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

Benjamin Goldman, Jamie Gracie, Sonya R. Porter

A Exposure

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

The weights used are shown in Figure A.4. The relative cohort weights are sex-specific and are constructed as the share of marriages involving individuals at a particular census tract distance or age gap, scaled by the number of marriages at the most common distance level or age gap nationally. Specifically, as shown in Figure A.4 Panel A, separate distance weights are applied for each tercile of the population density distribution, but for simplicity, this distinction is omitted from the notation in this section. Let M^k denote the number of marriages between individuals k tracts apart, and $M^{m,s(i)}$ denote the number of marriages involving individuals in relative birth cohort m for people of sex s. We calculate the combined weights by multiplying the distance and age weights as follows:

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

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

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

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

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

$$e_{i,g}^{\alpha} = \frac{N_{i,g}^{\alpha}}{N_{i}^{\alpha}}.$$

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

$$e_{i,g}^{a} = \frac{1}{10} \sum_{j=18}^{27} \frac{N_{i,g}^{j}}{N_{i}^{j}}.$$

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

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

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

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

B Relationship Between Treatment and Instrument

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

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

and the fraction who are from sex s as

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

The childhood neighborhood version of market tightness is

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

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

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

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

C Constructing the Sample of Tracts and Types

C.I Tract Sample

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

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

Including Census tracts in the type space adds substantial computational complexity to the model. The dimension of μ would be more than five trillion if the model were solved on the entire U.S. Instead, we solve the model on a sample of 150 census tracts so that the dimension of μ is approximately 23 million. We choose the sample of tracts shown in Figure XI Panel A and construct the sample to be nationally representative using the procedure described below. The nationally representative sample is important because the moments the model is being matched to are estimated nationally. Because the model does not allow preferences to vary by geographic area, the chosen sample matters only because it determines the topography of segregation, i.e., who lives where.

We estimate tract-level counts in each parent income quartile and race cell using publicly available Census-based tabulations rather than the underlying confidential microdata. Although we observe the true counts in the restricted Census environment, we rely on predicted values so that we can present maps of model-based predictions without disclosing tract-level counts. To do so, we begin with the data at the census tract and race level published in Chetty et al. (2025). We restrict the sample to tracts with at least 100 people living in them in the 2000 Census. Within each tract and race cell, these data also include the fraction of individuals who have parents with below median parental income and the mean parent income percentile.

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

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

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

C.II Unobserved Trait

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

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

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

$$\Omega_{t,q,r} = \omega^1_q + \omega^2_r + (\omega^3_q + \omega^4_r)p_t$$

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

D Model Estimation

D.I Parameterization of Marital Surplus

We model marital surplus,

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

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

Match value based on class, $\chi_{c(j),c(k)}$, is a function of the parent income quartile of type j, $c(j)$, and k, $c(k)$ (e.g., there is a surplus for a top-bottom quartile match, bottom-bottom, bottom-second, etc.). The $\rho_{r(j),r(k)}$ is the analogous component for race, where $r(t)$ is a function

[1] The rescaling is done by first multiplying the counts in each tract–race–parent-income-quartile cell by one constant per race group and then by one constant per parent income quartile to ensure that 25% of individuals belong to each quartile nationally. This single pass gets the joint distribution close to national race and income shares, but it is an approximation: adjusting the income distribution mechanically perturbs race totals (and vice versa). Exact satisfaction of all margins would require full iterative proportional fitting.

that gives the race group of type t. If k has the unobserved trait, the surplus includes that value that a person of race $r(j)$ and class $c(j)$ places on the unobserved trait, $\upsilon_{c(j),r(j)}$. If j has the unobserved trait, then surplus includes the type k's value of the unobserved trait, $\upsilon_{c(k),r(k)}$.

The distance cost portion of surplus is defined by

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

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

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

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

D.II SMM Estimation

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

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

We use three different types of moments in ψ. The first set of moments are the national marriage outcomes by race and class. For each parent income quartile, we compute the fraction of individuals married to a spouse from each of the other quartiles.[3] We compute the analogous marriage outcomes by race, but aggregate our data into four race groups for simplicity: non-Hispanic White, non-Hispanic Black, Hispanic, and other. National marriage outcomes by class and race are in Tables A.1 and A.2.

Our second set of moments are the descriptive associations between marriage outcomes and

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

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

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

Our third and final set of moments are the IV estimates of the causal effect of market tightness, i.e. exposure to opposite-sex vs. own-sex members of other race and class groups, on interclass and interracial marriage. For each parent income quartile, we include the IV effects of exposure on marriage to somebody from each of the other quartiles and construct the analogous estimates by race.[4] We include random sex ratio variation in our sample, which allows us to estimate a model analogue of our IV coefficients. Section III details the construction of the empirical version of these estimates and the IV coefficients for our four key race and class groups are summarized in Figure VIII. Figure A.17 shows the relationship between the simulated model moments and the data moments used in estimation.

Once μ is known for fixed θ, we can compute $\hat{\psi}_s(\theta)$, the model moments evaluated at θ. We repeat this process using a numerical solver, adjusting θ at each step until the objective function in Equation D3 is minimized. To ensure all three sets of moments contribute to estimation, we use a modified identity matrix for weighting. Specifically, we upweight moments corresponding to four key groups: children from the bottom and top quartiles of the parent-income distribution, White individuals, and Black individuals. Moments involving a single key group receive twice the baseline weight. Moments in which both own and spouse samples correspond to the same key group receive three times the baseline weight, while intergroup moments involving two key groups receive four times the baseline weight. Adjustments to this weight matrix have little impact on our counterfactual estimates, given their precision. In contrast, the “optimal" weight matrix (inverse of the covariance matrix) would place nearly all weight on national marriage moments, which are estimated with near-infinite precision. Doing so would assign negligible weight to our quasi-experimental moments, which are central to the counterfactual analyses, as shown in Appendix E.

[4]We exclude the IV effect of exposure to own-race on within race marriage. There is little sex ratio variation within race groups because of high-levels of racial segregation. We show in Figure A.18 that there is approximately 36 times more sex ratio variation across versus within race groups. This is not an issue for estimating IV effects within class—there is three times more own-class versus own-race variation.

D.III Computing Model Equilibrium

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

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

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

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

We can apply the quadratic equation to solve

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

and the analogous equations for the female equilibrium

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

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

D.IV Estimating the Variance of Counterfactuals

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

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

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

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

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

E Sorting on Unobservable in Matching Model

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

The intuition behind how our IV estimates disentangle sorting from exposure can be illustrated with a simple example with only two types, Packers and Vikings. There are nine total neighborhoods. Packers tend to live near Green Bay, while Vikings typically reside near Minneapolis. In Figure A.16 Panel A, we show the fraction of neighborhood residents who are Packers and Vikings as one moves from Minneapolis to Green Bay. In Minneapolis, 90% of residents are Vikings, and the proportion of Vikings gradually decreases until reaching Green Bay, where only 10% of residents are Vikings. Conversely, 10% of Minneapolis residents are Packers, compared to 90% in Green Bay. Approximately 35% of Vikings possess a trait, unobserved by the researcher, that makes them more attractive to Packers.

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

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

In Panel D, we show that despite producing similar observable relationships between intergroup marriage and cross-group exposure, the two scenarios lead to very different IV estimates. In the no-sorting case (purple), Vikings are much more likely to marry a Packer as the share of opposite-sex Packers in their neighborhood increases. However, when the gradient in Panel C is driven entirely by sorting, the relationship between the share of Vikings married to Packers and the share of nearby opposite-sex Packers disappears.[5]

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

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

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

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

Notes: This table presents observed rates of interclass marriage, measured at age 30, and two benchmarks. For each of the four class quartiles, we calculate the overall marriage rate, the fraction who have a spouse whose own parent(s) were in each of the four quartiles, as well as the fraction who have a spouse who is missing parent income. These spouses cannot be matched to parents using the procedure described in Section I. 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 the population, which we estimate as the share of spouses with missing parent income data.

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

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

Notes: This table presents observed rates of interracial marriage, measured at age 30, and two benchmarks. For each of the six race groups, we calculate the marriage rate, the fraction who have a spouse from each of the race groups, as well as what fraction have a spouse who is missing race. Spouses are assigned race following the procedure in Section I, regardless of whether or not they can be assigned to parents. The rates of missing spouse race are substantially lower than missing spouse class, for this reason. The fully random benchmark is the overall marriage rate multiplied by the fraction of the population from each race group. The random spouse benchmark is equal to the actual marriage rate for each race group multiplied by the fraction of the married population that comes from each race group. For both benchmarks, we include those with missing race the population, which we estimate as the share of spouses with missing race data. Note that to construct the benchmarks in this table, unlike Table II which uses the full sample, we exclude individuals who are themselves missing race (approximately 5% of individuals). They are computed only using individuals who are white, Black, Asian, Hispanic, AIAN, or in the other race group.

TABLE A.3: Cohabitation Outcomes by Class Relative to Random Matching Benchmark

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

Notes: This table presents observed rates of interclass cohabitation, measured at age 30, and two benchmarks. Cohabitation is defined using data from the ACS. Concretely, we take individuals in the main analysis sample who receive the ACS in the year they are 30. Individuals are cohabiting if they respond to the survey as the primary respondent and list a married spouse or a non-married partner in their household as an additional respondent or if they are listed as the spouse or partner of the primary respondent. The benchmarks are defined analogously to those in Table A.1, using cohabitation rather than marriage. The fully random benchmark is the overall cohabitation rate multiplied by the fraction of the population from each class group. The random partner benchmark is equal to the actual cohabitation rate for each race group multiplied by the fraction of the cohabiting population that comes from each class group.

