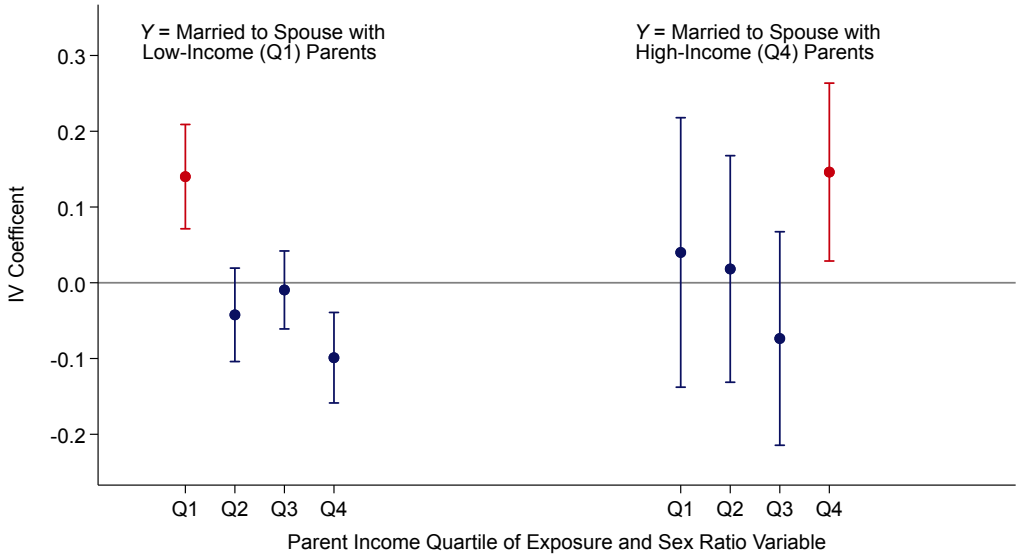
Notes: This figure summarizes the IV results for interracial marriage among White individuals. Each bar represents the effect of a 100pp increase in market tightness for a specific racial group on the probability that a White individual marries someone from that group. For details on the specification procedure, see Section IV, B, and the notes to Table VI.

