NBER WORKING PAPER SERIES

WHO MARRIES WHOM? THE ROLE OF SEGREGATION BY RACE AND CLASS

Benjamin Goldman
Jamie Gracie
Sonya Porter

Working Paper 35140
http://www.nber.org/papers/w35140

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
April 2026

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, CBDRB-FY24-0106, CBDRB-FY25-0334). 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 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, 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 the editor, four anonymous referees, and 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. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Who Marries Whom? The Role of Segregation by Race and Class
Benjamin Goldman, Jamie Gracie, and Sonya Porter
NBER Working Paper No. 35140
April 2026
JEL No. D31, J12, J15, J62, R23

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.

Benjamin Goldman
Cornell University
Economics and Public Policy
and NBER
bgoldman@cornell.edu

Sonya Porter
U.S. Census Bureau
sonya.rastogi.porter@census.gov

Jamie Gracie
Harvard University
jgracie@g.harvard.edu

A online appendix is available at http://www.nber.org/data-appendix/w35140
A Non-Technical Summary is available at https://benjamindgoldman.s3.us-east-2.amazonaws.com/papers/marria
A County-Level Marriage Data is available at https://www.benjamindgoldman.com/data/marriage_homophily/in

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; Chadwick and Solon, 2002; Fryer, 2007).

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. Alternatively, people may be open to marrying across group lines, but a lack of *exposure* to other groups prevents such marriages from forming.[1] If exposure is a key driver of marital homophily, policies that increase opportunities for cross-group interaction—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

[1]By *exposure*, we mean the presence of a group in a shared setting, such as a neighborhood, that creates opportunities for cross-group interaction. We use *exposure*, *residential segregation*, and *geographic proximity* interchangeably, though exposure need not imply actual contact. We focus on exposure because it is more directly shaped by policy, and contact conditional on exposure may reflect underlying preferences. We use *preferences* to refer to the propensity for within-group marriage conditional on exposure; these may be shaped by individual, familial, or community norms and life experiences (Allport, Clark and Pettigrew, 1954; Barlow et al., 2012; Corno, La Ferrara and Burns, 2022).

and Solon, 2002). 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 (Chetty et al., 2020). Our baseline sample consists of individuals born between 1982 and 1989, measuring marital status at age 30, when 35.0% 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 much lower than these benchmarks—there would be at least 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 67.6% 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.

Because neighbors make up a key part of the marriage market, we next examine how likely individuals are to live near people from different race and class backgrounds. The U.S. is well known for its residential segregation by race and income (e.g., Massey and Denton, 1988; Reardon

[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 and 23.9% are Black.

[3]Measuring class-based marital homophily requires longitudinal data tracking individuals from childhood to adulthood (e.g., Charles, Hurst and Killewald, 2013; Ermisch, Francesconi and Siedler, 2006), 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 (Fryer, 2007; Qian and Lichter, 2011).

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

and Bischoff, 2011). In our analysis, we measure segregation—or neighborhood exposure to other groups—based on the composition of the 50 nearest census tracts. As young adults (ages 18 to 27), individuals tend to live in highly segregated neighborhoods. On average, those from low-income families live in neighborhoods where 30.8% of nearby individuals are also from low-income families, compared to only 19.2% from high-income families, while high-income children live near far more high-income than low-income families (34.2% versus 19.2%). Racial exposure is also very uneven: on average, White children live in neighborhoods that are 75.6% White and just 10.8% Black, whereas Black children live near nearly equal shares of Black and White neighbors (40.5% and 43.9%, respectively).

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 is associated with 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. 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 share 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 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 a given group. Our tightness measure captures whether there is a relative surplus of opposite-sex members within a given race or class group, but it is distinct from what we refer to as market "thickness"—the group's overall share in a neighborhood. We return to this distinction after presenting our baseline empirical estimates.

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).[5] 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. In a combined sample of individuals from low- and high-income families, we find that a 10pp increase in market tightness across class lines raises the probability of an interclass marriage by 1.3pp. The 95% confidence interval allows us to rule out effects smaller than 0.8pp. 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 and Black individuals, a 10pp increase in market tightness across race lines raises the probability of interracial marriage by only 0.006pp. The 95% confidence interval allows us to rule out effects larger than 0.356pp. 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 a Bernoulli random variable. This pattern would not hold if parents sorted into neighborhoods based on the sex composition of children in nearby cohorts, 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

[5] 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.

across cohorts between siblings. Finally, we present evidence that our instrument influences intergroup marriage by changing the availability of potential partners, rather than by shaping preferences through exposure to opposite-sex individuals from particular groups.

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—which capture variation in market tightness driven by sex imbalances within a group—into estimates based on market thickness or 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.[6] 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 (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 3.0pp. A comparable increase in cross-race exposure increases interracial marriage by 1.0pp, 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

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

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 a no-segregation counterfactual, interclass marriage rates rise from 3.1% to 4.6%, closing 35.2% (S.E.=5.4%) of the gap between observed interclass marriage rates and the no-homophily benchmark, where all class groups marry at the same rates and choose spouses randomly. Meanwhile, rates of White-Black interracial marriage rise from 2.1% to 3.0%, closing only 4.3% (S.E.=0.6%) of the gap between observed interracial marriage rates and the no-homophily benchmark. These results suggest that limited 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. Indeed, we find that the causal effect of exposure on interclass marriage is strongest in areas with higher baseline levels of cross-class social interaction (Chetty et al., 2022). 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., 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; Fisman et al., 2008; Hitsch, Hortaçsu and Ariely, 2010). 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.[7] 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

[7] This literature includes studies on the impact of marital homophily within (Clark and Cummins, 2022; Eika, Mogstad and Zafar, 2019; Schwartz, 2010) and across generations (Bailey and Lin, 2024; Buckles et al., 2023; Chadwick and Solon, 2002; Ermisch, Francesconi and Siedler, 2006; Eriksson et al., 2023; Mare, 2000; Olivetti et al., 2024).

of "meeting opportunities" (e.g., Belot and Francesconi, 2013; Kalmijn and Flap, 2001; Wen and Mao, 2025), with notable recent advancements by Ciscato (2024).[8] 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 inequality and economic opportunity (Chetty et al., 2022; Fischer and Mattson, 2009; Putnam, 2016). 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; Cutler and Glaeser, 1997; Derenoncourt, 2022; Fernández and Rogerson, 2001; Kremer, 1997).

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

I 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 (US Census Bureau, n.d.*c*,n,n). 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).

[8]Ciscato (2024) uses a dynamic matching model that separately identifies preferences and search frictions from descriptive data on relationship durations, concluding that high rates of same-race marriage are largely driven by search frictions. Although our findings differ, they are not formally inconsistent, as the studies focus on different types of marriages and consider different scopes of search frictions. Ciscato (2024) examines search frictions broadly defined, while we focus on residential segregation—one important and policy-relevant friction.

I.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 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.

I.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

[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.

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

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.

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 Alaska 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 unmarried 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 with 12.9% imputed from nearby years.

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 as reported on Form 1040. Individual income is defined as wage income reported on the W-2 plus self-employment and other non-wage income reported on Form 1040; for married couples, this non-wage income is split evenly between spouses. 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 (Goldman, Gracie and Porter, 2026).

I.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 95.1% 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.[11]

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 (Chetty et al., 2020). 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 (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 (Tolbert and Sizer, 1996).

[11] We define income quartiles at the national level to capture interclass marriages between individuals from low- and high-income families, regardless of location. A regional definition would obscure any effects of cross-region segregation on intergroup marriage and make the definition of interclass marriage location-dependent.

II Descriptive Facts on Marriage, Homophily, and Neighborhoods

II.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.

To quantify homophily, we compare observed interclass marriage rates to two benchmarks in which individuals select spouses without regard to class. These benchmarks, shown in Figure I Panel A, 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.[12] In reality, interclass marriage is far less common, indicating substantial marital homophily—interclass marriage would be 2–3 times more prevalent if spouses were selected without regard to class background (see Table A.1 for benchmarks and rates across all four parent income quartiles).

Figure I Panel B presents the corresponding statistics for interracial marriage. 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 holding racial differences in marriage rates constant, interracial marriage would be more than four times higher if individuals chose spouses

[12] These benchmarks are computed allowing for a separate category of spouses whose parental income cannot be linked (approximately 15% of married individuals), so each observed parent income quartile represents less than 25% of the population.

from the pool of married individuals without regard to race (see Table A.2 for marriage outcomes and benchmarks for all groups).

We next assess whether the low rates of interracial and interclass marriage are sensitive to how we define marriage. In recent decades, Americans have increasingly delayed marriage or chosen cohabitation as an alternative (Kennedy and Bumpass, 2008; Lundberg, Pollak and Stearns, 2016). Measures of homophily may therefore vary depending on the age at which marriage is measured and whether cohabiting relationships are included—particularly if individuals who marry later or cohabit are more likely to form interclass or interracial unions.

To address these concerns, we examine individuals in the ACS at age 30, including both married and cohabiting couples. We also analyze marriage outcomes for the 1982 birth cohort at age 37 using 2019 tax data. In both cases, we find higher overall rates of partnership compared to our primary estimates. However, the proportional gaps between observed rates of interclass and interracial marriage and the no-homophily benchmarks remain similar. For example, under the random spouse (or cohabiting partner) benchmark, the share of Black individuals with a White partner is 4.6 times higher than the observed share, compared to 4.1 times higher in our baseline estimates. Full statistics for all race and class groups are reported in Tables A.3 and A.4 (cohabitation), and Tables A.5 and A.6 (marriage at age 37).

Rates of interclass and interracial marriage are low for both sexes. Among individuals from low-income families, 2.9% of men and 3.3% of women have a spouse from a high-income family. While there is modest evidence of hypogamy—i.e., women being slightly more likely to "marry up" the class distribution—the 0.4pp difference is small, and interclass marriage is rare for both sexes (see Table A.7 for comparable statistics across all class groups). Among Black individuals, 2.6% of men and 1.7% of women have a White spouse (see Table A.8). For the remainder of our analyses, we focus on the baseline estimates at age 30, pooling both sexes.

II.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; 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. 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. 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 (see Figure A.1 for maps illustrating these and other examples of this definition). 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, only 0.07% of marriages would occur between individuals who lived within this radius of each other.