TABLE A.4: Cohabitation Outcomes by Race Relative to Random Matching Benchmark

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

Notes: This table presents observed rates of interracial cohabitation, measured at age 30, and two benchmarks. Cohabitation is defined using data from the ACS. Concretely, we take individuals in the main analysis sample who receive the ACS in the year they are 30. Individuals are cohabiting if they respond to the survey as the primary respondent and list a married spouse or a non-married partner in their household as an additional respondent or if they are listed as the spouse or partner of the primary respondent. The benchmarks are defined analogously to those in Table A.2, using cohabitation rather than marriage. The fully random benchmark is the overall cohabitation rate multiplied by the fraction of the population from each race group. The random partner benchmark is equal to the actual cohabitation rate for each race group multiplied by the fraction of the cohabiting population that comes from each race group.

TABLE A.5: Marriage Outcomes by Class Relative to Random Matching Benchmark; Sample: 1982 Birth Cohort

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

Notes: This table presents statistics on rates of interclass marriage measured at age 37. The sample contains individuals in the 1982 birth cohort, the oldest in the main analysis sample, who are 37 in 2019, the last year available in the tax data. For details on construction, see the notes to Table A.1.

TABLE A.6: Marriage Outcomes by Race Relative to Random Matching Benchmark; Sample: 1982 Birth Cohort

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

Notes: This table presents statistics on rates of interracial marriage measured at age 37. The sample contains individuals in the 1982 birth cohort, the oldest in the main analysis sample, who are 37 in 2019, the last year available in the tax data. For details on construction, see the notes to Table A.2.

TABLE A.7: Marriage Outcomes by Class and Sex

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

Notes: This table presents statistics on rates of interclass marriage separately by sex. For details on construction, see the notes to Table A.1.

TABLE A.8: Marriage Outcomes by Race and Sex

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

Notes: This table presents statistics on rates of interracial marriage separately by sex. For details on construction, see the notes to Table A.2.

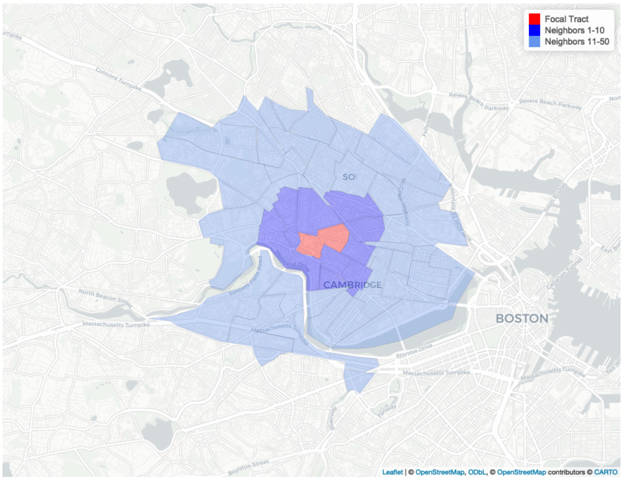
TABLE A.9: Distance from Eventual Spouse

	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	0.290	0.398	0.511	0.963	0.846	0.725	0.522	0.357
50 Census Tracts	0.235	0.340	0.451	0.956	0.815	0.676	0.461	0.298
25 Census Tracts	0.183	0.279	0.386	0.946	0.779	0.622	0.395	0.240
10 Census Tracts	0.123	0.201	0.295	0.931	0.726	0.543	0.304	0.168
B. Percent of Couples that Ever Lived in Radius								
25 Miles	0.375	0.473	0.579	0.972	0.882	0.784	0.595	0.437
10 Miles	0.240	0.337	0.440	0.955	0.816	0.678	0.456	0.299
5 Miles	0.143	0.225	0.316	0.936	0.746	0.572	0.332	0.192
C. Percent of Couples that Ever Lived in Same Geographic Area								
State	0.558	0.650	0.747	0.987	0.942	0.886	0.756	0.620
Commuting Zone	0.398	0.491	0.593	0.971	0.885	0.791	0.609	0.457
County	0.274	0.371	0.475	0.959	0.831	0.702	0.491	0.334
Census Tract	0.027	0.052	0.094	0.893	0.599	0.364	0.112	0.040

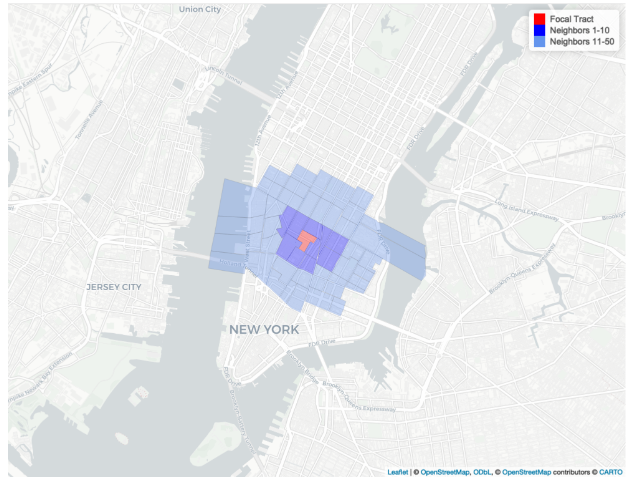
Notes: This table presents statistics on the geographic proximity of married couples in childhood and adulthood. All statistics are constructed from a dataset that contains one row per couple. Unlike Table III, this table includes couples who are living at the same address. The ages are evaluated relative 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.