FIGURE A.6: Impact of Race and Class Exposure on Spouse Characteristics

A: Children with Low-Income (Q1) Parents

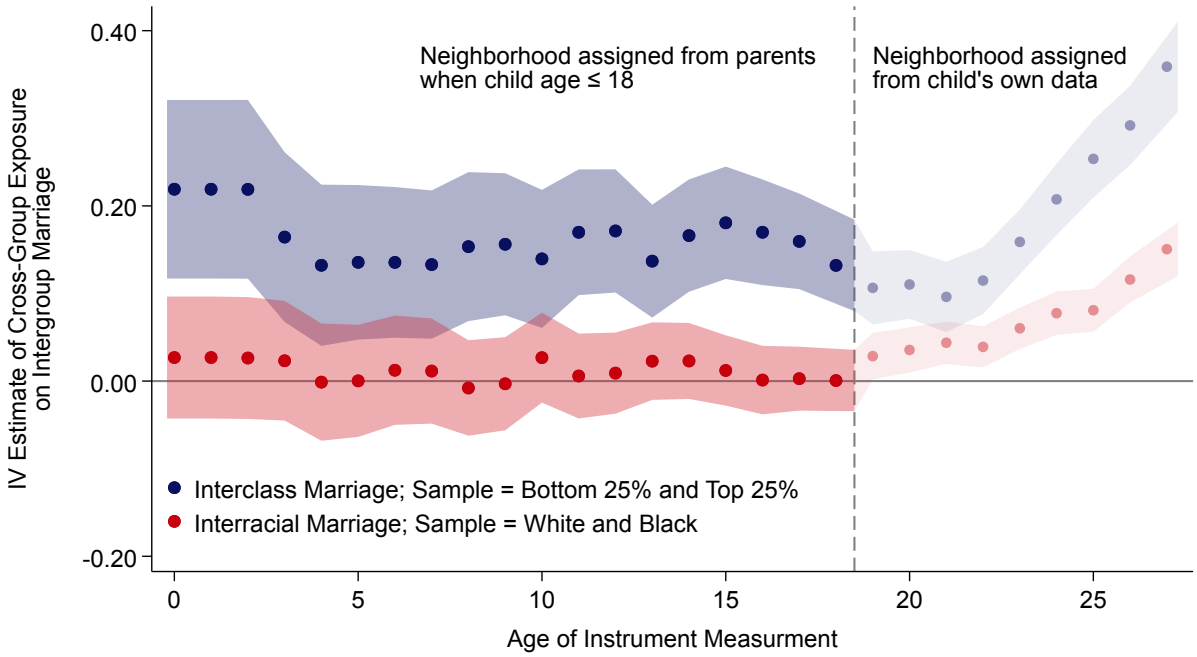


B: Children with High-Income (Q4) Parents



Notes: This figure presents a set of multivariate IV regressions for interclass marriage outcomes, analyzed separately for individuals from low parent income families (Panel *A*) and high parent income families (Panel *B*). In each case, the marriage outcome—indicated above the points—is regressed on four endogenous exposure measures. These exposure measures are instrumented using four sex ratio instruments, one for each parent income quartile. The regression controls are consistent with those used in the baseline IV specification; see Section IV.B for additional details.

FIGURE A.7: IV Estimates by Age of Instrument Measurement



Notes: This figure presents IV coefficients from separate regressions where the instrument is measured at different ages throughout childhood, analyzed separately for class and race. The outcome in each regression is the fraction of individuals in an interclass (or interracial) marriage at age 30. The endogenous exposure variable is defined as the difference between the fraction of opposite-sex neighbors (measured over ages 18–27) who belong to the other group and the fraction of own-sex neighbors who belong to the other group. The instrument is the fraction of neighbors from the opposite group who were opposite-sex, with neighbors measured at each age of childhood, from birth to age 27. Neighborhood assignments are based on parent data through age 18 and child data after age 18. Coefficients where the instrument is measured at age 18 correspond to the baseline results presented in Figure VIII. See Section IV.B for additional details on the estimation procedure.

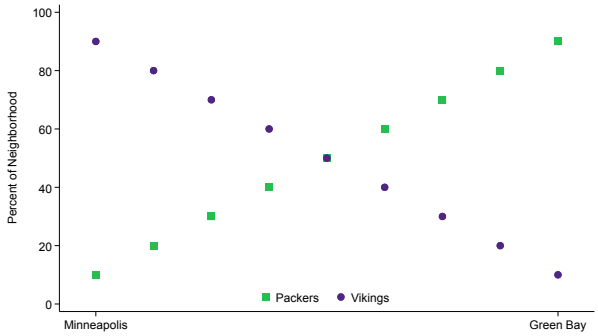
FIGURE A.8: Within vs. Outside Childhood Neighborhood Marriage Bottom and Top Quartile



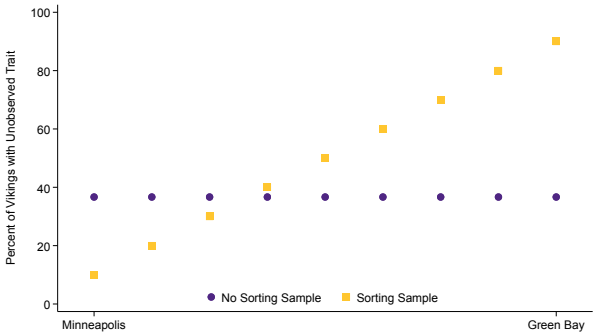
Notes: This figure examines whether variation in neighborhood-level class-group exposure affects the likelihood of marrying someone from within the neighborhood versus outside the neighborhood. The left half of the figure shows estimates for own-class exposure and the likelihood of having an own-class spouse, while the right half focuses on opposite-class exposure and its effect on the likelihood of having an opposite-class spouse. The sample includes individuals from families in the bottom and top quartiles of parental income. Each half of the figure presents coefficients from four separate regressions. The right-hand-side variables are consistent with the baseline IV specification; see Section IV.B for additional details on the estimation procedure. The outcome variables measure the likelihood of having a spouse in the relevant class group either within or outside the 50 or 100 nearest neighbor tracts to an individual's age 18 neighborhood.

FIGURE A.9: Simulations With and Without Sorting on Unobserved Trait

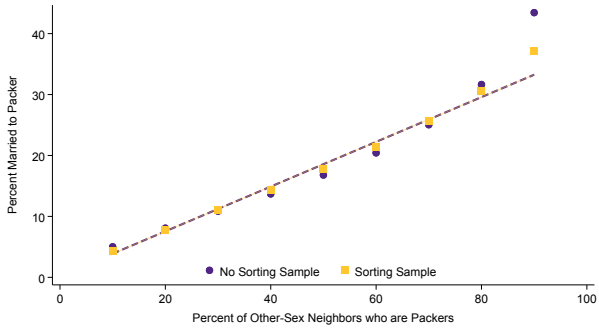
A: Types by Neighborhood



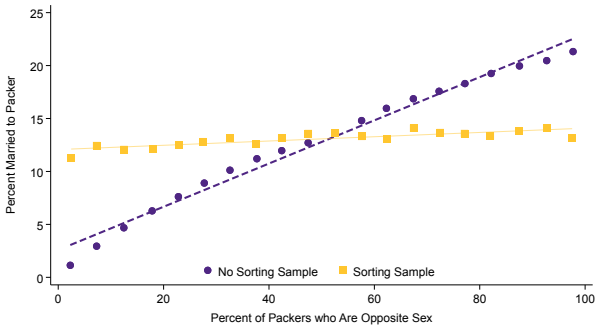
B: Sorting on Unobservable by Neighborhood