Although these neighborhoods encompass only a small share of the overall population, we find that a substantial fraction of marriages occur between individuals who had lived within 50 census tracts of each other 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, 67.6% 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, 58.2% 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), 45.1% 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 III presents these statistics across various geographic definitions, ages, and years relative to marriage (see Table A.9 for comparable statistics that include co-residing couples). Remarkably, 23.5% of marriages are between two people who were born within 50 census tracts of each other. Across different ages, the share of couples who lived within 50 census tracts of each other is similar to the share who lived within 10 miles or in the same county—alternative ways to define neighborhoods. Despite some variation by race and class, we find that marrying a partner who has lived nearby is common across all groups (see Figure A.2).

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 combined probability of marrying someone from one's own tract or the nine nearest tracts is equal to the probability of marrying someone located between 10 and 50 tracts away. This pattern holds for both within-group and intergroup marriages (see Figure A.3).

II.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 using the 50 nearest census tracts, considering 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.

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 (see Figure A.4 Panel A)

For age, we allow 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 full set of weights for both sexes are shown in Figure A.4 Panel B. The distribution of age differences are similar for within-group and intergroup marriages (details in Figure A.5). Further details on the construction of the geographic and age weights are provided in Appendix A.

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 33.6% 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 17.3% of neighbors are from high-income families. This means the median person from a high-income family has nearly twice as many high-income neighbors as the median person from a low-income family. Similarly, individuals from low-income families have 1.7 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 78.4% of their neighbors are White, whereas the median Black person lives in a neighborhood where only 44.9% 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 the share 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.

II.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 average 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 is associated with 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. Taken together, these results suggest that the class composition of one's neighborhood during young adulthood is a strong predictor of the class background of one's eventual spouse.

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 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, meaning a 10pp increase in the share of neighbors who are White is associated with a 0.9pp (42.9%) increase in the probability of having a White spouse. This relationship is steepest in majority-White neighborhoods, indicating that marginal increases in cross-race exposure are most strongly associated with increases in interracial marriage in the most segregated areas.

We summarize these relationships 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.

III Quasi-Experimental Effects of Neighborhood Exposure

The association between segregation and marriage outcomes documented in Section II.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.

III.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; Brainerd, 2017; Goni, 2022). A concurrent study by Zaremba (2024) also leverages the random assignment of sex at birth but uses it to examine how bargaining power in matching markets influences downstream maternal health outcomes. In contrast, we focus on the direct impact of this variation on marriage market outcomes.

As discussed in Section II, individuals living in nearby neighborhoods and of similar age constitute a significant component of the marriage market. Because people often live close to where they grew up, chance variation in local sex ratios—such as Black families having more opposite-sex than same-sex children in certain years—can have lasting effects on the composition of the marriage market. Such variation can increase the number of potential Black partners for individuals born in a given time and place.

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 from the perspective of opposite-sex relationships.

The second fact, as shown in Section II.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. In the language of Alvarez and Toneto (2024), our design estimates an average causal response function for the "marginal compliers"—the individuals who remain as adults near where they grew up. This excludes, for instance, those who move far away for college and do not return. We argue that these marginal compliers are an important group, as they represent the majority of Americans. 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 refer to 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] From the perspective of a White male, we describe the market as "tighter" when the sex imbalance among Black neighbors favors women, thereby increasing the probability of an interracial match relative to a scenario in which the imbalance favors men. Segregation, by contrast, is about the overall fraction of neighbors from a given race or class group. In Section IV, 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 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 sex ratios of children in nearby cohorts, 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

[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 peers (Anelli and Peri, 2019; Hoxby, 2000).

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. Notably, it even matches the discrete jumps caused by repeated integer values in some cells, which arise from the small number of Bernoulli draws within a particular tract-group-cohort cell. Consistent with this result, the instrument is uncorrelated across groups within a tract and unrelated to tract-by-birth-cohort covariates, such as mean parental income (see Figure A.6 and A.7).

III.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 in the share of opposite-sex and own-sex neighbors who belong to race or class group g (e.g., Black or high-income parents) in i's adult neighborhood 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 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 superscript a indicates that the measure is based on adult neighborhoods. The weights, which account for census tract distance and relative age, are described in detail in Section II.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. The superscript c indicates that the measure is based on the childhood neighborhood.

The first-stage and reduced-form equations examining how market tightness within group g affects the likelihood of marrying someone from group g are given by:

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

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

We instrument for market tightness among group g in the adulthood marriage market, $T_{i,g}$, with sex ratios among group g in person i's childhood neighborhood, $Z_{i,g}$. The o subscript indicates own group: in specifications focused on class, o indicates own-parent quartile, and for race, o indicates own race group. The first-stage and reduced-form coefficients, β^f and β^r, are pooled across the full sample, though we also present specifications that allow these coefficients to vary by o. 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 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. In our baseline specification by class, the sample includes individuals from families in the top and bottom quartiles of the parental income distribution. We set $g = -o$, where $-o$ denotes the "other" class group, so the outcome, treatment, and instrument variables are $(Y_{i,-o}, T_{i,-o}, Z_{i,-o})$. For example, $Y_{i,-o} = 1$ 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). We also report estimates separately for each group rather than pooling them. In the baseline specification by race, $Y_{i,-o} = 1$ 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 $N^c_{i,-o}$ across individuals 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. As expected, the variation in $Z_{i,g}$ in equations 1 and 2 is greatest in areas with smaller values of $N^c_{i,-o}$ (see Figure A.8).

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 $N_{i,-o}$ 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 $T_{i,-o}$ on indicators for each ventile bin of $Z_{i,-o}$. We plot the coefficients on each ventile bin against the mean value of $Z_{i,-o}$ in that ventile bin and normalize the mean of the coefficients and the mean of $Z_{i,-o}$ to zero. This method allows us to non-parametrically evaluate the first-stage relationships before proceeding with the 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.

Using the analogous method from Cattaneo et al. (2024), we plot the reduced-form relationships in Panels B and D of Figure VII, corresponding to the first-stage results in Panels A and C. We also 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 (S.E.=0.03) 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.[14]

Our exposure measure captures neighborhood-level demographics and, while policy relevant, does not directly capture interpersonal contact—likely a key channel through which proximity influences interclass marriage. We show that exposure has nearly twice the effect in counties where cross-class friendships are more common, using this as a proxy for interpersonal contact. We rely on friending bias from Chetty et al. (2022), which measures the extent to which people form cross-class Facebook friendships, holding exposure constant. Splitting the sample at the median of this measure, we find that exposure has a significantly larger effect in low-bias (i.e., high-contact) counties, suggesting that interpersonal interactions may be an important channel through which proximity increases interclass marriage (see Figure A.9). This finding is not mechanical: proximity could matter through other channels—for instance, by shaping the pool of potential partners on online dating platforms, irrespective of who individuals form friendships with.

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. The pooled coefficient is 0.001 (S.E.=0.018), and the 95% confidence interval rules out effects larger than 0.036. While we cannot rule out small effects of exposure on interracial marriage, the estimates decisively reject effects comparable in magnitude to those observed for class.

While our main analysis focuses on White-Black interracial marriages, we also estimate 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

[14] In theory, including the $Z_{i,g}$ variables from groups other than $-o$ should improve precision, even though our baseline estimates are unbiased since the group-specific sex ratios are uncorrelated. In practice, however, the precision gains when including them as controls or instrumenting for them are negligible. The standard errors are nearly identical to those in the baseline (see Figure A.11). We therefore use the parsimonious model in Equations 1 and 2 as our baseline specification.

effects on interracial marriages involving other racial combinations, with magnitudes resembling those observed for interclass marriages (see Figure A.10). This distinction underscores the unique dynamics of White-Black interracial marriage and suggests that other interracial pairings may be more responsive to exposure. We note that the near-zero effect for White-Black marriage is consistent with a stronger preference for same-race marriage or the possibility that neighborhood exposure does not translate into meaningful interpersonal contact.[15]

Nonetheless, our IV estimates underscore a central result: lack of residential 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 IV, where we directly quantify the overall impact of increasing exposure through reductions in residential segregation on interclass and interracial marriage. Before turning to that analysis, we document several additional patterns and robustness checks related to our IV results.

We find that our baseline estimates by class and race are similar—and statistically indistinguishable—across sexes (see Figure A.12). This aligns with the relatively small sex differences in intergroup marriage rates shown in Tables A.7 and A.8.

Market tightness influences interclass marriage by shaping whom people marry rather than affecting their overall likelihood of being married. Column 5 of Table IV shows that market tightness has a minimal and statistically insignificant impact on the likelihood of being married. This limited effect arises because new interclass marriages largely replace marriages to individuals from other class groups who grew up in different neighborhoods (see Figure A.13). When tighter markets within one's own class lead to more same-class marriages, these typically substitute for cross-class marriages within the neighborhood. Consistent with this substitution pattern, market tightness operates in a class-specific way. For example, tight markets among neighbors from bottom-quartile parent income families most strongly increase the likelihood of marrying someone from the same bottom quartile. In contrast, market tightness in other quartiles has small negative effects on marrying into the bottom quartile (see Figure A.14).

In addition to demonstrating that the distribution of our instrument aligns with random Bernoulli sex assignment, Table IV shows that our IV estimates remain consistent when using within-family variation in market tightness, ruling out any role for family-level 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 isolate variation exclusively from birth timing among siblings with the same childhood neighborhood. In both cases, the estimates closely align with our baseline results in column 2.

[15] Friending bias data from Chetty et al. (2022) is not available by race, limiting our ability to test heterogeneity in contact. Nonetheless, the relatively precise null effects of exposure make it unlikely that large positive effects exist in above-median-contact counties, as this would imply large negative effects elsewhere. This does not rule out a role for contact—it may simply be that interracial contact conditional on exposure is rare in most places.

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. 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 (see Figure A.15).

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'. 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 (see Figure A.15).

The exclusion restriction requires that sex ratios in childhood neighborhoods influence marriage outcomes only through partner availability in adult neighborhoods, not through alternative channels such as preference formation. If exposure to opposite-sex peers of a particular race or class shaped preferences, we might expect an increase in opposite-sex peers from other class groups to raise interclass marriage rates, even with individuals outside the cohorts and neighborhoods used to define the instrument. Figure IX evaluates the effect of cohort-specific sex ratios on cohort-specific marriage outcomes. Panel A shows the impact of own-class sex ratios on marriage to an own-class spouse born in the prior cohort. Panel C repeats this for other-class sex ratios and interclass marriage. Panels B and D examine the effects on marriage to a spouse born in the subsequent cohort.