FIGURE A.1: Examples of Nearest 50 Census Tracts

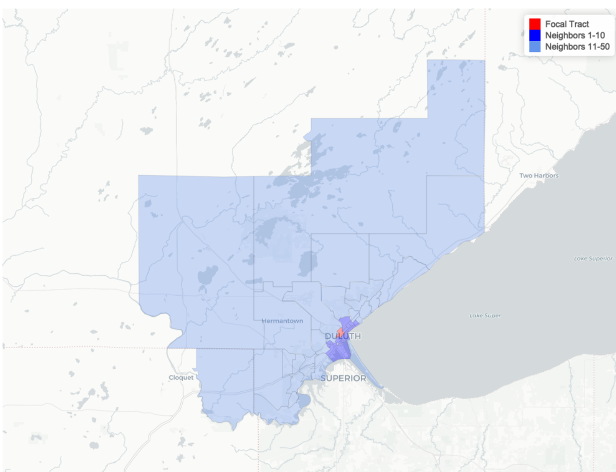
A: Cambridge, MA



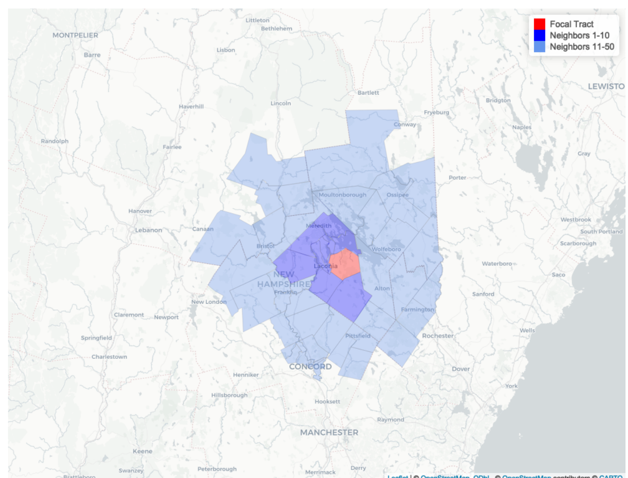
B: New York City, NY



C: Duluth, MN



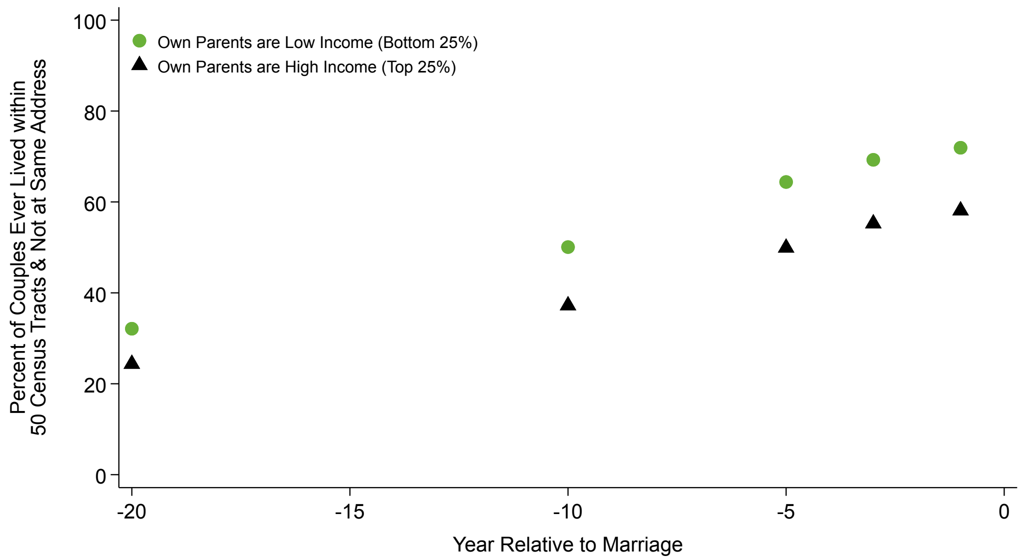
D: Gilford, NH



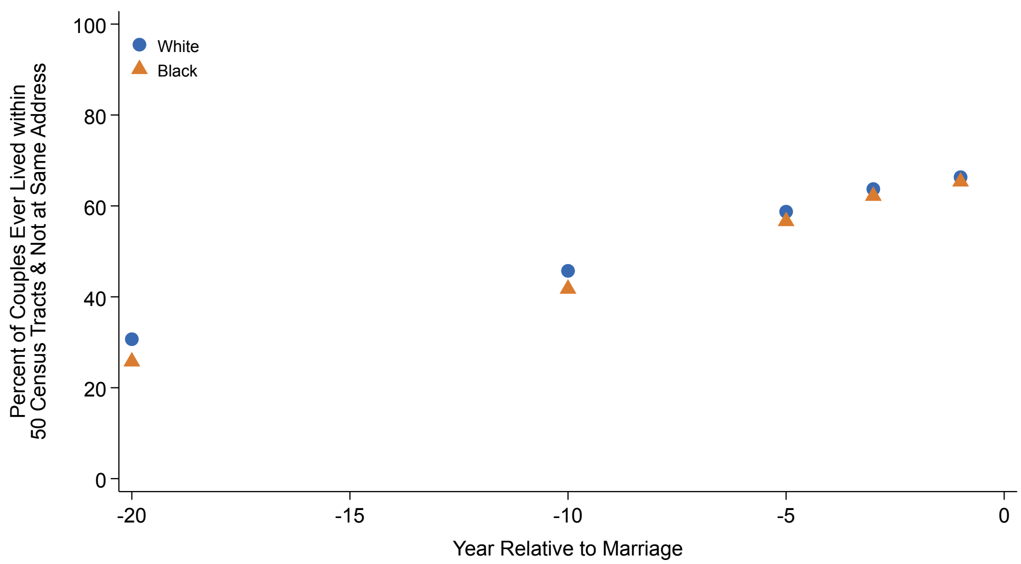
Notes: This figure depicts the nearest 50 Census tract neighbors of a particular focal tract in four U.S. cities. In each case, the neighbors are defined using the distance between Census tract centroids. The focal tract is shown in red, neighbors 1-10 in dark blue, and neighbors 11-50 in light blue.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

FIGURE A.2: Distance From Spouse by Class and Race

A: Class, Bottom 25% and Top 25%



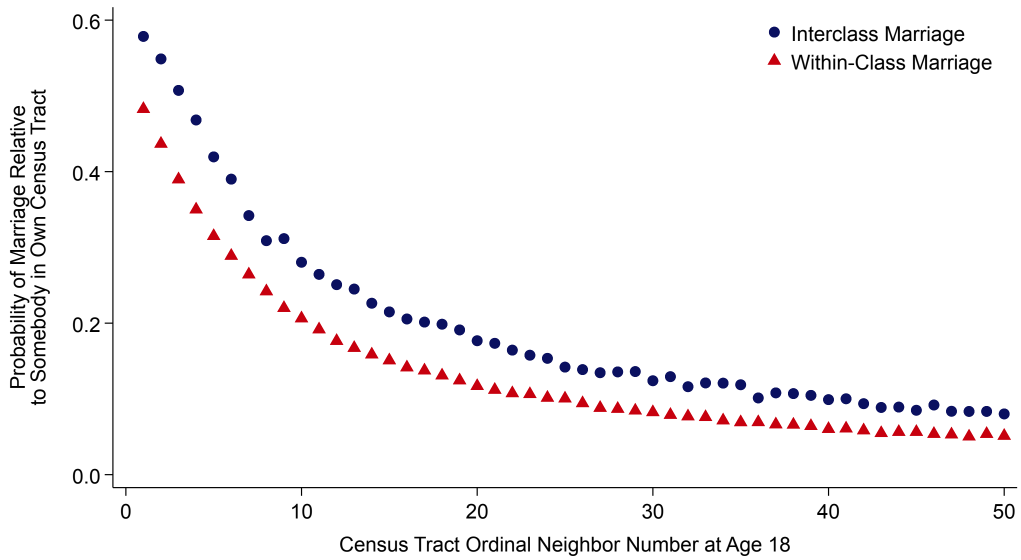
B: Race, White and Black



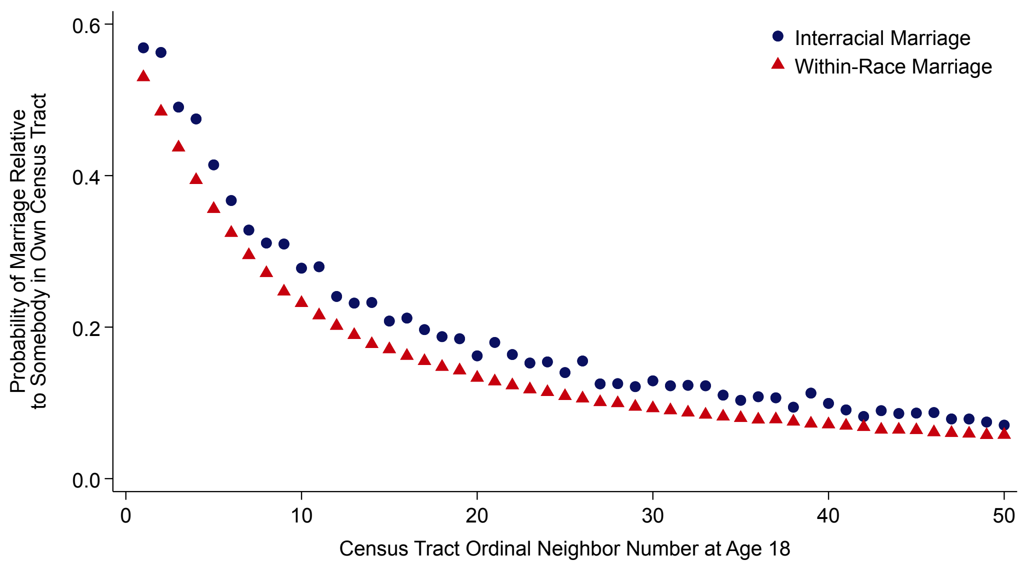
Notes: This figure presents statistics on the local nature of marriage markets, following the approach in Figure II Panel A, but disaggregated by class (Panel A) and race (Panel B). We plot the share of couples who had ever lived within 50 census tracts of each other by a given year, relative to the year they married. Years when couples lived at the exact same address are excluded.

