Migration, Segregation, and Interracial Marriage

Cameron Deal*

September 19, 2025

Abstract

Interracial marriage increased substantially in the late 20th century, concurrent with declines in regional and residential segregation. This paper studies whether these trends are related using two historical quasi-experiments and a stylized model of the marriage market. I use a shift-share instrument to find that the Great Migration increased interracial marriage in receiving communities, though this effect was absent in highly segregated cities. Using railroad density as an instrument, I also find that residential segregation decreased interracial marriage. Together, these findings indicate that internal migration and declining segregation contributed to the rise of interracial marriage in the non-Southern United States.

Keywords: Interracial Marriage, Segregation, Great Migration, Marriage Markets

JEL Codes: J12, J15, N32

*Harvard University. cdeal@g.harvard.edu. I thank Bhash Mazumder, Daniel Aaronson, William Collins, Christopher Carpenter, Michelle Marcus, Marcus Dillender, Steven Hill, Ben Harrell, Samuel Mann, Laura Nettuno, Santiago Deambrosi, Geoffrey Kocks, Arthur Wickard, Craig Sylvera, Claire Chen, Jay Garg, Ellora Derenoncourt, Marco Tabellini, Benny Goldman, Leo Saenger, Alex Chan, and especially Brian Beach for their comments, feedback, and guidance, as well as seminar participants at ASHEcon 2023, SEA 2023, and the Federal Reserve Bank of Chicago, and many anonymous referees. I also thank Ellora Derenoncourt for graciously sharing data.

1 Introduction

In 1950, the interracial marriage rate (IMR) between Black and white Americans was approximately 1.5 per 1000 marriages—50 years later, it was 18 times higher at 26.5 per 1000 marriages.[1] However, this is only a fifth of the IMR expected under random assignment to spouses.[2] Evidently, the US marriage market has made sizeable progress towards integration, yet substantial racial sorting in partner choice persists.[3]

Interracial marriage is a useful marker of social integration—perhaps the most intimate decision one makes is who their partner will be. Moreover, the segregation of the marriage market reinforces persistent racial disparities in economic outcomes by amplifying the intergenerational transmission of wealth within racial groups (Margo, 2016; Derenoncourt et al., 2023).[4] Additionally, interracial marriage serves as a valuable indicator of social integration between groups, particularly in contexts where other measures, like friendship networks, are unavailable. A substantial sociological and historical literature has traced the legal, cultural, and demographic determinants of interracial unions (see Section 3 for a brief review), and economists have documented the rise of interracial marriage and used structural models of the marriage market to study racial homophily (e.g., Wong (2003); Fryer (2007); Chiappori et al. (2016)). This paper builds on this literature and examines how two historical forces—regional migration and residential segregation—have shaped Black–white intermarriage and social integration at scale.

Between 1940 and 1970, regional racial segregation declined significantly. In 1940, 50% of African Americans would have needed to move across Census regions to equalize population distribution with whites, but by 1970, this figure had dropped to 25%, largely due to the Black Southern exodus of the Great Migration (see Figure 1). Additionally, urban residential segregation reached a peak in 1960 and started falling thereafter. Thus, interracial marriage rates increased dramatically following decreases in regional segregation and concurrent with drops in urban residential segregation. These relationships are also borne out in the cross-section—in 2000, migration-induced Black population increases are positively associated with IMRs, and residential segregation is negatively associated with IMRs (see Figure 2). I build on this time series and cross-sectional evidence by using quasi-experimental designs to assess how historical forces, such as the Great Migration and residential segregation, have influenced the prevalence of interracial marriage in the

[1] I define the interracial marriage rate (IMR) to be the share of Black same-race marriages, Black and white marriages, and white same-race marriages that are between Black and white spouses. For more detail, see Section 3.

[2] This rate is 132.3 per 1000 marriages in 2000, calculated using the racial composition of the population. For more detail, see Section 3.

[3] I treat random matching as a descriptive counterfactual to quantify racial sorting in partner choice. I do not take a position on what the 'right' level of intermarriage should be.

[4] As Margo (2016) highlights, low levels of racially mixed households can, in combination with high intergenerational transmission and large initial gaps of human capital, create "intergenerational drag," making racial inequality persistent. While economic and health disparities between Black and white Americans have reduced, they remain high—the white-to-Black wealth ratio remains at 6 to 1 and Black Americans live on average 3.6 fewer years than their white counterparts (Schwandt et al., 2021; Derenoncourt et al., 2023). Marriage can be one pathway to reducing these inequalities.