Across all four panels, the sex ratio with the strongest influence on marriage outcomes is the one measured within the same cohort as the outcome (e.g., if the outcome is interclass marriage to someone one year younger, then the relevant sex ratio is in the cohort one year younger). For instance, an increase in opposite-sex individuals from the other class group in one's own cohort does not significantly affect the probability of interclass marriage in earlier or later cohorts. If the effects were driven by preference formation, we would expect an increase in opposite-sex peers from one's own cohort to raise the likelihood of interclass marriage across all other cohorts. Moreover, our IV estimates primarily reflect marriages between individuals from the same childhood neighborhood (Figure A.13), consistent with an exposure or partner-availability channel. Importantly, however, local exposure need not operate strictly within neighborhood boundaries. Because marriage is a one-to-one matching process, changes in partner availability in one neighborhood can generate spillovers that reallocate matches across nearby neighborhoods. These spillovers naturally generate changes in marriage patterns outside the focal neighborhood—consistent with the increase in out-of-neighborhood interclass marriages in response to local interclass availability (see Figure A.13 Panel A). As discussed in Section IV, a matching framework provides a coherent way to track these spillovers while enforcing the one-to-one constraint. Finally, because we focus on sex ratios, our estimates hold constant the overall level of exposure to each race or class group and therefore do not capture any effects of growing up in a more diverse environment on long-run preferences (e.g., Corno, La Ferrara and Burns, 2022; Gordon and Reber, 2018).

III.C From Market Tightness to Market Thickness

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.

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 is the conceptual distinction between market tightness and market thickness. Our IV estimates exploit variation in market tightness, which holds fixed the overall share of a group and varies the sex ratio within the group. In contrast, reducing segregation affects market thickness, which holds fixed the sex ratio within the group and varies the overall share of the group across neighborhoods.

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 using a simple three-neighborhood example in Figure X. Individuals are of two types—those with low-income parents and those with high-income parents—and they have preferences over both the class background of their spouse and distance to potential partners (i.e., the cost of marrying outside one's neighborhood). We consider a policy that moves residents with low-income parents from neighborhood C to neighborhood A. Solid black lines denote marriage matches that would occur absent the move, while dashed orange lines denote counterfactual matches after the move. For illustrative purposes, assume person 1 has a stronger-than-average preference for interclass marriage, whereas person 2 has a stronger preference for within-class marriage.

Figure X demonstrates how the same partial-equilibrium change within the treated neighborhood can have different implications for the total number of interclass marriages, depending on spillovers across neighborhoods. Panel A depicts a case in which, on average, preferences over class dominate distance costs. After the move, person 1—who values interclass marriage—can marry person 3 at lower distance cost. Person 2, who prefers within-class marriage, marries person 4, even though this requires marrying outside the neighborhood. In this scenario, the individual labeled $\star$ chooses a within-class marriage with person 6 in neighborhood C after the move. Overall, the move from C to A generates one new interclass marriage in total, which also appears as one new interclass marriage among residents of neighborhood A, the neighborhood "treated" by the move.

Panel B instead depicts a case in which, on average, distance costs dominate class preferences. In this scenario, the individual labeled $\star$ prefers to marry person 5 across class lines rather than incur the higher distance cost of a within-class match. As a result, the same move generates two new interclass marriages in total. However, from the perspective of neighborhood A alone, the increase in interclass marriage remains one in both panels. Thus, focusing only on partial-equilibrium effects within the treated neighborhood can mask economically meaningful general-equilibrium spillovers across neighborhoods. This example underscores the importance of estimating the total general-equilibrium effect of reducing segregation, which is the focus of Section IV.

IV A Spatial Model of the Marriage Market

We develop 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. The model identifies the underlying matching parameters consistent with our IV estimates, and then uses this structure to simulate equilibrium responses to counterfactual policies such as reductions in segregation. The goal is to provide economic magnitudes and context that reduced-form estimates alone cannot capture, building on related approaches that map changes in marriage outcomes to parameters of marriage matching models (e.g., Adda, Pinotti and Tura, 2025). 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., 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., 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 useful step before applying it to larger counterfactuals.

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, 2010). This counterfactual corresponds to the effect of eliminating segregation while holding fixed downstream changes to preferences or endogenous resorting. Importantly, the effects of eliminating segregation may differ from those of small reductions if there are diminishing returns—moving from near-zero cross-group exposure to a modest level may have a larger marginal impact on intergroup marriage than subsequent increases in already diverse settings.

IV.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, 2018; Mourifié and Siow, 2021; Siow, 2015). While such assumptions may be appropriate in some contexts, they are not in ours: individuals disproportionately marry neighbors, and hyperlocal changes in neighborhood composition meaningfully affect marriage outcomes. Ignoring geography would therefore conflate exposure with preference, attributing low intergroup marriage rates entirely to preferences even if they are partly driven by proximity. Expanding the type space to include geography introduces computational challenges, which we address in our discussion of 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 might not hold in our setting. Incorporating the unobserved trait also allows exposure effects to vary across neighborhoods with identical observable demographics, making the model more realistic. The unobserved trait, similar to a "latent type" (e.g., Arcidiacono, Muralidharan and Singleton, 2024), serves as a reduced-form way to capture neighborhood-level heterogeneity in preferences, unobserved characteristics, or even the likelihood of interaction conditional on exposure.

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. 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 (see Figure A.16).

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. The two-type example in Appendix E illustrates how the IV estimates based on sex ratio variation help distinguish between these mechanisms. If selection alone drives the observed pattern, the IV coefficients will tend toward zero. More broadly, the simulation shows how model parameters map to reduced-form estimates: under the matching model, reduced-form estimates based on sex ratio variation are larger when proximity has a greater weight in the utility function and when individuals are more open to intergroup marriage.

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 set 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, thereby maximizing total surplus in the marriage market. These transfers need not be monetary but may include implicit, non-monetary exchanges between partners. 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.

Men's preferences,

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

and women's preferences,

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

are the sum of the systematic component (α), the transfer (τ), and the idiosyncratic component (ε). The idiosyncratic component follows a type one extreme value distribution:

$$\begin{aligned} \varepsilon_{k,i} &\sim \text{Gumbel}(0,1), \\ \varepsilon_{j,i} &\sim \text{Gumbel}(0,1). \end{aligned}$$

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.

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,

$$\mu_{j,k} = \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}}, \tag{5}$$

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,

$$m_j = \mu_{j,0} + \sum_{k=1}^{T} \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}} \tag{6}$$
$$w_k = \mu_{0,k} + \sum_{j=1}^{T} \sqrt{e^{\gamma_{j,k}} \mu_{j,0} \mu_{0,k}}. \tag{7}$$

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. However, for any candidate solution, the model also implies a full set of bilateral matches, $\mu_{j,k}$, which contains on the order of T^2 endogenous outcomes. 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 adjust the sample within these tracts to be nationally representative. A map of the selected tracts, drawn from Boston, MA, is shown in Figure XI Panel A. Because our empirical moments are based on national estimates, we reweight the Boston sample to make it nationally representative by race and class. We apply a single scaling factor across all tracts, preserving baseline segregation patterns while ensuring overall representativeness (e.g., that 25% of individuals fall into each parent income quartile). 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.

In a typical application of the Choo and Siow (2006) model, the types, m_j and w_k, are known, and Equation 5 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. We estimate the model using the simulated method of moments (SMM) (Laffont, Ossard and Vuong, 1995; McFadden, 1989). We use a diagonal weighting matrix that assigns equal weight across moment categories (national means, OLS, and IV) and, within each category, places greater weight on moments for the four key race and class groups emphasized in our baseline results. This ensures that the quasi-experimental moments for these focal groups meaningfully influence estimation. In contrast, weighting by the inverse variance matrix would place nearly all weight on national marriage outcomes, which are estimated with extremely high precision and would crowd out the quasi-experimental variation required to identify counterfactuals (see Appendix D.II for details).

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. At the neighborhood level, we model the prevalence of the trait as a function of average neighborhood income, but crucially, we allow the parameters of this function to differ by race and class. This approach captures the idea that the types of individuals living in high- versus low-income neighborhoods are likely to differ, and that these differences may vary by group. Details are provided 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 6 and 7 can be used to solve for each $\mu_{j,k}$, the number of type-specific matches.[16] Using μ, we compute the simulated moments analogous to those covered in Sections II and III, 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. The model provides a close match for the data moments (see Figure A.17). Details on the SMM objective function and the included moments are provided in Appendix D.

[16]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).

IV.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. 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.[17] 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., 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. This corresponds to the population-weighted mean change in interclass marriage rates per average change in exposure across the entire market. The model predicts that moving a small number of individuals from the most segregated neighborhoods would produce a slope of 0.30, implying that a 10pp increase in exposure to the other class group would cause a 3.0pp (S.E.=0.8pp) increase in interclass marriage rates.[18] 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

[17] We move a racially representative resident of each class based on the national share of race groups in bottom and top quartile families.

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

White suburbs. 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 (S.E.=0.2pp) 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. We view this as an important out-of-sample validation of the model's performance, particularly given the limited statistical power to estimate causal effect of exposure in the subset of highly segregated neighborhoods with our baseline IV strategy.

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 III—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 by 1.5pp (S.E.=0.2pp), from 3.1% to 4.6%. Eliminating segregation closes 35.2% (S.E.=5.4%) of the gap between the observed interclass marriage rates and the rates that would occur if individuals married at the same rate and chose their spouse without regard to class. Comparing instead to the rates that would occur under current marriage rates and random spouse selection, removing segregation would close 46.8% (S.E.=7.1%) of the gap. In both cases, 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 by 0.9pp (S.E.=0.1pp), 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% (S.E.=0.6%) of the gap between observed interracial marriage rates and the rates predicted under the fully random benchmark, and 13.2% (S.E.=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). 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).

V Conclusion

Marital homophily can exacerbate economic inequality between race and class groups, reinforcing disparities both within and across generations (Binder et al., 2022; Chadwick and Solon, 2002; Mare, 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 social 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; Doob, 2015; Fischer and Mattson, 2009; Putnam, 2016).

Reducing marital homophily by race is likely to require efforts that go beyond addressing racial segregation. One possible reason for the limited impact of desegregation is that it may not generate meaningful interpersonal interactions between racial groups. Understanding the conditions under which exposure translates into interaction—and how such interactions influence preferences for within-group versus intergroup marriage—remains an important area for further study. 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.

Adda, Jérôme, Paolo Pinotti, and Giulia Tura. 2025. "There's more to marriage than love: the effect of legal status and cultural distance on intermarriages and separations." *Journal of Political Economy*, 133(4): 1276–1333.

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

Alvarez, Luis AF, and Rodrigo Toneto. 2024. "The interpretation of 2SLS with a continuous instrument: A weighted LATE representation." *Economics Letters*, 237: 111658.