FIGURE A.3: Marriage Probability by Distance and Marriage Type

A: Sample: Bottom 25% and Top 25%



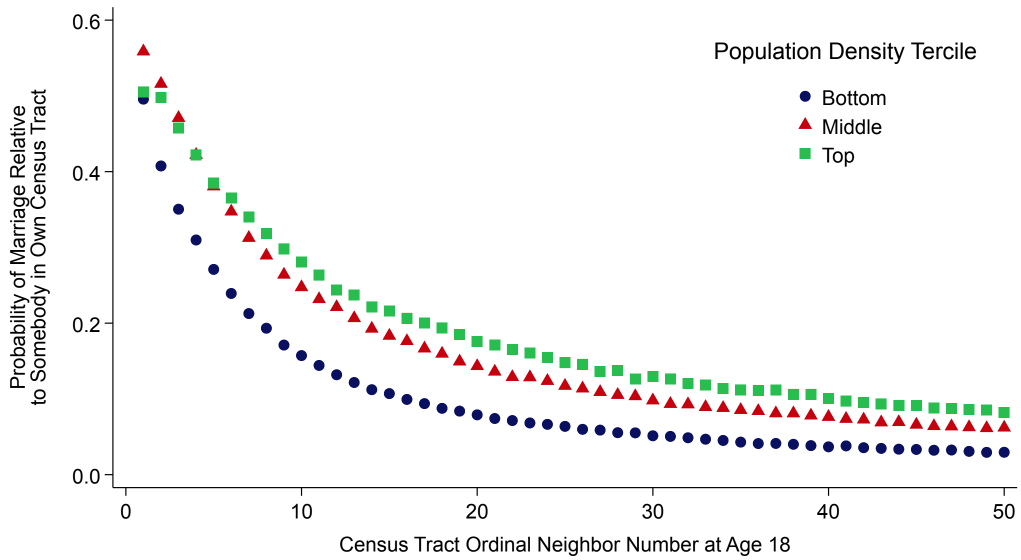
B: Sample: White and Black



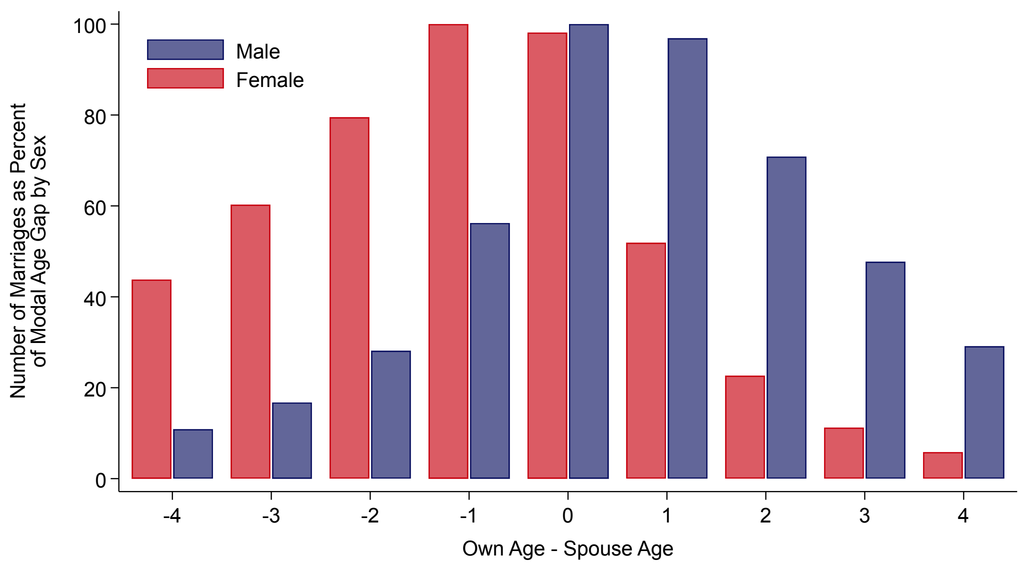
Notes: This figure shows how the likelihood of marrying someone from nearby Census tracts varies by whether individuals marry within or across class and race groups, following the approach in Figure II Panel B. We plot the probability of marrying someone from each of the 50 nearest Census tracts to a person's age-18 address, normalized by the probability of marrying someone from their exact tract. Panel A focuses on individuals from low- and high-income families. Interclass marriage is defined as marrying someone from the opposite income group. Panel B includes Black and White individuals, with interracial marriage defined as a Black-White marriage. In both panels, within-group marriage refers to marrying someone from the same class or race group.

FIGURE A.4: Distance and Age Weights Used to Construct Exposure Measure

A: Marriage Probability Decay by Distance and Population Density



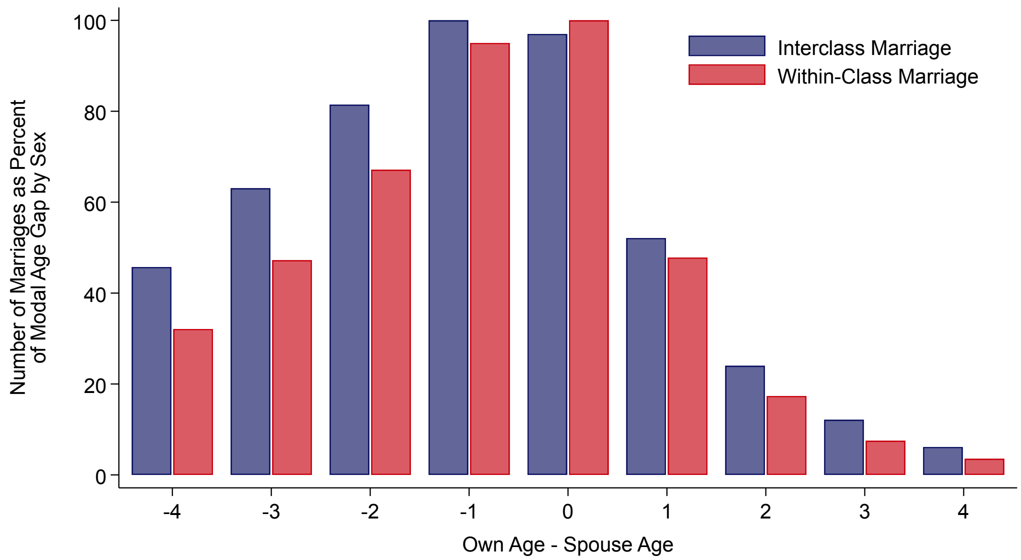
B: Difference in Age From Spouse by Sex



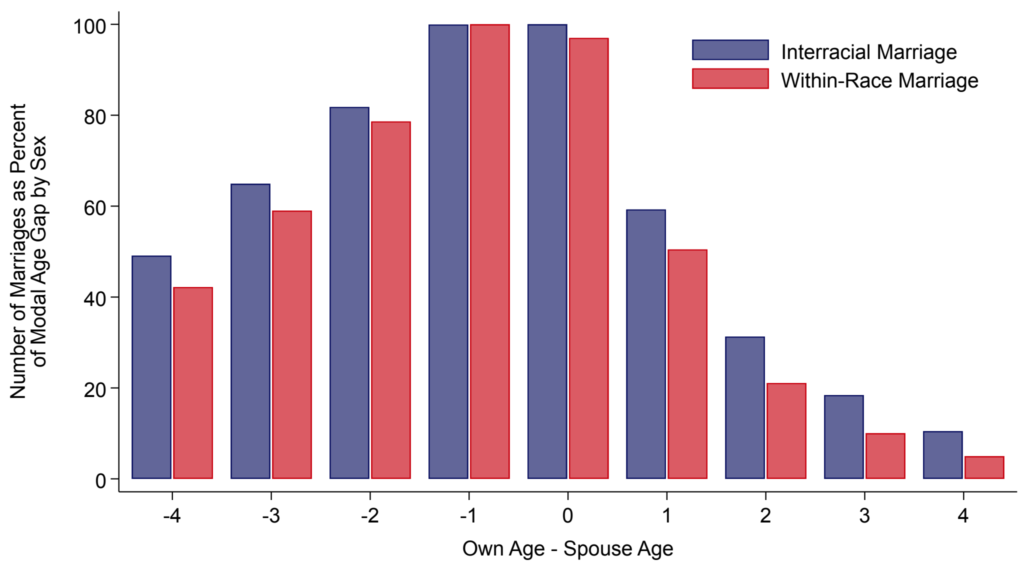
Notes: This figure presents the weights used to construct the exposure measure described in Section II.C. In Panel A, we present the distance weights, separately for each tercile of population density, measured in 2000. A pooled version of this figure is shown in Figure II Panel B; see figure notes for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

FIGURE A.5: Difference in Age From Spouse by Marriage Type

A: Sample: Female, Bottom 25% and Top 25%



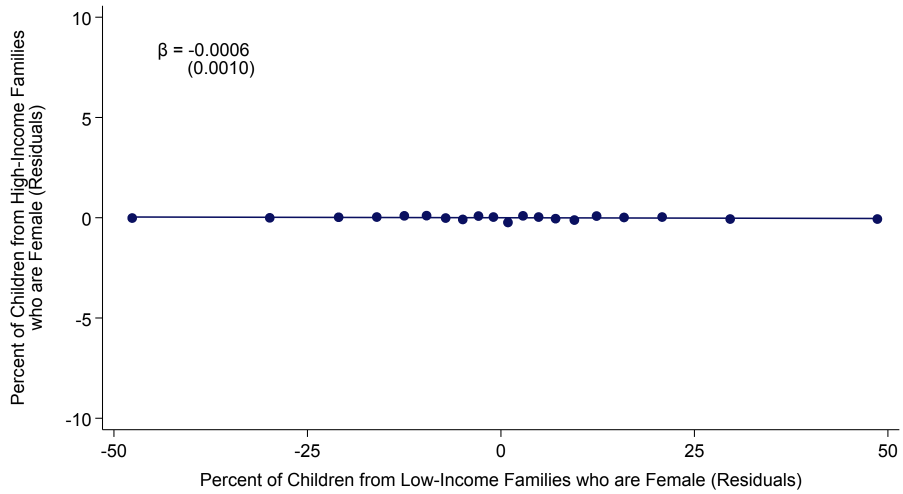
B: Sample: Female, White and Black



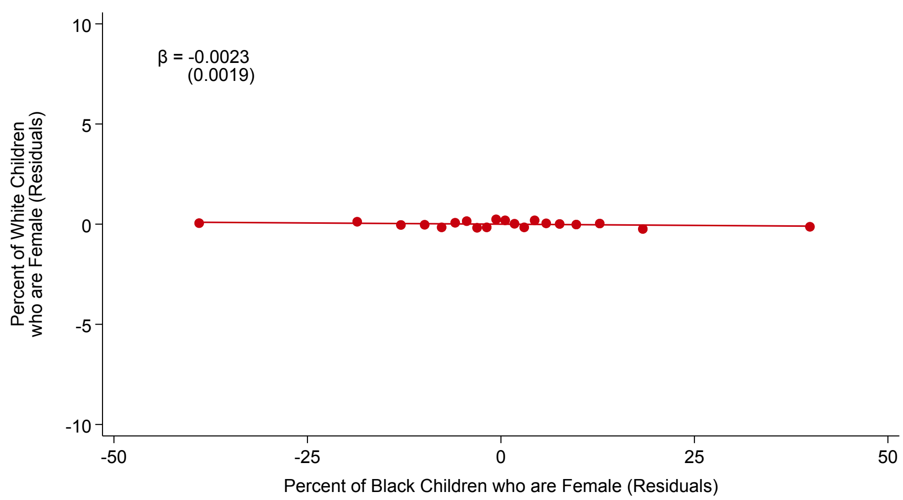
Notes: This figure shows the frequency of spouse age gaps separately by marriage type for females. For each type, the modal age gap is assigned a value of 100. For all other age gaps between -4 and 4, the height of the bar reflects the frequency of the age gap as a percentage of the frequency for the modal age gap. Panel A focuses on individuals from low- and high-income families. Interclass marriage is defined as marrying someone from the opposite income group. Panel B includes Black and White individuals, with interracial marriage defined as a Black-White marriage. In both panels, within-group marriage refers to marrying someone from the same class or race group. A pooled version of this figure, for both males and females, is shown in Figure A.4 Panel B; see figure notes for additional details.