United States.

First, I outline a stylized model of the marriage market and derive several comparative statics that I test in the data.[5] Potential matches are generated according to a meeting technology where cross-race contact is increasing in minority population share and decreasing in segregation frictions. I consider segregation in three roles: a wedge that reduces cross-race contact and interracial marriage directly (tested by the direct effects of residential segregation), a mediating friction that limits the effects of population change (examined through heterogeneity in the effects of the Great Migration), and an endogenous response to population change (implicitly measured via the total effect of the Great Migration). Conditional on meeting, potential matches are accepted according to acceptance probabilities that are decreasing in homophily preferences and the social costs to interracial marriages. I use this model to derive several predictions about the effects of migration and segregation on interracial marriage: 1) Raising the Black population share increases white outmarriage,[6] decreases Black outmarriage, and has ambiguous effects on IMR; 2) Residential segregation decreases all measures of interracial marriage; 3) Segregation dampens the effects of Black population increases on interracial marriage; 4) Declining social stigma amplifies the effects of Black population increases on interracial marriage; 5) Increasing Black population share increases white non-marriage.

Then, I use the Second Great Migration as a large-scale natural experiment to assess how minority in-migration and Black population increases affected social integration in receiving communities. While the Great Migration increased exposure between Black and white Americans by significantly reducing regional segregation, white Americans in receiving communities acted through labor markets, local governments, and residential choices to limit the progress of Black migrants, potentially offsetting the effects of increased exposure (Derenoncourt, 2022; Boustan, 2010, 2016). Previous work has documented increases in racial animus and white native-immigrant marriage in response to both waves of the Great Migration, indicating backlash (Derenoncourt, 2022; Fouka et al., 2022). Indeed, one especially important manifestation of these actions is increases in residential segregation in response to Black in-migration (Massey and Denton, 1993). Whether these endogenous responses outweighed the increases in exposure is ultimately an empirical question. As a result, the aggregate effect of in-migration on social integration and interracial marriage is unclear. I evaluate the effects of increasing Black population shares using a shift-share instrument constructed with the Second Great Migration (Derenoncourt, 2022). I find that a 10-percentile increase in Black population change[7] causes 1.32 additional interracial marriages per 1000 marriages (10% increase) in 2000. In the aggregate, exogenous increases in minority in-migration increase social integration, suggesting that endogenous responses and mediating frictions are dominated by exposure effects. Additionally, effects grow over time in

[5] I build heavily on previous work studying search and matching in marriage markets Wong (2003); Choo and Siow (2006).

[6] I define the outmarriage rate as the share of Black individuals marrying a white spouse (and vice versa).

[7] Or a 3.5 percentage point increase in Black population share—see Figure 2 of Derenoncourt (2022).

the series and within birth cohort, consistent with the model prediction of evolving social norms amplifying the effects of Black population increases.

However, these aggregate positive effects mask substantial heterogeneity. Receiving areas with above-median residential segregation saw no increase in interracial marriage in response to Black population change driven by the Great Migration, in accordance with the model prediction of segregation limiting the effects of Black population increases. This heterogeneity suggests that mediating frictions, particularly residential segregation, may have constrained the social integration gains associated with the Great Migration. In addition to analyzing raw IMRs, I construct several additional outcomes that can account for population racial composition and decompose effects into mutually exclusive categories such as outmarriage, same-race marriage, other-race marriage, and single status. First, I construct a marital integration outcome to account for differences in population racial composition (for example, comparing a city that is 50% Black and 50% white to one that is 5% Black and 95% white). This outcome can be thought of as the fraction of interracial marriages expected under random assignment that are observed in the data.[8] When using this outcome, I find that a 10-percentile increase in Black population change causes a 0.15-point decrease in marital integration in 1990—seemingly in contrast to the positive effects on interracial marriage. This pattern can be attributed to the fact that while interracial marriage increases in cities that receive more Great Migration migrants, this increase is an order of magnitude lower than expected given the magnitude of Black population change. I interpret this difference in magnitude between expected and realized increases in interracial marriage as further evidence for substantial mediating frictions and endogenous responses in response to minority in-migration. Additionally, I decompose effects on IMR by race and find that the Great Migration increased outmarriage and decreased same-race marriage for white individuals, while it had opposite effects for Black individuals, in line with model predictions. In the aggregate, the prevalence of interracial marriage increased because the white population is much larger. Interestingly, I also confirm the model prediction that Black population increases might produce a higher white nonmarriage rate, suggesting some white individuals were induced into remaining unmarried when there was a large influx of Black migrants into their geographic area.