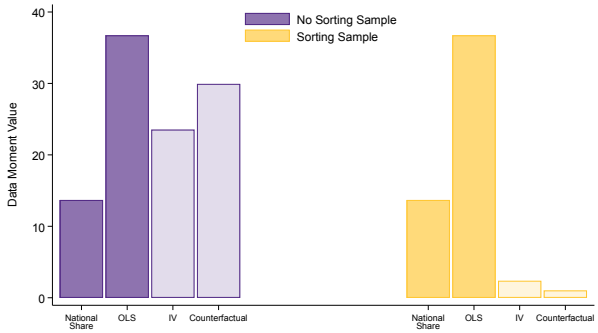
C: OLS



D: Reduced Form

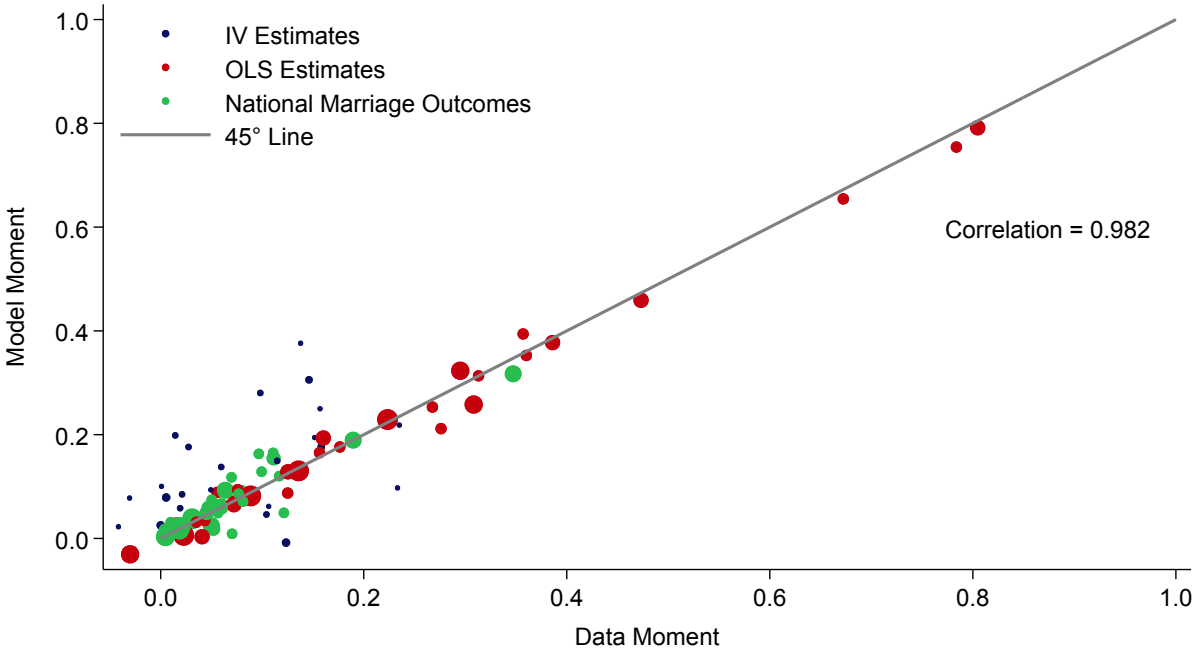


E: Data Moments by Sorting Pattern



Notes: This figure presents a simulation from a simplified version of the spatial matching model discussed in Section V. In the simulation, described in detail in Section E, there are two types, Packers and Vikings, and nine neighborhoods, from Minneapolis to Green Bay. Panel *A* shows the fraction of neighborhood residents who are Packers and Vikings as one moves from Minneapolis to Green Bay. Panel *B* shows the share of Vikings, by neighborhood, who have the unobserved trait in two contrasting cases, a sorting and no-sorting case. Panel *C* shows the OLS relationship between share of Vikings married to a Packer and neighborhood exposure to packers, analogous to Figure V. Panel *D* shows the reduced form relationship between share of Vikings married to a Packer the share of Packers in their neighborhood who are opposite-sex, analogous to Figure VII. Panel *E* summarises the data moments in the sorting and no-sorting case and also shows the counterfactual effects of a 10pp increase in neighborhood exposure to Packers on the percent of Vikings married to Packers.