FIGURE A.6: Within-Neighborhood Correlation of Sex Ratios Across Groups

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

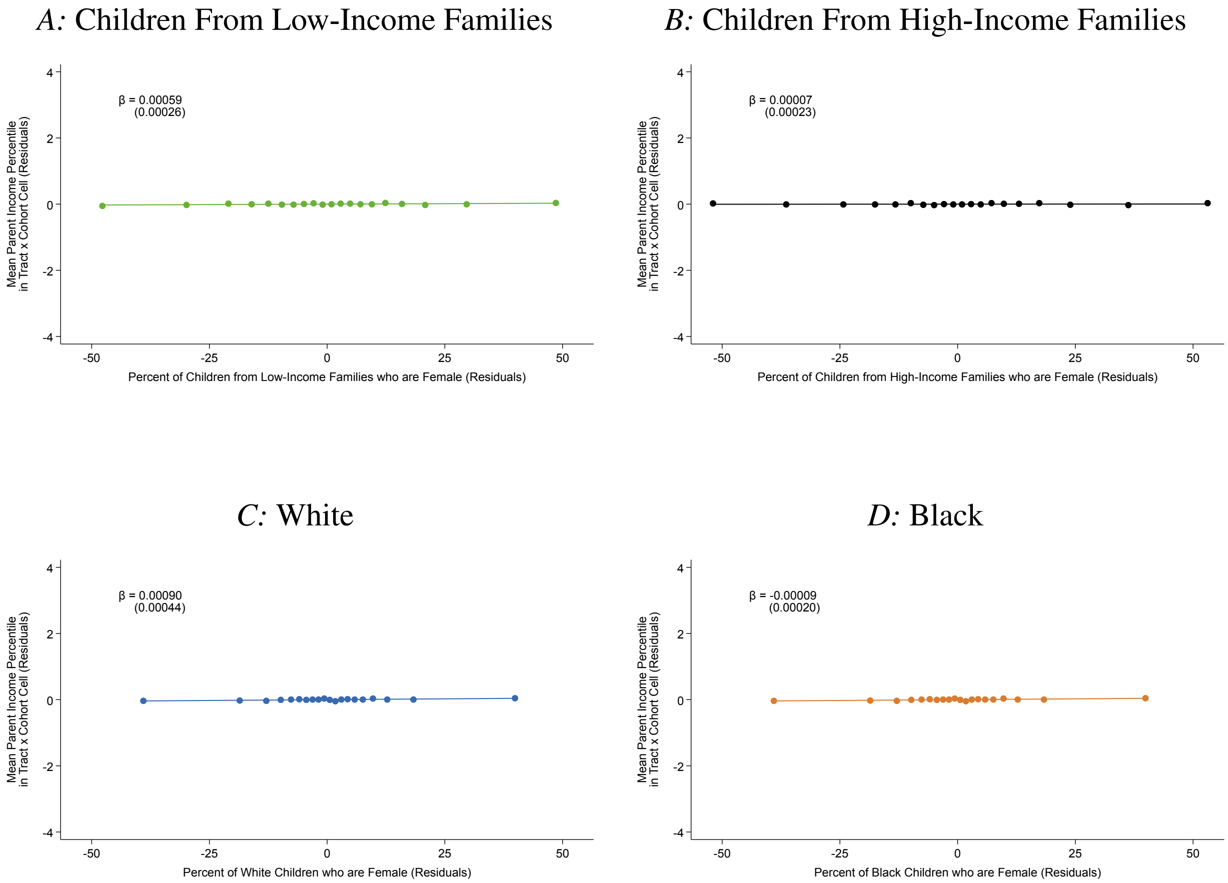


B: White and Black



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

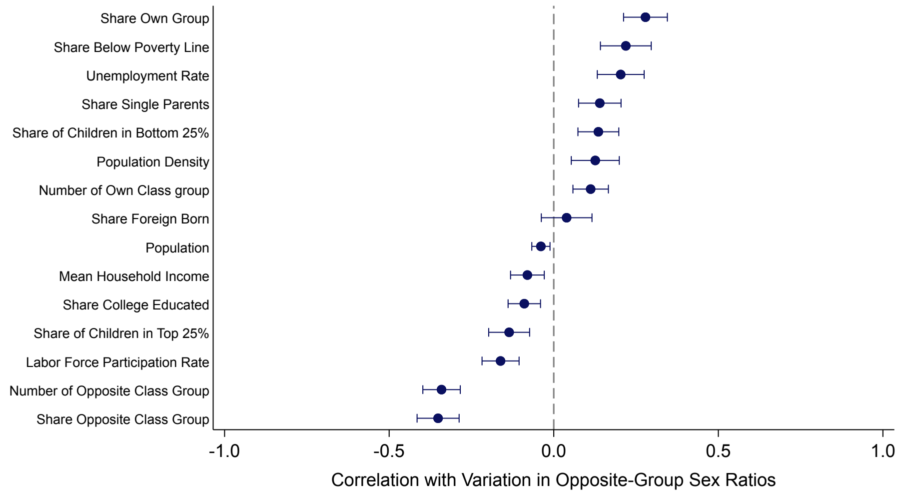
FIGURE A.7: Tract × Cohort Sex Ratio Balance on Mean Parental Income



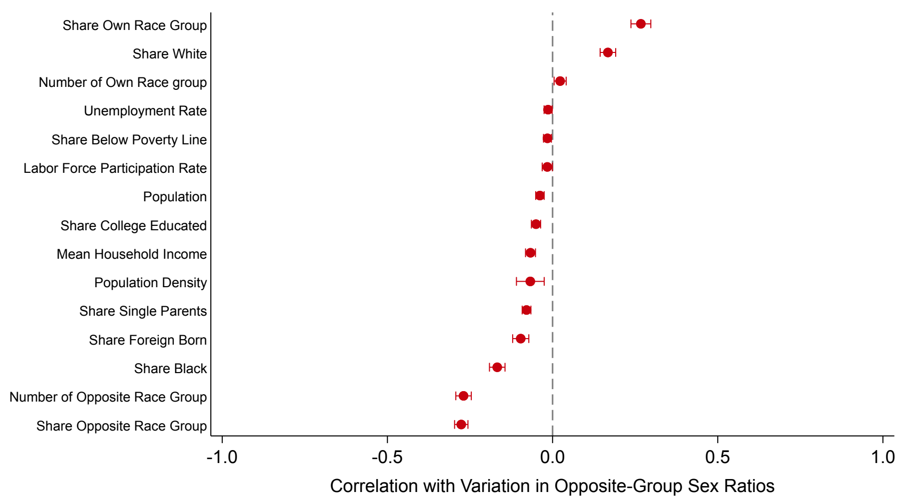
Notes: This figure presents binned scatter plots of the within- neighborhood relationship between sex-ratios for specific subgroups and mean parent income rank. We show results for individuals from low-income families (Panel A), high income families (Panel B), White (Panel C), and Black individuals (Panel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