In a complementary analysis, I examine the causal effect of residential segregation on social integration and interracial marriage. Residential segregation greatly reduces the opportunities for individuals of different races to interact, and has negative effects on Black-white inequality and Black economic opportunity, which

[8] I focus on marriage as my primary outcome rather than a broader definition of interracial relationships, in part because marriage is consistently recorded in the population Censuses for the time period I study. The 1990 Census added "unmarried partner" as an option for relationship status, but there is concern that this measure will not cover many cohabiting couples who would instead prefer terms like "boyfriend" or "fiancè" (Manning and Smock, 2005; Kennedy and Fitch, 2012). Additionally, historians have argued that the formality of marriage conveys something substantively different about social integration in relation to more illicit or informal relationships between races, which occur frequently throughout history (Hodes, 1999b). To the extent this affects my estimates, it suggests my findings are a lower bound for the true effects of the Great Migration and residential segregation on interracial relationships more broadly.

could negatively impact social integration and interracial marriage (Ananat and Washington, 2009; Ananat, 2011; Chyn et al., 2022). I instrument for contemporary segregation using the placement of historical railroad tracks. To do so, I construct novel measures of railroad density at the commuting zone level using historical railroad placement data from Atack (2016). I find that a 1 standard deviation increase in residential segregation causes 6.77 (41%) less interracial marriages per 1000 marriages and a decrease of 0.47 points of marital integration in 2000. When I decompose the effects of segregation, I find increases in same-race marriage among both Black and white individuals, and these are offset by decreases in outmarriage and other-race marriage among Black individuals, in accordance with model predictions. These results are related with those concerning the Great Migration—confirming that residential segregation has a negative causal effect on interracial marriage suggests that some endogenous responses will take place in response to minority in-migration given previous literature establishing the segregation response to the Great Migration (Massey and Denton, 1993).[9] Additionally, I replicate these results using the original Ananat (2011) sample and find similar estimates (See Appendix D).

In sum, I find that both trends—decreasing regional and residential segregation—contributed to the increase in interracial marriage in the second half of the 20th century.[10] I use the Great Migration to study the social response to minority in-migration and find aggregate positive effects on social integration, suggesting that exposure effects overwhelmed mediating frictions and endogenous responses in receiving communities. However, evidence on heterogeneity by segregation of receiving areas suggests that some mediating frictions were present. Additionally, I study the effects of residential segregation on interracial marriage and find negative effects, confirming the cross-sectional relationship seen in the data.[11] My results largely confirm the predictions of my stylized model, suggesting that the framework accurately captures the major forces determining interracial marriage.

My findings contribute to the economic literature on interracial marriage by examining the impact of historical forces on the rise of interracial marriage. Previous studies have documented broad increases in

[9]Additionally, Ananat (2011) highlights that the placement of railroad tracks only predicts residential segregation after the Great Migration, as the ease of segregation responses to the Great Migration is what gives the instrument its predictive power. As long as Northern historical railroad track placement is orthogonal to exogenous increases in Black population share stemming from Southern economic shocks, this relationship should not pose a challenge to my identification strategies.

[10]One potential concern is that concurrent legal changes might confound the relationships I establish, as legal access to interracial marriage was greatly expanded in the mid-20th century. Some claim this changing legal landscape lifted previously binding constraints, allowing interracial couples who previously desired marriage to realize it (Moran, 2003; Newbeck, 2008). Relatedly, studies of access to same-sex marriage find increases in marriage for same-sex couples following expansion of legal access (Dillender, 2014). In related work, I find little evidence that permanent repeals of anti-miscegenation statutes affect the rate of either interracial marriage or marital integration. I leverage the staggered repeal of anti-miscegenation laws across states and time to estimate the effects of legal access to interracial marriage in difference-in-difference and event study frameworks and find consistent evidence of null effects. Additionally, I can rule out effects of the size recorded in the same-sex marriage literature. See Deal (2024) for more details and results.

[11]Because my identification strategies rely on exogenous Black population shifts and variation in residential segregation that are unique to the non-South, my findings are limited to interracial couples and social dynamics outside the South, where there is also a rich history of interracial relationships (Mills, 1981). In related work, I show that trends in the prevalence of interracial marriages are qualitatively similar across regions, suggesting that understanding trends outside the South may have implications for national trends (Deal, 2024).

interracial marriage over time and used structural models to analyze the determinants and welfare benefits of interracial marriage (Koh, 2024; Fryer, 2007; Wong, 2003). Additionally, other studies have demonstrated how immigration and institutional change can shape the marriage market (Fouka et al., 2022; Ager et al., 2021; Carlana and Tabellini, 2018). I build on this work by showing that the broader trends of decreasing regional and residential segregation increased interracial marriage rates and social integration.