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.

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.

Arcidiacono, Peter, Karthik Muralidharan, and John D Singleton. 2024. "Experimentally Validating Welfare Evaluation of School Vouchers." National Bureau of Economic Research.

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.

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.

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.

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

Chetty, Raj, John N Friedman, Nathaniel Hendren, Maggie R Jones, and Sonya R Porter. 2025. "The opportunity atlas: Mapping the childhood roots of social mobility." *The American Economic Review*, forthcoming.

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. "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. 2018. "Bidimensional Matching with Heterogeneous Preferences: Education and Smoking in the Marriage Market." *Journal of the European Economic Association*, 16(1): 161–198.

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.

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.

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. "Exposure, Integration, and Interracial Marriage: Evidence from The Great Migration and Residential Segregation." Working Paper.

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

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

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.

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.

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.

Goldman, Benjamin, Jamie Gracie, and Sonya R. Porter. 2026. "Data and Code for: Who Marries Whom? The Role of Segregation by Race and Class." *Inter-university Consortium for Political and Social Research [distributor].*

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.

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.

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

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.

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.

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

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

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.

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. 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.

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.

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.

Tolbert, Charles M., and Molly Sizer. 1996. "U.S. Commuting Zones and Labor Market Areas: A 1990 Update." U.S. Department of Agriculture, Economic Research Service Staff Report AGES-9614.

US Census Bureau. n.d.*a*. "1989, 1994, 1995, and 1998–2015 Federal Income Tax Returns." [restricted data].

US Census Bureau. n.d.*b*. "Census 2000 and 2010 Short Forms." [restricted data].

US Census Bureau. n.d.*c*. "Census 2000 long form and 2005–2015 American Community Surveys (ACS)." [restricted data].

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.

Wen, Jiayi, and Jiaming Mao. 2025. "Assortative Marriage and Geographic Sorting." *Available at SSRN 5143049*.

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	0.078	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 on the sample 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 individual's incomes, respectively. Panel C provides information on the availability of address data and the moving patterns of individuals in adulthood. Panel D presents descriptives on marriage outcomes measured at age 30. All monetary values are measured 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 I for more detail on the matching 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 those with missing class as participating in the marriage market, which we estimate as the share of spouses with missing parent income data (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 race groups listed. Note that the pooled row, with the exception of the benchmark calculations, includes the missing race group which is approximately 5% of the sample; therefore sample shares listed at the bottom do not sum to 1. The benchmarks are defined as in Table I. See the notes to Table I for additional details.

TABLE III: Distance from Eventual Spouse; Excluding Couples at Same Address

	By Age			By Year Prior to Marriage				
	0	10	18	1	3	5	10	20
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
A. Percent of Couples that Ever Lived in Nearest Census Tracts								
100 Census Tracts	0.289	0.398	0.509	0.725	0.700	0.650	0.514	0.356
50 Census Tracts	0.235	0.339	0.449	0.658	0.631	0.582	0.451	0.298
25 Census Tracts	0.183	0.279	0.383	0.579	0.552	0.504	0.382	0.239
10 Census Tracts	0.123	0.200	0.291	0.463	0.437	0.393	0.287	0.167
B. Percent of Couples that Ever Lived in Radius								
25 Miles	0.375	0.472	0.577	0.804	0.782	0.733	0.589	0.437
10 Miles	0.240	0.336	0.438	0.667	0.641	0.588	0.446	0.299
5 Miles	0.143	0.225	0.312	0.517	0.490	0.441	0.317	0.192
C. Percent of Couples that Ever Lived in Same Geographic Area								
State	0.558	0.649	0.746	0.908	0.894	0.861	0.753	0.619
Commuting Zone	0.398	0.490	0.591	0.811	0.790	0.743	0.604	0.457
County	0.274	0.370	0.472	0.696	0.671	0.621	0.481	0.333
Census Tract	0.026	0.051	0.089	0.184	0.166	0.139	0.086	0.040

Notes: This table presents statistics on the geographic proximity of married couples in childhood and adulthood. All statistics are constructed from a dataset that contains one row per couple. The ages are evaluated relative to the younger member of the couple, if the individuals were not born in the same year. Parent addresses are used for ages 18 and below. Addresses at later ages are defined using the individual's own addresses. Missing years are imputed, separately for childhood and adulthood locations. For more information on the construction of the address panel, see Section I. Columns 1-3 present the fraction of eventual couples that had ever lived within a defined geographic area by age 0, 10, and 18, respectively. Columns 4-7 present analogous statistics, but evaluated at particular points in time relative to the year the couple marries. Years in which couples reside at the exact same address, defined using the MAFID, are excluded.

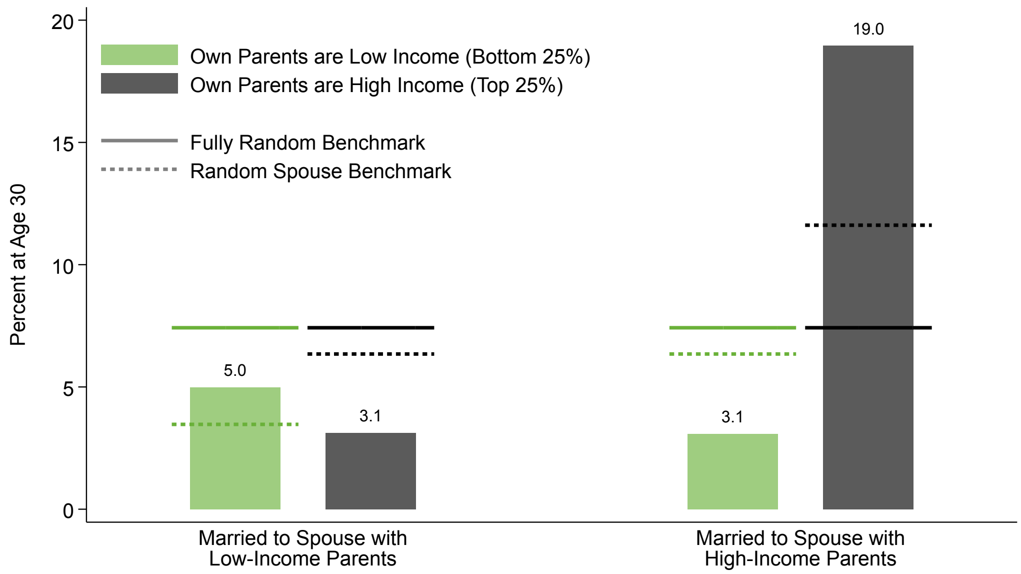
TABLE IV: Effect of Market Tightness Among Other Race or Class Group on Intergroup Marriage and Overall Marriage Rates

		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

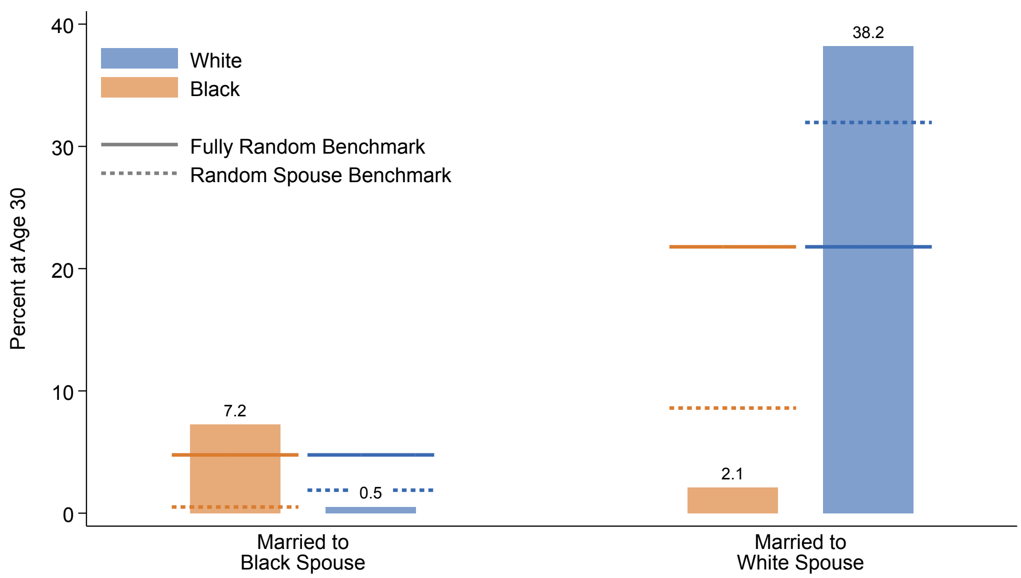
Notes: This table presents regression coefficients for the effect of exposure to different class and race groups on marriage outcomes. In Panel A, we use a sample of individuals from low-income (bottom quartile) and high-income (top quartile) families. In Panel B, we use a sample of white and Black individuals.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

FIGURE I: Marital Homophily by Class and Race

A: Rates of Interclass vs. Within-Class Marriage



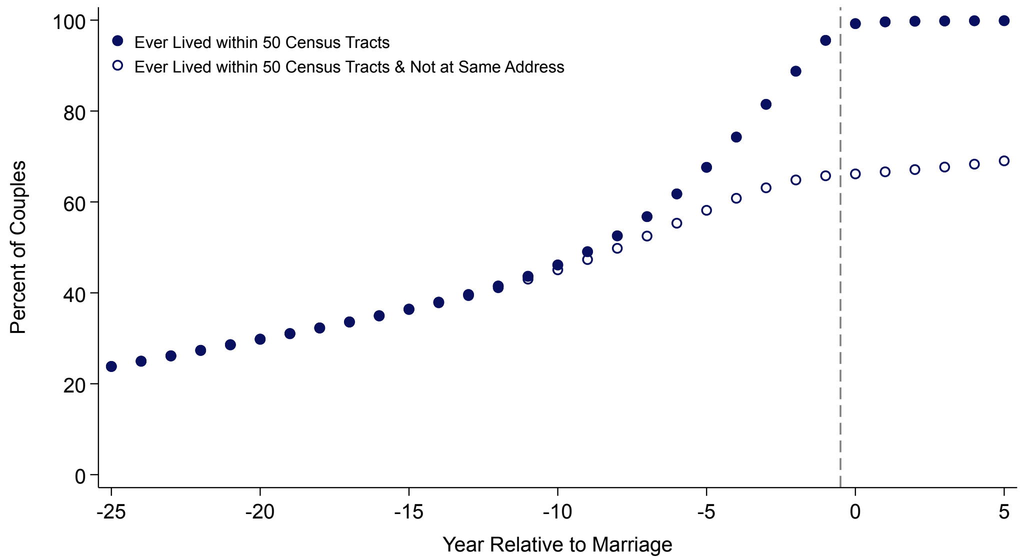
B: Rates of Interracial vs. Within-Race Marriage