FIGURE A.8: Characteristics of Places with Opposite-Group Sex Ratio Variation

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

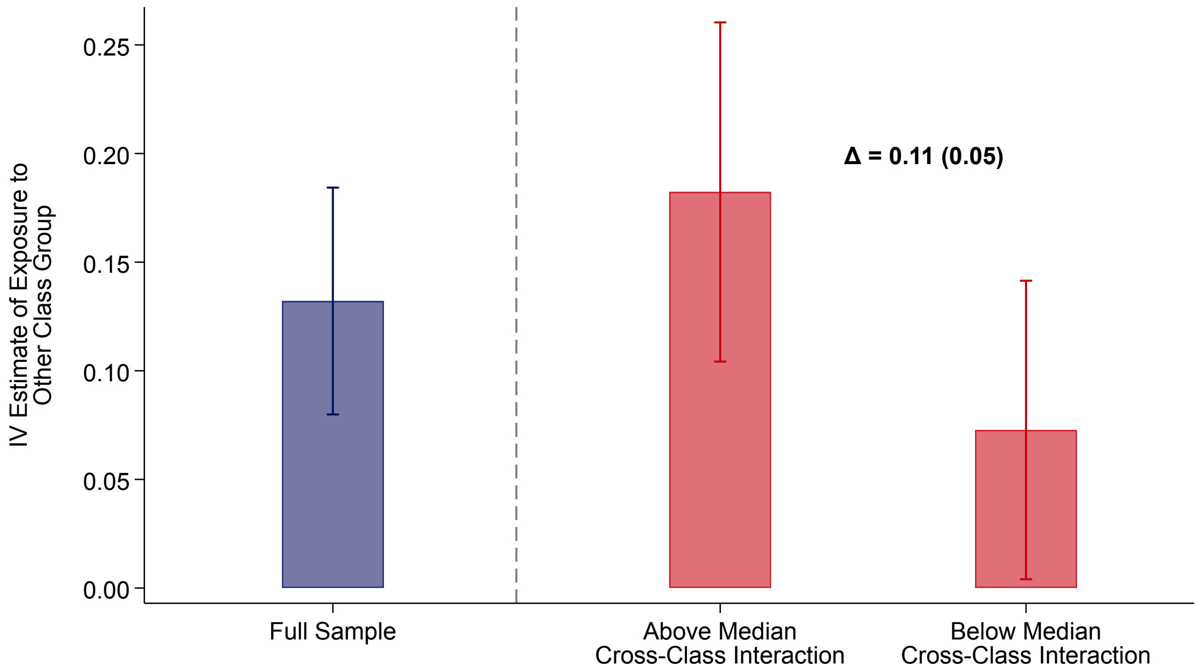


B: White and Black



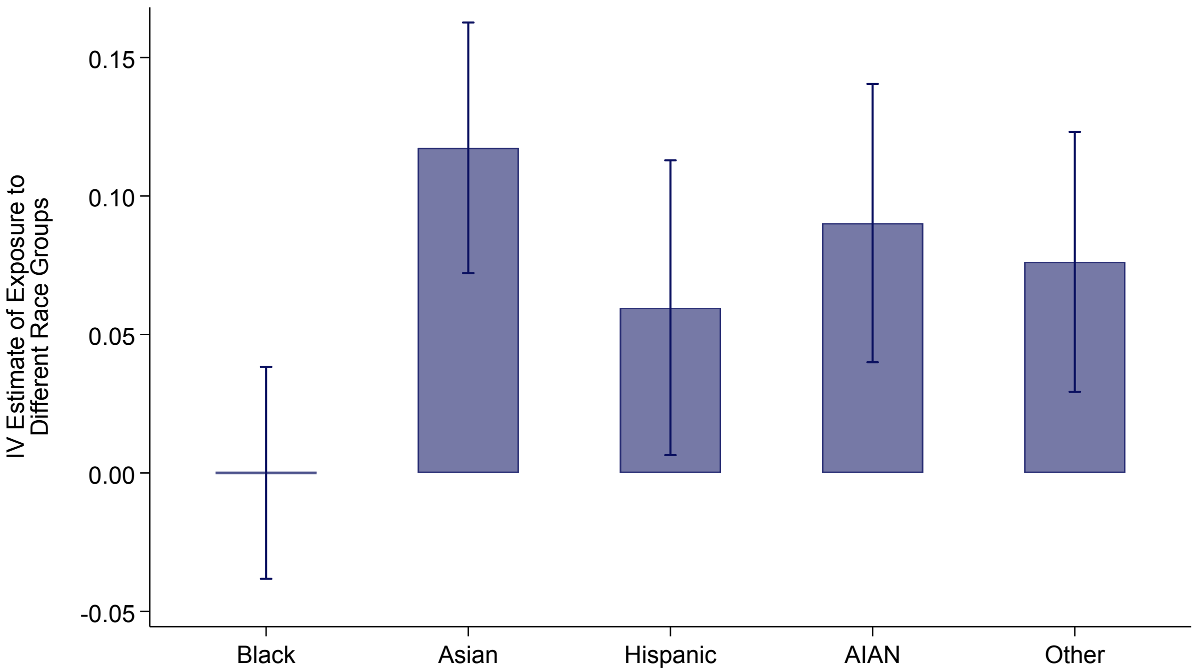
Notes: This figure plots the correlations between Census-tract characteristics and the amount of residual variation in our sex-ratio instrument. These correlations speak to which places have more or less *variation* in the instrument; they are *not* tests of whether the instrument's level is related to the same characteristics. Panel A shows results for individuals whose parents are in the bottom or top income quartile. The instrument is the share of opposite-sex neighbors who belong to the opposite class or race group in an individual's childhood neighborhood (age 18). We residualize the instrument 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. We then take the mean of the squared (residualized) instrument within a tract and correlate this measure with tract-level characteristics. Panel B repeats the analysis for White and Black individuals. Group-composition measures (e.g., shares of own and opposite groups) are calculated from our baseline sample, and all other tract attributes come from the 2000 Decennial Census as compiled by Chetty et al. (2025).

FIGURE A.9: IV Estimates of Cross-Class Exposure by County-Level Friending Bias in Cross-Class Friendship Formation



Notes: This figure presents our IV estimates for interclass marriage, separately for counties with above- and below-median levels of "friending bias," as defined by Chetty et al. (2022). Friending bias measures the extent of cross-class friendships on Facebook, relative to what would be expected given the class composition of the county. Counties with low friending bias exhibit higher rates of cross-class friendship conditional on exposure. The first bar reports the pooled estimate from Figure VIII. The difference between the coefficients in above- versus below-median counties is shown between the bars, alongside the associated standard error. For details on the specification procedure, see Section III.B and the notes to Table IV.

FIGURE A.10: IV Estimates of Cross-Race Exposure for White Individuals



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

FIGURE A.11: IV Estimates of Cross-Group Exposure Including All Sex Ratios

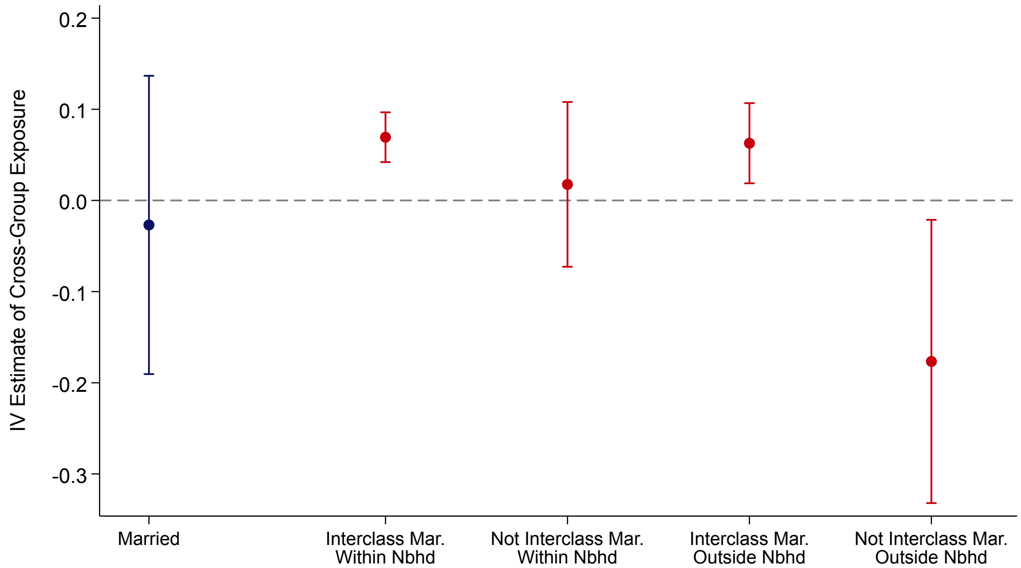
Notes: This figure examines the sensitivity of the baseline IV estimates for interclass and interracial marriage to the inclusion of sex ratios from other groups, beyond the focal group used in the baseline. The baseline estimates match the pooled results from Figure VIII. The "Other SRs as Controls" bars add the sex ratios for all class or race groups as controls in the IV regression. The "Other SRs as Instruments" bars include each group's sex ratio as an additional instrument, alongside the group-specific market tightness variable. For details on the specification, see Section III.B and the notes to Table IV.

FIGURE A.12: IV Estimates of Cross-Group Exposure by Sex

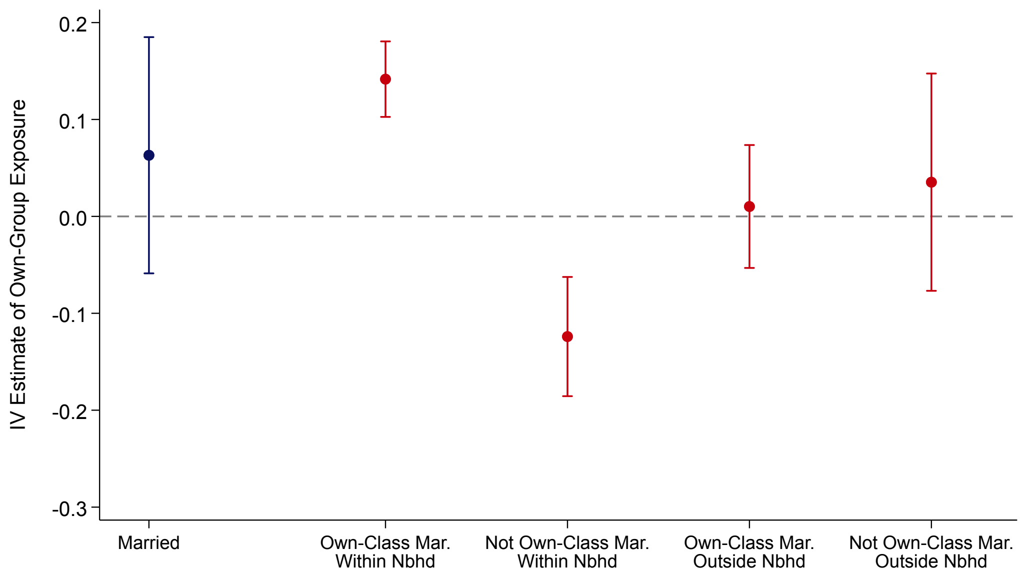
Notes: This figure summarizes the key IV results for interclass and interracial marriage, separately by sex. It shows the pooled estimates from Figure VIII, along with corresponding estimates for males and females. For details on the specification procedure, see Section III.B and the notes to Table IV. Additional robustness checks are provided in Table IV.