Another strand of research highlights how exposure to people of different races impacts interracial marriage. For instance, exposure to racially diverse classmates has been shown to increase interracial relationships later in life, while studies on school desegregation reveal mixed effects on interracial marriage and mixed-race births (Merlino et al., 2019; Shen, 2018; Gordon and Reber, 2017). Concurrent work explores how marriage market tightness and residential segregation shape interracial marriage, finding that exposure to diverse neighbors can increase such marriages but with limited aggregate effects (Goldman et al., 2023). This work is better suited to capture pure exposure effects, while my approach captures broader equilibrium effects within areas.[12]

I also contribute to the literature on the Great Migration, focusing on its second wave (1940–1970). This migration reshaped destination cities' demographics, economics, and politics (Boustan, 2009; Calderon et al., 2022; Shi et al., 2022). I find that while the Great Migration increased interracial marriage, its effects were muted in highly segregated areas, reflecting barriers to broader integration. These findings align with previous research documenting economic gains for migrants but also adverse impacts on upward mobility and urban racial dynamics (Boustan, 2010; Derenoncourt, 2022; Collins, 2021). Additionally, I evaluate the Great Migration shift-share design using recently-developed econometric tools, finding strong support for its validity. (Borusyak et al., 2021; Goldsmith-Pinkham et al., 2020; Borusyak et al., 2024).

Lastly, my work intersects with prior literature on residential segregation, which predominantly examines economic and political outcomes. Higher segregation has been linked to reduced upward mobility, increased poverty, and widening racial disparities in crime and political efficacy (Collins and Margo, 2000; Cutler et al., 1999; Ananat, 2011; Chyn et al., 2022; Ananat and Washington, 2009; Cox et al., 2022). I instead focus on segregation's impact on social integration, highlighting its role in limiting interracial marriage.

The remainder of this paper proceeds as follows. In Section 2, I develop a stylized model of the marriage market and derive comparative statics. In Sections 3 and 4, I describe the setting and data, respectively. I

[12]These studies are much better positioned to identify pure exposure effects, as they rely on variation at the individual or neighborhood level and thus do not have to account for endogenous responses that could also affect interracial marriage. However, they are also limited by their reliance on specific spheres of interaction for identification—for example, while exposure to more Black peers may impact interracial marriage, education is only one context wherein individuals meet partners or interact with people across race. They provide motivation for why the trends presented in Figure 1 may be causal, but cannot account for responses to increasing exposure that occur in the aggregate. The present study offers a historical and quasi-experimental approach to measuring the effect of minority in-migration on social integration that can capture equilibrium effects within commuting zones.

outline my empirical strategy in Section 5 and present the main results and robustness tests in Section 6. Section 7 concludes.

2 Stylized Model of the Marriage Market

I outline a stylized model of the marriage market featuring: 1) a meeting technology in which cross-race contact is increasing in the share of opposite-race individuals and decreasing in racial segregation; 2) preferences for interracial marriage conditional on contact are shaped by homophily and slow-moving norms. I use this model to derive several comparative statics that predict the effects of the Great Migration and residential segregation on interracial marriage.

2.1 Meeting Technology

Consider a commuting zone (CZ) where $x \in [0,1]$ share of the population is Black and $1-x$ is white.[13] Additionally, there is a segregation friction $s \in [0,1]$ that determines the difficulty of meeting conditional on population shares—in my setting, this could be proportional to a dissimilarity index or other residential segregation measure. I can also allow segregation to respond to Black population share as $s(x)$. I re-derive my comparative statics on outmarriage with this change in Appendix B.

Agents meet potential partners according to a simple contact technology, where q_{WB} is the probability that a meeting for a white individual is with a Black partner, and q_{BW} is the converse, defined below:

$$q_{WB}(x,s) \equiv (1-s)\,x, \qquad q_{BW}(x,s) \equiv (1-s)\,(1-x),$$

Within-race meeting probabilities are defined as $q_{WW} = 1 - q_{WB}$ and $q_{BB} = 1 - q_{BW}$. These formulas capture the intuition that the probability of cross-race contact is increasing in the share of the population that is of the opposite race and decreasing in the segregation of the commuting zone. This structure captures the "contact" channel emphasized in social psychology (e.g. Allport (1954)) and in economic matching with search.[14] Next, I consider preferences for interracial marriage conditional on cross-race contact.