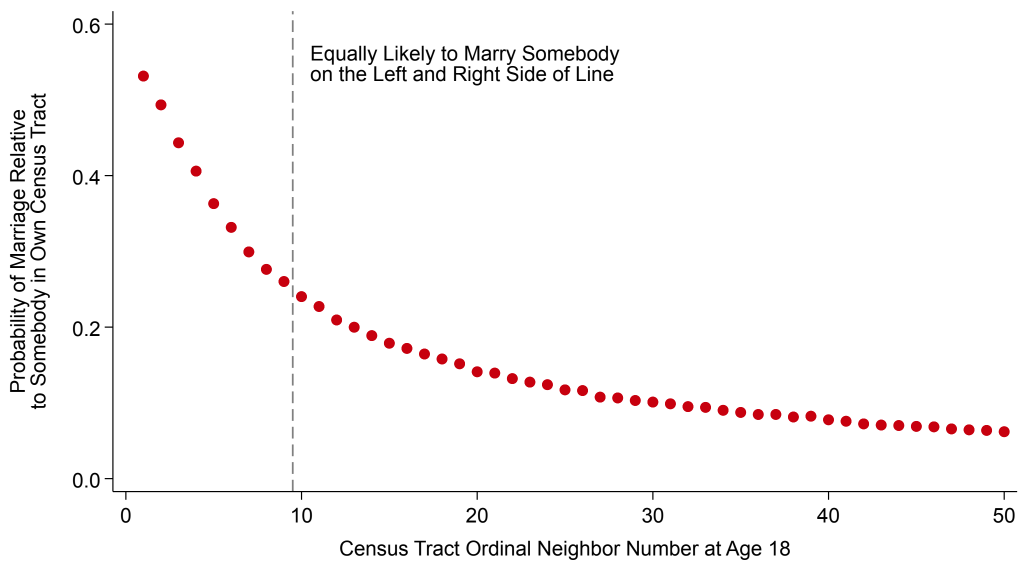
Notes: This figure summarizes key patterns of marital homophily by class (parent income) and race. In Panel A, we plot the fraction of individuals married at age 30 to a spouse from a low-income (bottom parent quartile) or high-income (top quartile) family, separately for individuals from low- and high-income 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 A.1 and A.2, respectively.

FIGURE II: The Role of Neighborhoods in Marriage Markets

A: Distance From Spouse by Years from Marriage

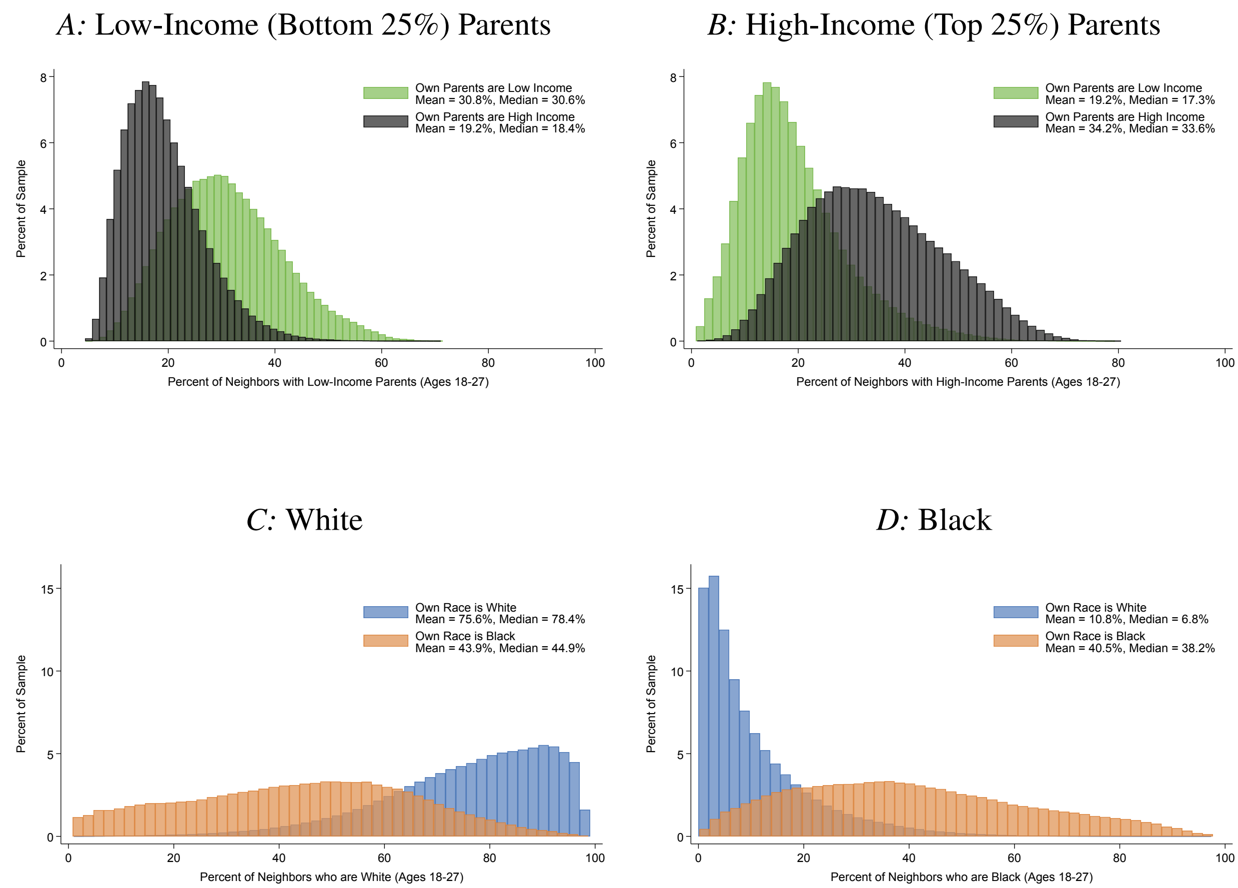


B: Marriage Probability Decay by Distance



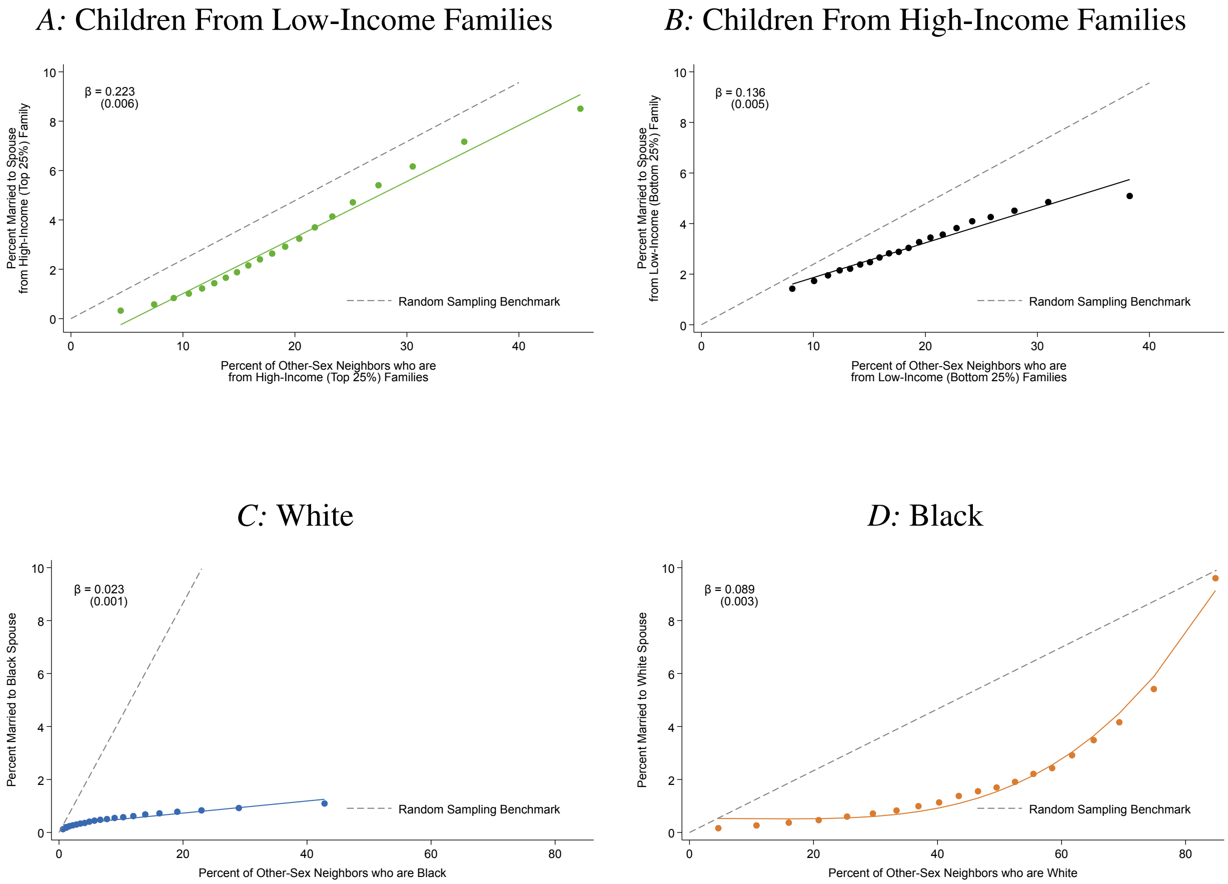
Notes: This figure presents statistics on the local nature of marriage markets. In Panel A, we plot the fraction of couples who had ever lived within 50 Census tracts by a given year relative to the year they married. The solid circles represent individuals who lived near each other, including those 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 III. 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.4 Panel A. That version provides the decay weights used to weight neighbors, as described in Section II.C.

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



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 I. 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.4B. For more details on the construction of the exposure measure, see Section II.C.