FIGURE A.13: Effect of Exposure on Spouse Characteristics

A: Effect of Exposure to Other Class Group



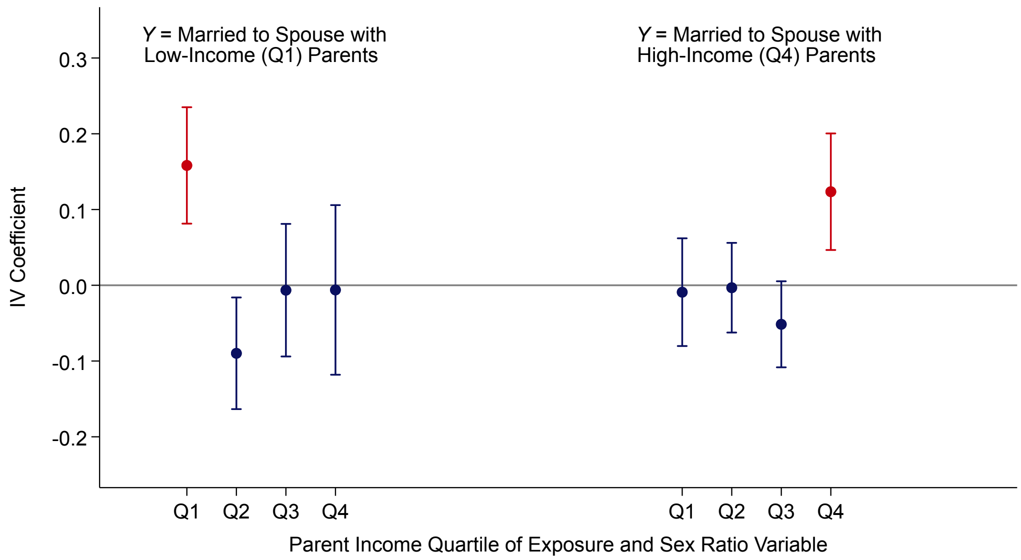
B: Effect of Exposure to Own Class Group



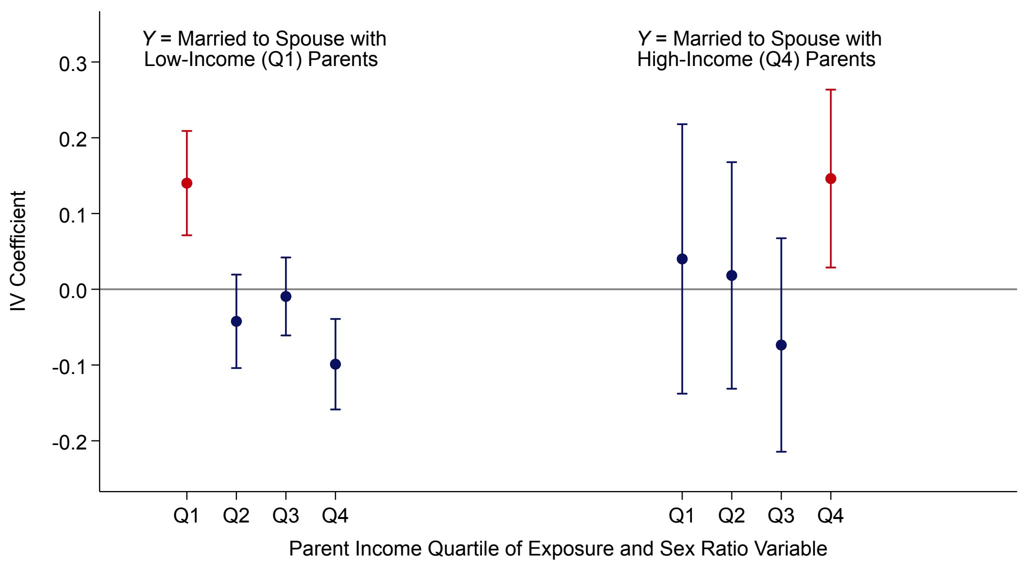
Notes: This figure examines how neighborhood-level variation in class-group exposure affects marriage outcomes by neighborhood and class group. The sample includes individuals whose parents are in the bottom or top income quartile. Panel A shows IV estimates (based on the baseline "Pooled" specification in Figure VIII) for mutually exclusive outcomes: marrying someone from the opposite or a different class group, within or outside a 50-tract radius of the individual's age 18 neighborhood. Panel B repeats the analysis using exposure to one's own class group as the treatment. In both panels, we also show the estimate for overall marriage, which is the sum of the four subgroup outcomes. For further details on the estimation procedure, see Section III.B.