2.2 Preferences and Acceptance Probabilities

I define preferences that encode homophily and social norms against interracial marriage, in line with the historical literature summarized below. A potential same-race match yields deterministic surplus $v_S > 0$. In

[13] I abstract away from other races for simplicity and consider only individuals who are part of the effective marriage market.
[14] See Becker (1973), Wong (2003), Fryer (2007), and Chiappori et al. (2017) for classic and modern analyses of marriage markets with sorting/search; for norm updating via social exposure see, e.g., Bursztyn et al. (2020).

contrast, an interracial match yields $v_I(\kappa_t) = v_S - h - \kappa_t$, where $h \geq 0$ measures preferences for homophily and $\kappa_t \geq 0$ is a time-varying stigma (social-norm) cost to marrying across race. Let ε be an idiosyncratic i.i.d. taste shock with cumulative distribution function F.[15] A proposed match is accepted if $v + \varepsilon \geq 0$. These matches arise from the meeting technology outlined above.

I define the acceptance probabilities of a proposed match as:

$$a_S \equiv \Pr\{\varepsilon \geq -v_S\} = 1 - F(-v_S) \in (0,1), \qquad a_I(h,\kappa_t) \equiv \Pr\{\varepsilon \geq -v_I(\kappa_t)\} = 1 - F(h + \kappa_t - v_S) \in (0,1),$$

with $\frac{\partial a_I}{\partial \kappa_t} < 0$ and $a_I(h,\kappa_t) < a_S$ whenever $h + \kappa_t > 0$. These conditions suggest that the acceptance probability of interracial matches is decreasing in stigma and that acceptance probabilities for proposed interracial matches will be lower than same-race matches whenever the sum of stigma costs and homophily costs is positive.

2.3 Outmarriage Propensities and Observed IMR

For each race $g \in \{W, B\}$, let $\text{Out}_g(x, s, \kappa_t)$ denote the probability that an individual of group g marries out of their race. Under the meeting/acceptance structure above,[16]

$$\text{Out}_W(x,s,\kappa_t) = \frac{q_{WB}(x,s)\, a_I(h,\kappa_t)}{q_{WB}(x,s)\, a_I(h,\kappa_t) + [1 - q_{WB}(x,s)]\ a_S}, \tag{1}$$

$$\text{Out}_B(x,s,\kappa_t) = \frac{q_{BW}(x,s)\, a_I(h,\kappa_t)}{q_{BW}(x,s)\, a_I(h,\kappa_t) + [1 - q_{BW}(x,s)]\ a_S}. \tag{2}$$

These probabilities offer simple predictions about how outmarriage will be shaped by racial composition, segregation, and time-varying norms.

Additionally, I can aggregate these outmarriage rates into an interracial marriage rate (IMR) that closely corresponds to the measure I use in the data. Let $\mu_o(x, s, \kappa_t)$ denote the observed share of Black–white marriages among all Black–white and same-race Black/white marriages in a given commuting zone. With balanced sex ratios, the individual-level share is simply a weighted average of the outmarriage rates, with weights corresponding to racial population shares. Intuitively, interracial marriage is driven more by white outmarriage when white individuals are a large share of the population.

$$\mu_o(x,s,\kappa_t) \;=\; \frac{x\,\text{Out}_B(x,s,\kappa_t) + (1-x)\,\text{Out}_W(x,s,\kappa_t)}{2}. \tag{3}$$

[15]This could represent idiosyncratic preferences for specific partners irrespective of race. I require that F is strictly increasing and continuous.

[16]See Appendix B for derivation and discussion.

Additionally, I can construct a benchmark of random matching under balanced sex ratios—the expected interracial share conditional on random assignment within the CZ—as:

$$\mu_r(x) \;=\; 2x(1-x) \tag{4}$$

This benchmark corresponds to model predictions when acceptance rates are equal for same-race and interracial proposed matches and there are no segregation frictions, so matching is driven only by population composition. Finally, I can use the ratio of observed and random interracial marriage rates to adjust for population composition and measure "marital integration":

$$m(x, s, \kappa_t) \;\equiv\; \mu_o(x, s, \kappa_t)/\mu_r(x). \tag{5}$$

I outline the measurement and precise definitions of these quantities in Section 4. Additionally, I relax the assumption of balanced sex ratios. Next, I examine the comparative statics this stylized model generates and take those predictions to the data with my quasi-experiments.