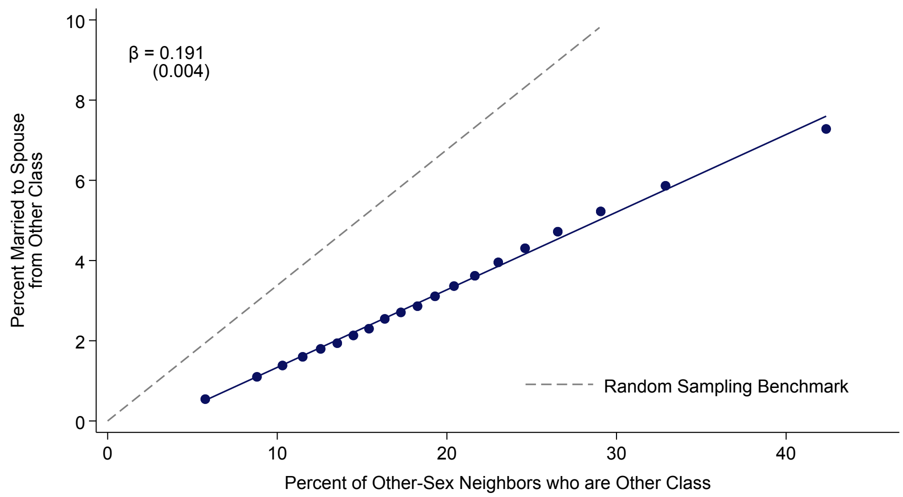
FIGURE IV: Association of Cross-Group Exposure and Cross-Group Marriage



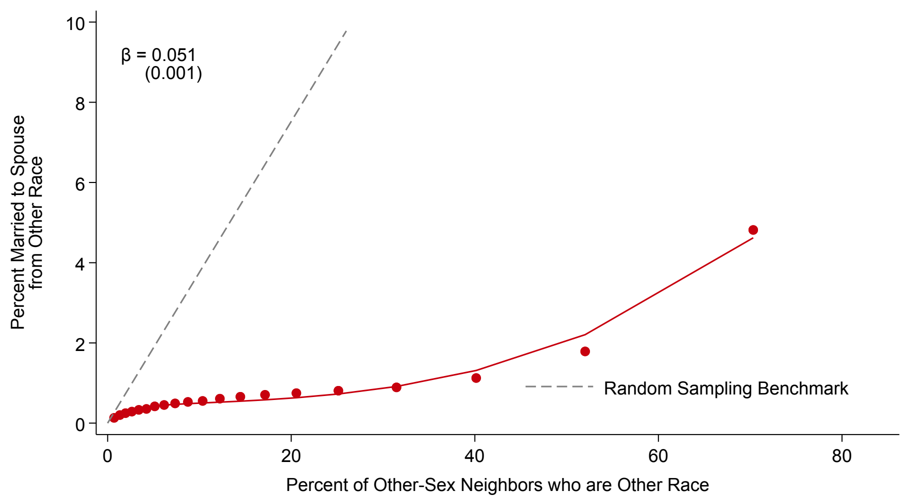
Notes: This figure presents binned scatter plots of the relationship between cross-group exposure and cross-group marriage. In Panel A, we show the relationship between the fraction of individuals from low parent income families who are married to a spouse from a high parent income family and the fraction 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 II.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 White 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%



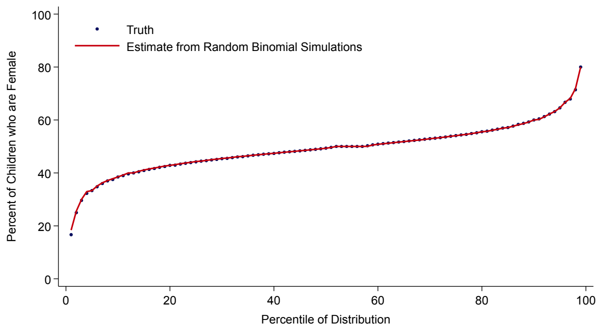
B: Rates of Interracial Marriage vs. Exposure to Other Race
Sample: White and Black



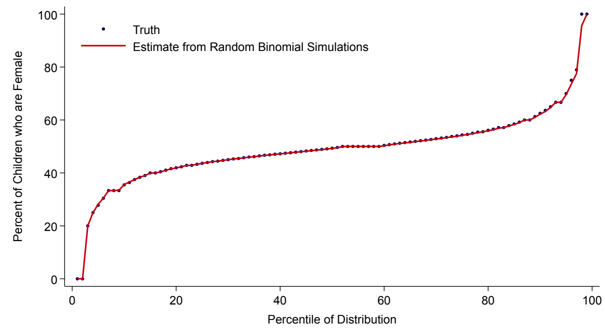
Notes: This figure presents binned scatter plots of the relationship between cross-group marriage and cross-group exposure, separately for class (Panel A) and race (Panel B). In Panel A, we combine the data used to create Figure IV Panels A and B to construct a single binned scatter plot. The sample 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

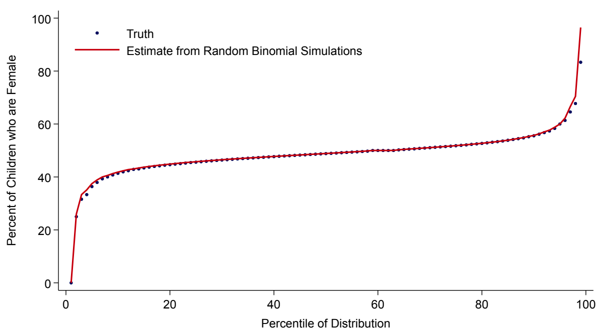
A: Children From Low-Income Families



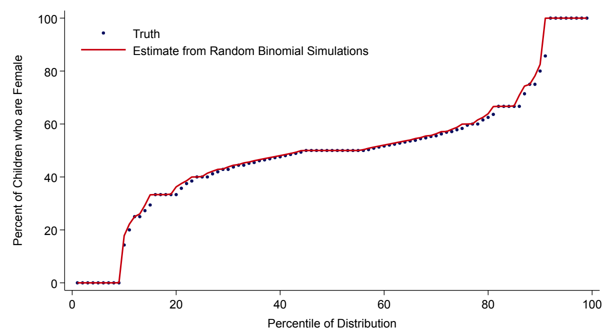
B: Children From High-Income Families



C: White



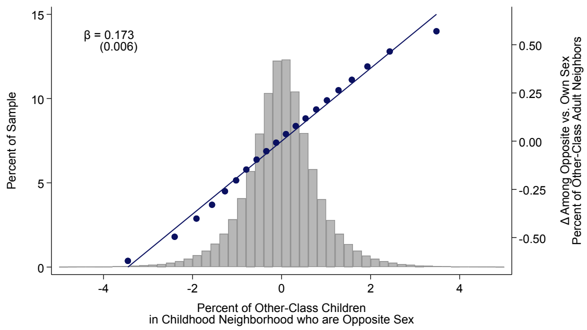
D: Black



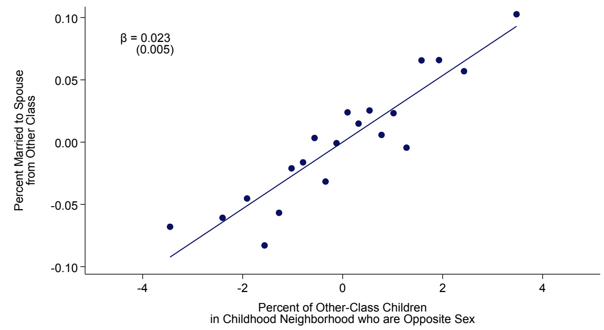
Notes: This figure examines whether observed sex ratio variation is consistent with random variation drawn from a binomial distribution, separately for individuals from low- and high-income families (Panels A and B) and for White and Black individuals (Panels C and D). The sex ratio is calculated as 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

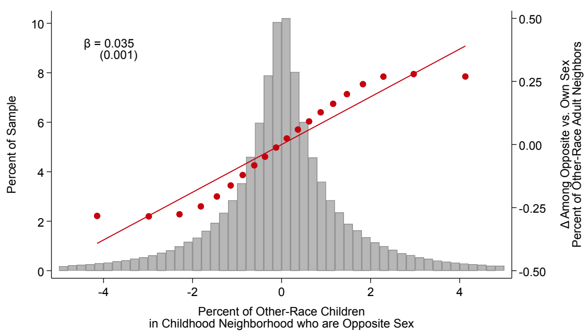
A: First Stage and *Z* Distribution
Sample: Bottom 25% and Top 25%



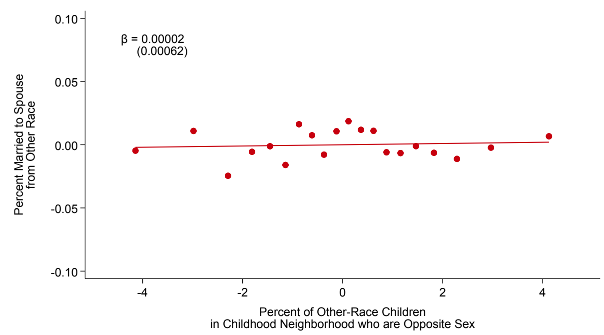
B: Reduced Form, Interclass Marriage
Sample: Bottom 25% and Top 25%



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

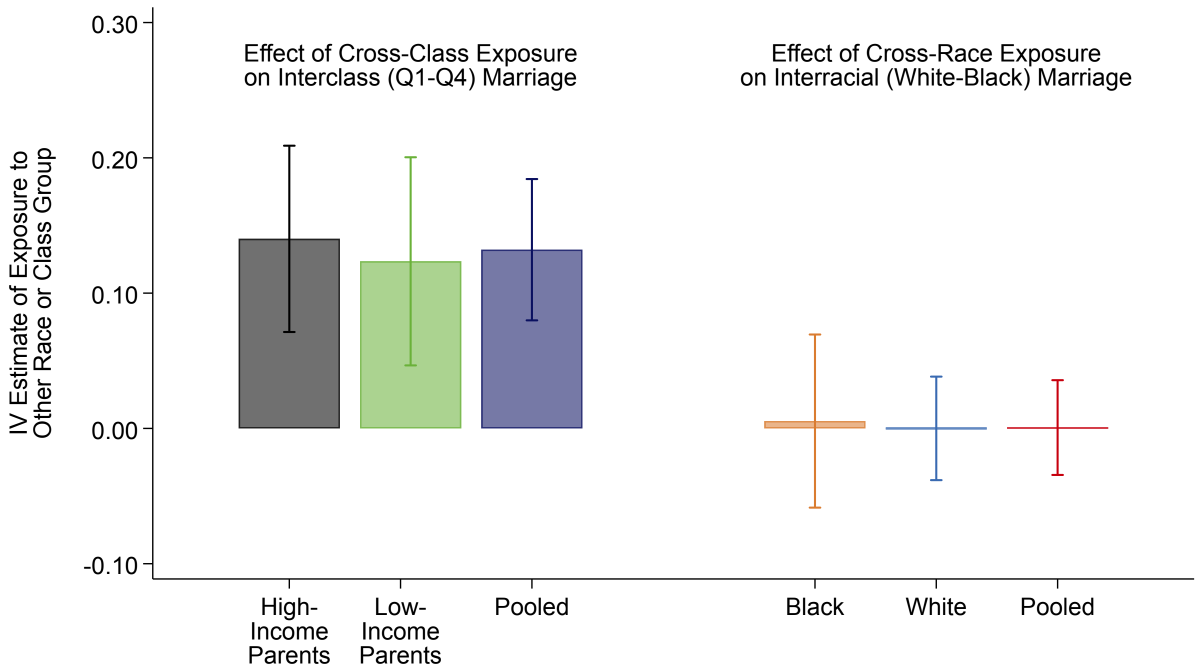


D: Reduced Form, Interracial Marriage
Sample: White and Black



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. (2024). 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.