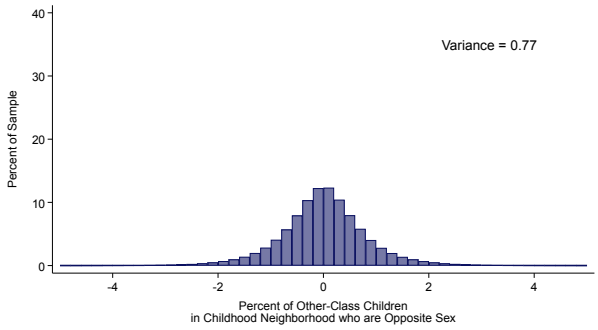
FIGURE A.10: Simulated Model Moments versus Observed Data Moments



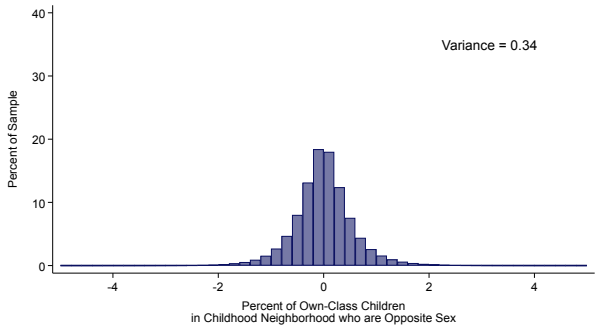
Notes: This figure shows the relationship between the data moments used in estimating the matching model, as described in Section V, and the corresponding simulated moments at the model equilibrium. The size of each point is scaled according to the weight assigned to the data moment during estimation. The weighted correlation between the model and data moments is reported. Details on the model setup, weighting, and estimation are provided in Section V and Appendix D.

FIGURE A.11: Distribution of Instrument Within vs. Across Groups

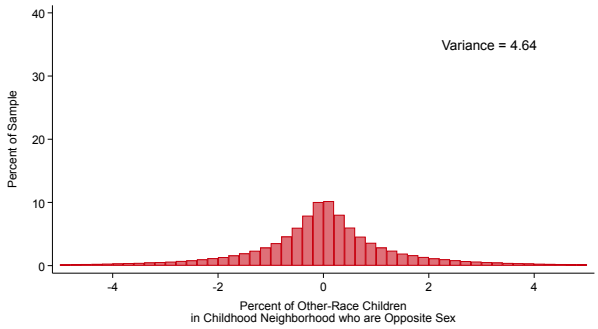
A: *Z* Distribution Across Class
Sample: Bottom 25% and Top 25%



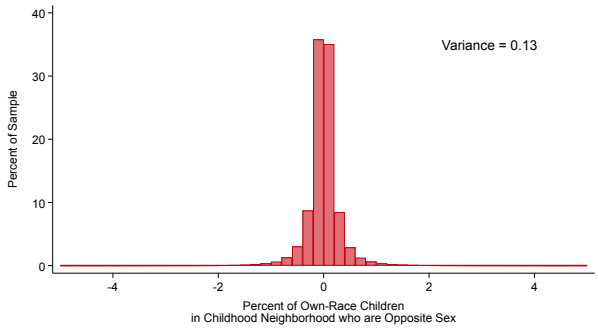
B: *Z* Distribution Own Class
Sample: Bottom 25% and Top 25%



C: *Z* Distribution Across Race
Sample: White and Black



D: *Z* Distribution Own Race
Sample: White and Black



Notes: This figure shows the distribution of four versions of the sex-ratio instrument. In Panels *A* and *B*, the instrument is constructed using a sample of individuals from the bottom and top 25% of the parent income distribution. In Panel *A*, the instrument represents the fraction of neighbors (measured at age 18) from the other class group who are of the opposite sex. In Panel *B*, it represents the fraction of own-class neighbors who are of the opposite sex. In Panels *C* and *D*, the instrument is constructed using a sample of White and Black individuals. In Panel *C*, the instrument is the fraction of neighbors from the other race group (e.g., Black neighbors for White individuals) who are of the opposite sex. In Panel *D*, it is the fraction of own-race neighbors who are of the opposite sex. The histograms are analogous to those presented in Figure VII. Panels *A* and *C* replicate the histograms shown in Figure VII Panels *A* and *C*, respectively. See Figure VII notes for further details.