FIGURE A.14: Impact of Exposure on Spouse Class by Quartile

A: Children with Low-Income (Q1) Parents

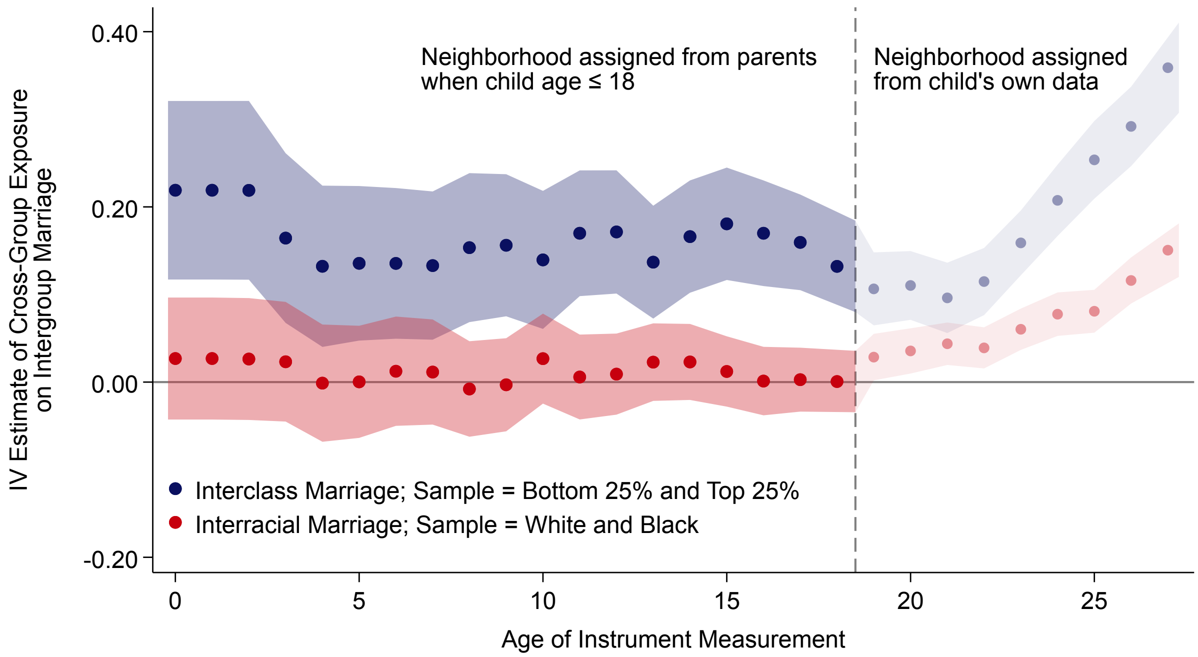


B: Children with High-Income (Q4) Parents



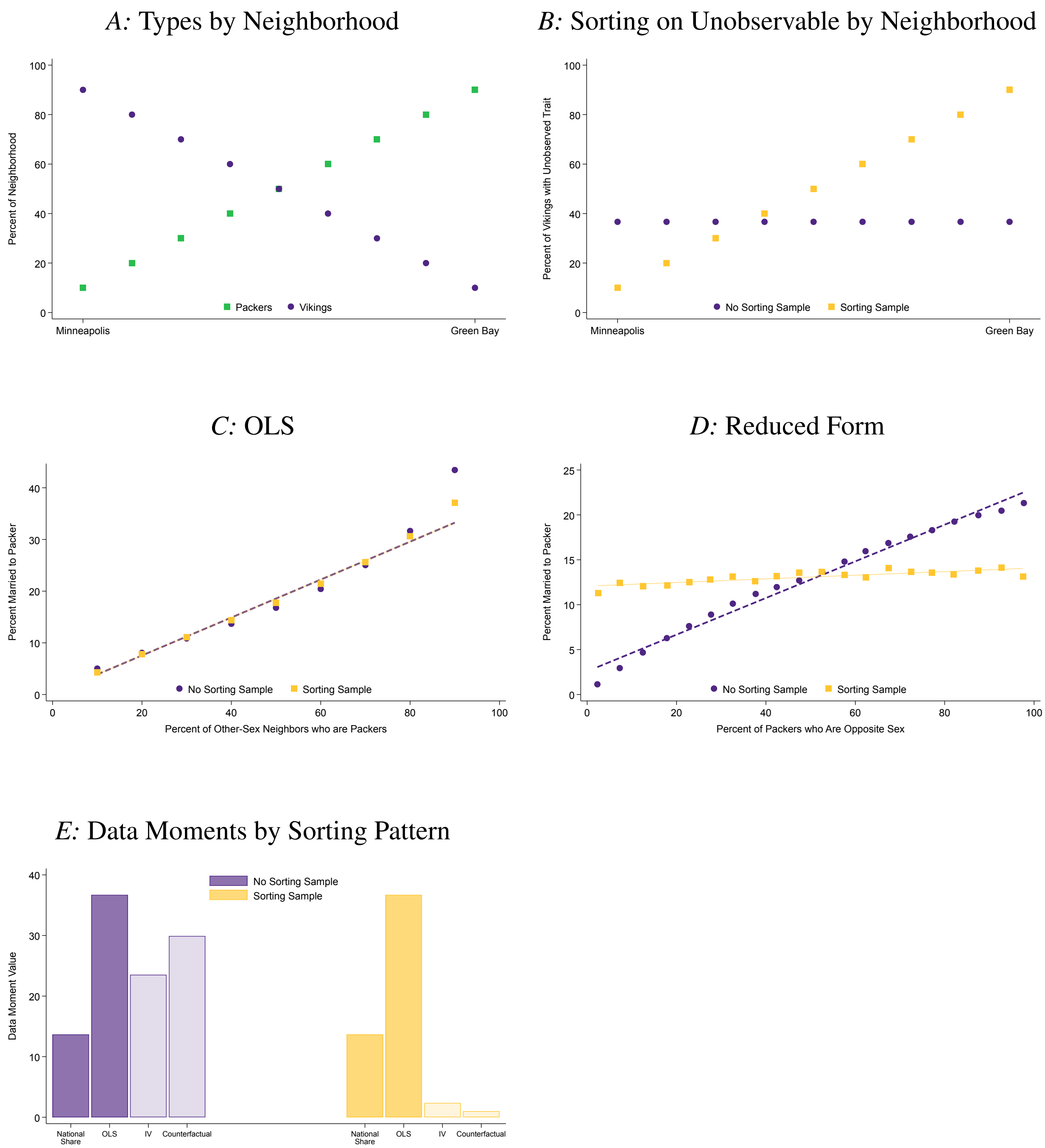
Notes: This figure presents a set of multivariate IV regressions for interclass marriage outcomes, analyzed separately for individuals from low parent income families (Panel A) and high parent income families (Panel B). In each case, the marriage outcome—indicated above the points—is regressed on four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

FIGURE A.15: IV Estimates by Age of Instrument Measurement



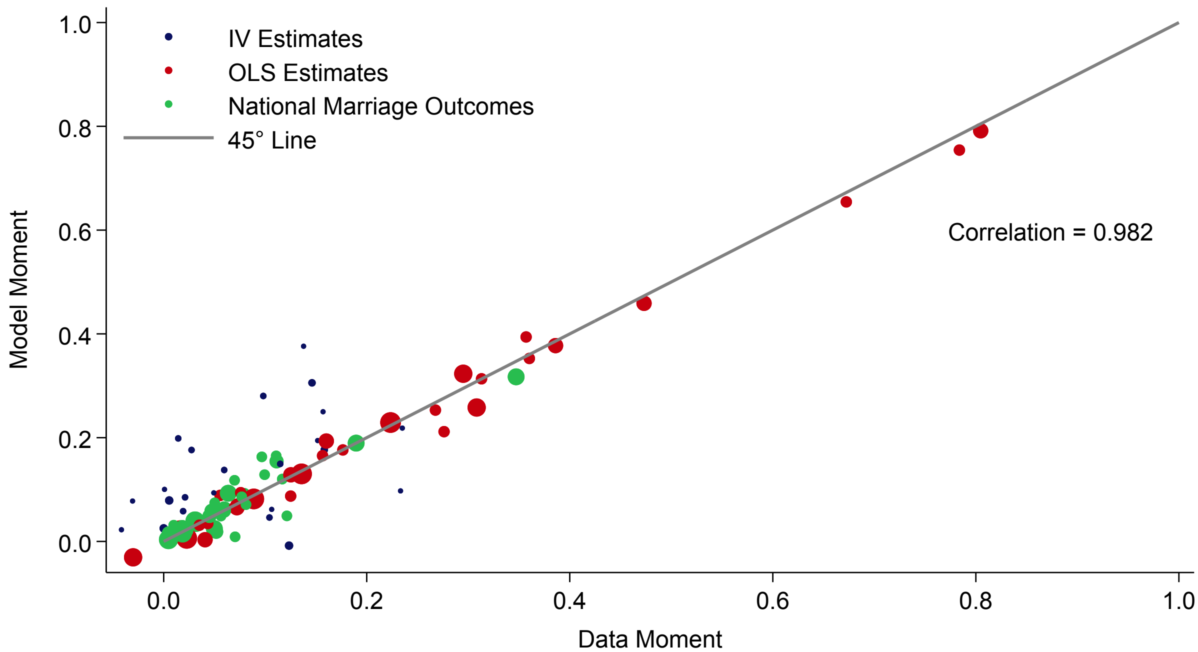
Notes: This figure presents IV coefficients from separate regressions where the instrument is measured at different ages throughout childhood, analyzed separately for class and race. The outcome in each regression is the fraction of individuals in an interclass (or interracial) marriage at age 30. The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

FIGURE A.16: Simulations With and Without Sorting on Unobserved Trait



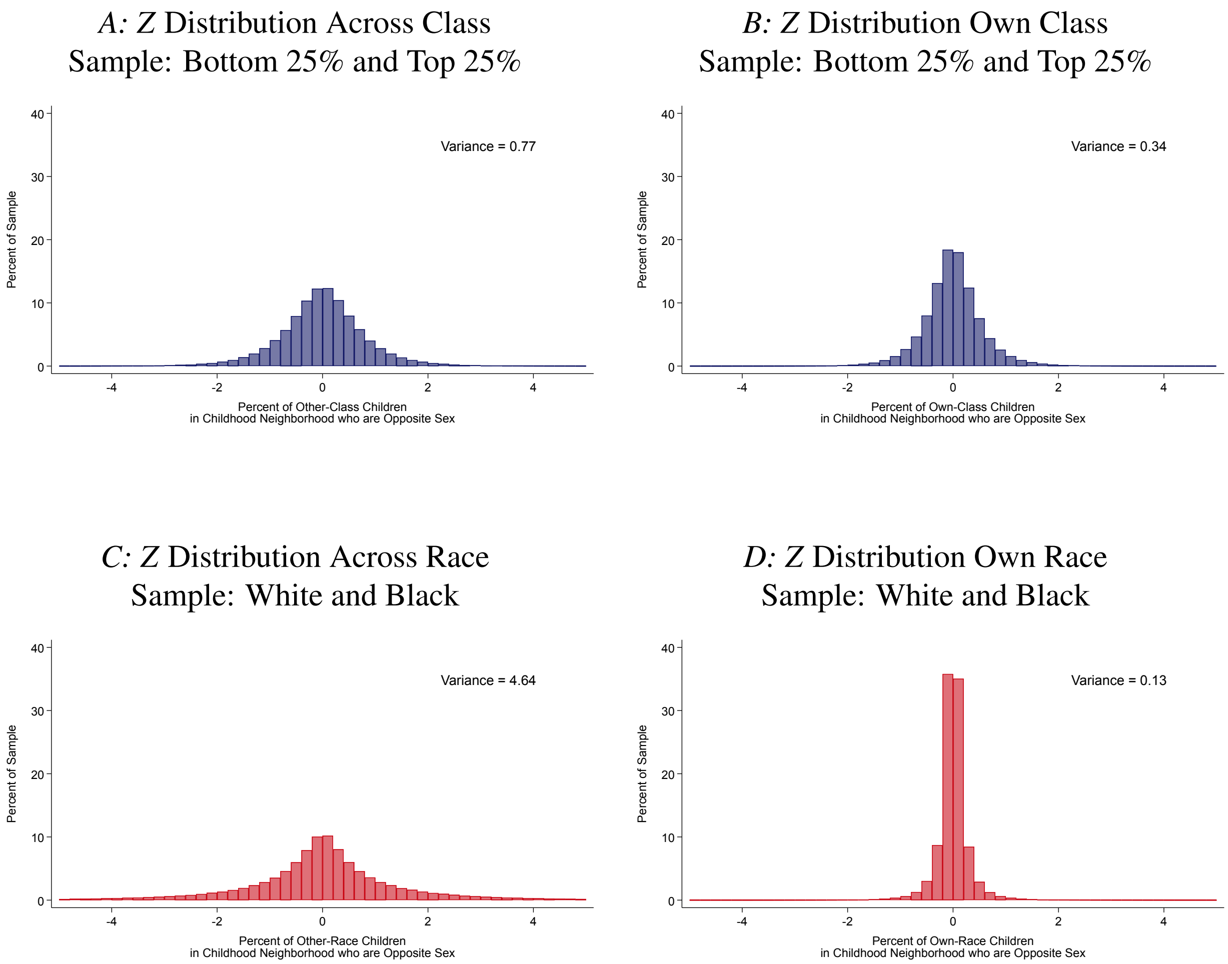
Notes: This figure presents a simulation from a simplified version of the spatial matching model discussed in Section IV. In the simulation, described in detail in Section E, there are two types, Packers and Vikings, and nine neighborhoods, from Minneapolis to Green Bay. Panel A shows the fraction of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

FIGURE A.17: Simulated Model Moments versus Observed Data Moments



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

FIGURE A.18: Distribution of Instrument Within vs. Across Groups



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