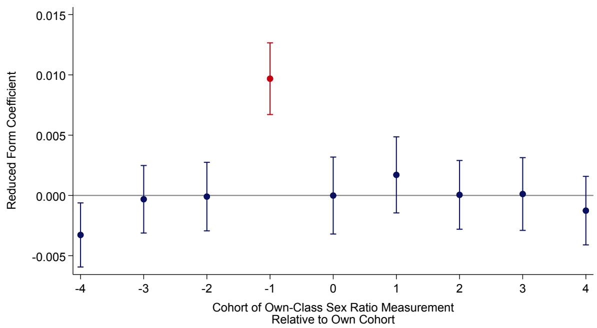
FIGURE VIII: IV Estimates of Cross-Group Exposure on Cross-Group Marriage



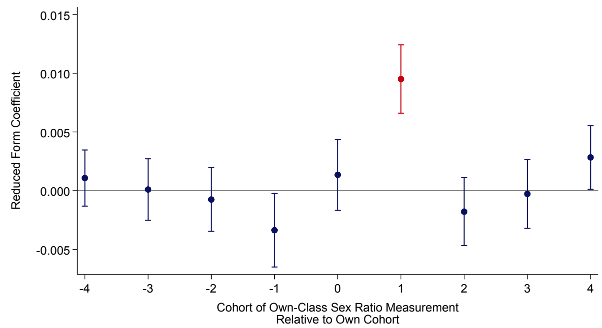
Notes: This figure summarizes the key IV results for interclass and interracial marriage. Results are presented separately for individuals from low- and high-income families, as well as for White and Black individuals, and are pooled for class and race, respectively. For details on the specification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

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

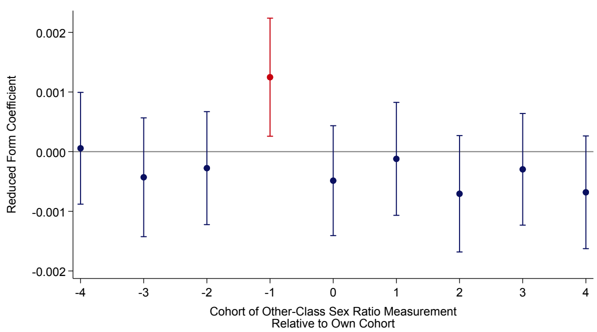
A: Own-Class Marriage to Spouse Born in Relative Cohort -1



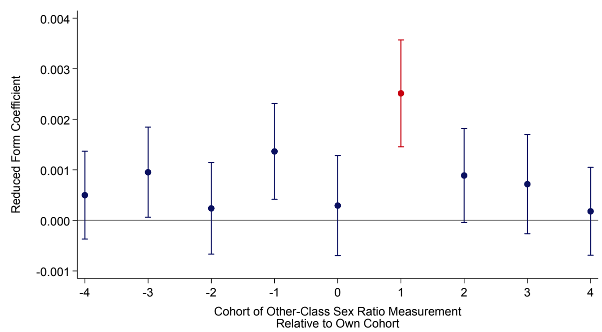
B: Own-Class Marriage to Spouse Born in Relative Cohort +1



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



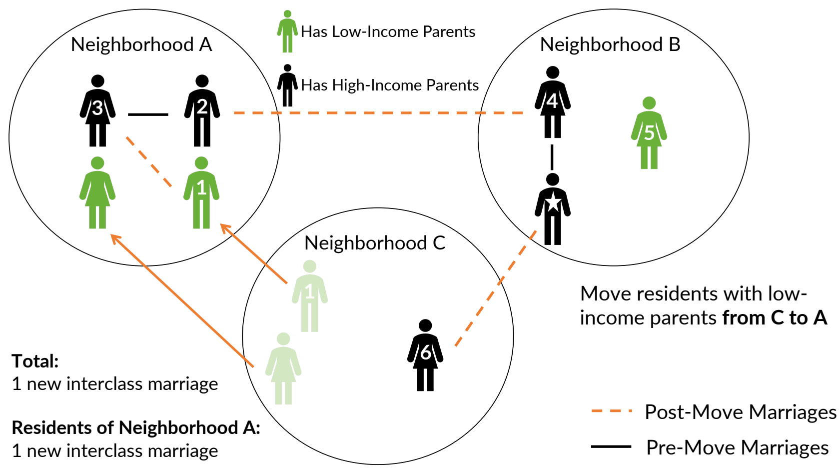
D: Interclass Marriage to Spouse Born in Relative Cohort +1



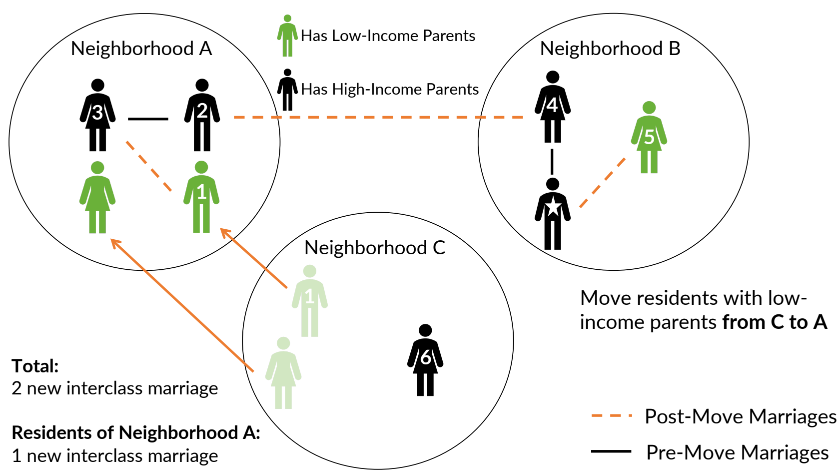
Notes: This figure presents a series of reduced form, outcome-based placebo tests to demonstrate the cohort specificity of the sex ratio variation. In each panel, the outcome variable is defined as having a spouse in a particular birth cohort by class group. In Panels A and B, the outcome is an own-class 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 III.B and the notes to Table IV for further details on the specification.

FIGURE X: Three Neighborhood Example of Spillover Effects

A: Class Preferences > Distance Costs: ★ Marries Person 1



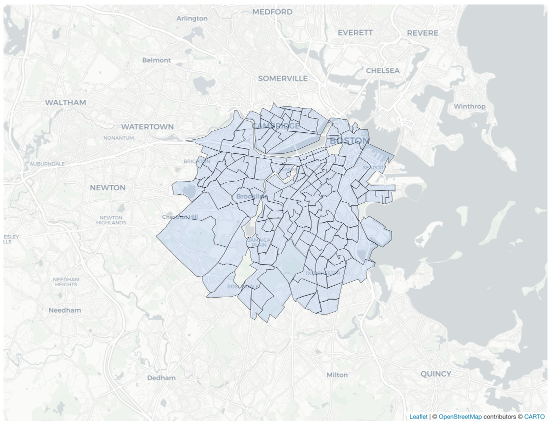
B: Distance Costs > Class Preferences: ★ Marries Person 2



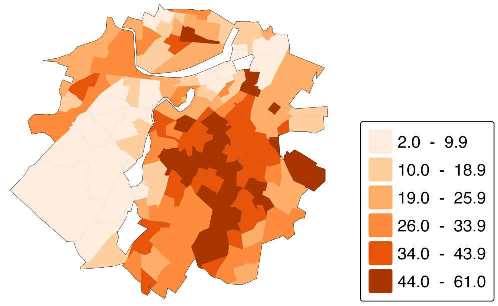
Notes: This figure presents a stylized three-neighborhood example illustrating how a policy that reduces residential segregation can generate spillovers in marriage outcomes across neighborhoods. Individuals are of two types—those with low-income parents (green) and those with high-income parents (black)—and have preferences over both the class background of their spouse and distance (the cost of marrying outside one's neighborhood). Solid black lines denote marriages that would occur absent the policy, while dashed orange lines denote counterfactual marriages after relocating residents with low-income parents from Neighborhood C to Neighborhood A (the treated neighborhood). Panel A considers a case in which, on average, preferences over class dominate distance costs. Following the move, the individual marked with a ⋆ marries person 6, meaning this reduction in segregation generated one new interclass marriage overall, which also appears as one new interclass marriage among residents of Neighborhood A. Panel B considers a case in which distance costs dominate class preferences. In this scenario, the starred individual marries person 5, resulting in two new interclass marriages in total, though still only one involving a resident of Neighborhood A. Thus, while the partial-equilibrium effect in the treated neighborhood is identical across panels, the total number of interclass marriages differs due to spillovers affecting residents of untreated neighborhoods.

FIGURE XI: Model Setup and Demographics of Chosen Neighborhoods

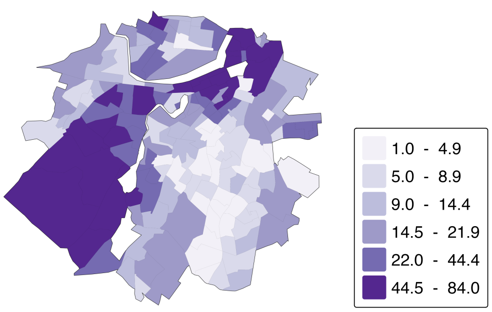
A: Analysis Sample of Census Tracts



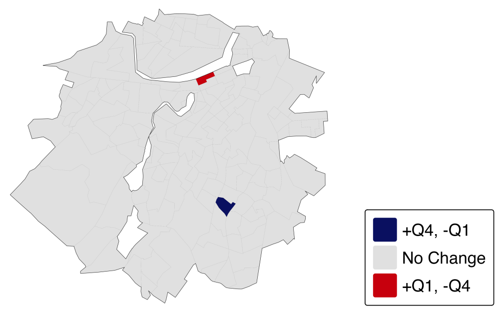
B: Percent from Low-Income Family



C: Percent from High-Income Family

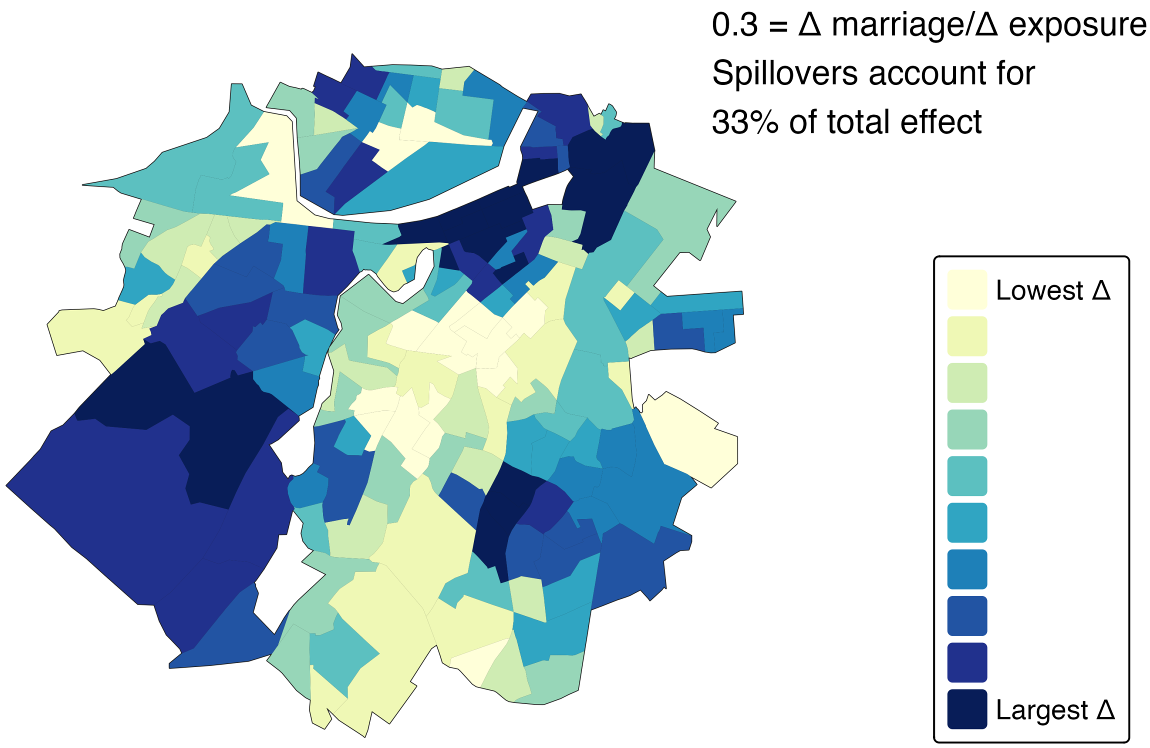


D: Simulated Moving Experiment



Notes: This figure shows the region of Boston used to estimate the model described in Section IV and the policy experiment described in Section IV.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 IV.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.

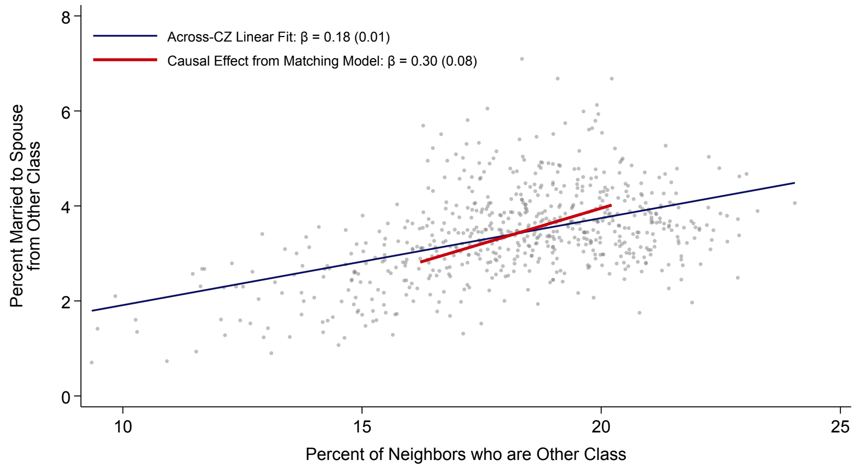
FIGURE XII: Impact of Policy Experiment on Interclass Marriage



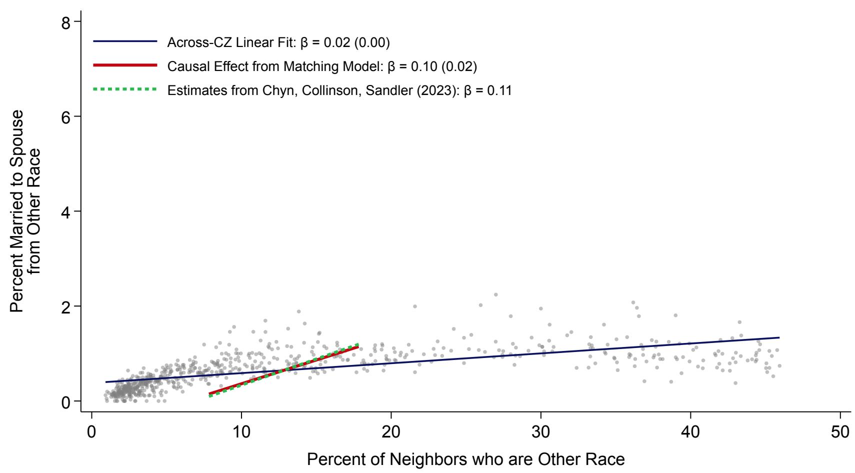
Notes: This figure illustrates the effect of the policy experiment depicted in Figure XI, Panel D, on interclass marriage rates. In the experiment, racially representative individuals from low-income families (one male and one female) are relocated from a tract in Dorchester to a tract in Back Bay. They 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%



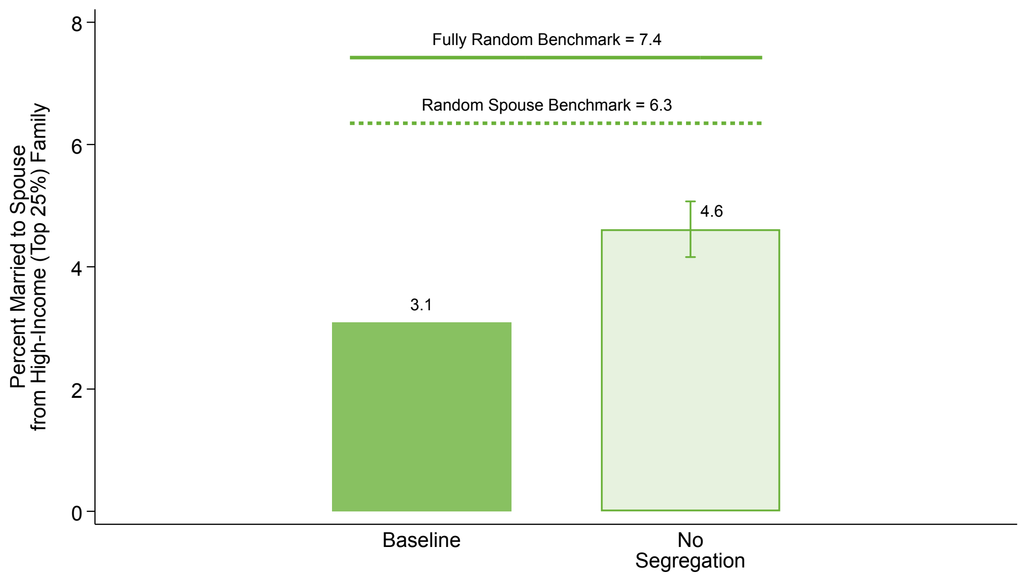
B: Rates of Interracial Marriage vs. Exposure to Other Race
Sample: White and Black



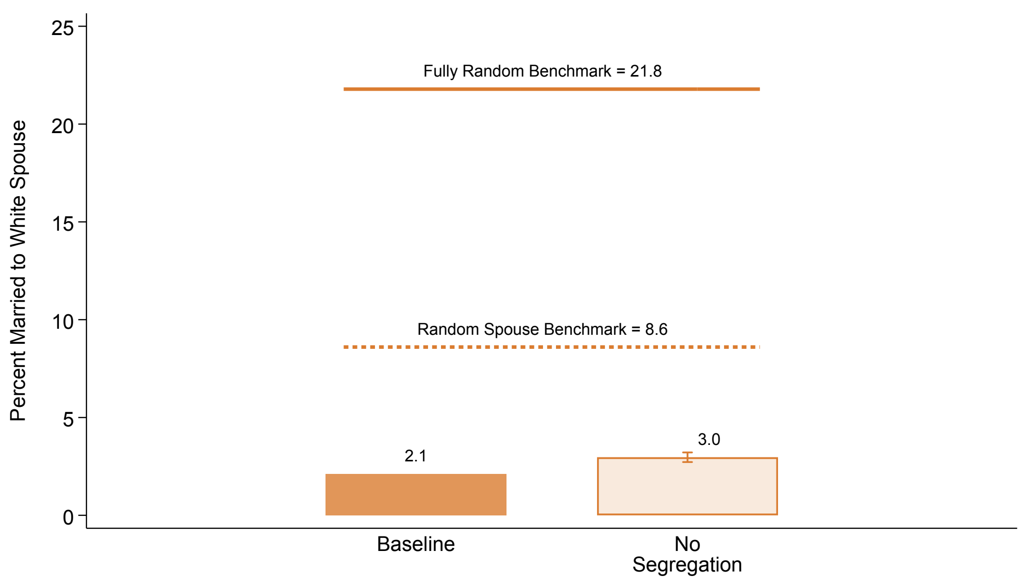
Notes: This figure plots the relationship between interclass (Panel A) and interracial (Panel B) marriage at the commuting zone level against exposure to neighbors of different class and race groups. This relationship is compared to estimates from the matching model developed in Section IV. To construct 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 SectionIV.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 IV. 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 IV.B. The standard error bars report 95% confidence intervals based on the procedure detailed in Appendix D.IV.