2.4 Comparative Statics

I derive several qualitative predictions about the effects of increasing Black population share and segregation below. I can denote $a_I = a_I(h, \kappa_t)$ and $q_{WB} = (1-s)x$, $q_{BW} = (1-s)(1-x)$. Define $D_W \equiv q_{WB}a_I + (1 - q_{WB})a_S$ and $D_B \equiv q_{BW}a_I + (1 - q_{BW})a_S$. These quantities are the denominators in the outmarriage rate and represent overall probabilities that a proposed match to an individual is accepted, regardless of whether it is interracial or same-race.

Increasing Black population share increases white outmarriage and decreases Black outmarriage: First, I examine the partial derivative of outmarriage rates Out_W and Out_B with respect to Black population share x:

$$\frac{\partial \text{Out}_W}{\partial x} = \frac{(1-s)a_I a_S}{D_W^2}; \quad \frac{\partial \text{Out}_B}{\partial x} = -\frac{(1-s)a_I a_S}{D_B^2}$$

These partial derivatives suggest that increasing the Black population share will increase white outmarriage rates while decreasing Black outmarriage rates, implying the Great Migration, which caused large increases in Black population shares across the non-Southern US, may have increased the propensity for interracial marriage for white individuals while decreasing interracial marriage among Black individuals (as their effective same-race marriage market expanded). When I allow for segregation responses, if $s'(x) > 0$ (in-migration increases segregation), this dampens and can even reverse the effect on white outmarriage and amplifies the effects on Black outmarriage.[17]

Increasing Black population share can increase or decrease IMR and marital integration: Next, I can aggregate these effects on outmarriage and derive predictions about effects of increasing Black population share on the observed IMR and marital integration:

$$\frac{\partial \mu_o}{\partial x} = \frac{\text{Out}_B - \text{Out}_W}{2} + \frac{x\frac{\partial \text{Out}_B}{\partial x} + (1-x)\frac{\partial \text{Out}_W}{\partial x}}{2}$$

The first term is a composition effect, where increasing the Black population share puts more weight on the Black outmarriage rate, while the second is a weighted average of the outmarriage effects, weighted by the population racial composition. Thus, the effect on observed interracial marriage is positive if the increase in white outmarriage is large enough to overwhelm the decrease in Black outmarriage (and compositional shifts).

For marital integration, I find that the partial effect is:

$$\frac{\partial m}{\partial x} = \frac{\frac{\partial \mu_o}{\partial x}\mu_r - \frac{\partial \mu_r}{\partial x}\mu_o}{\mu_r^2}$$

The sign of this object is ambiguous—as outlined above, $\frac{\partial \mu_o}{\partial x}$ has ambiguous sign. Additionally, $\frac{\partial \mu_r}{\partial x}$ is of ambiguous sign—if $x > 0.5$ (the Black share of the population exceeds 50%), then increases in the Black population share will decrease the interracial marriage rates expected under randomization. In my sample, the Black population share rarely exceeds 50%, so the sign of this partial should usually be positive. Thus, if there is a large mechanical increase in μ_r, this can outweigh an increase in observed interracial marriage and drive down marital integration.

Segregation decreases outmarriage, IMR, and marital integration: Additionally, I examine the direct effects of segregation on outmarriage rates:

$$\frac{\partial \text{Out}_W}{\partial s} = \frac{-xa_Ia_S}{D_W^2}; \quad \frac{\partial \text{Out}_B}{\partial s} = \frac{-(1-x)a_Ia_S}{D_B^2}$$

Both of these partial derivatives will be weakly negative, suggesting segregation will decrease outmarriage rates for both groups. Additionally, I can examine effects on the observed IMR, which will simply be a weighted average of the outmarriage effects, and marital integration, which will deflate by the random

[17] See Appendix B for more details.

matching benchmark:

$$\frac{\partial\mu_o}{\partial s} = \frac{x}{2}\frac{\partial \mathrm{Out}_B}{\partial s} + \frac{1-x}{2}\frac{\partial \mathrm{Out}_W}{\partial s}; \quad \frac{\partial m}{\partial s} = \frac{1}{\mu_r}\frac{\partial \mu_o}{\partial s}$$

Again, both of these partial derivatives will be weakly negative. In sum, segregation decreases interracial marriage for all groups and all measures in this model.

Segregation moderates the effects of increasing Black population share: I can also calculate the cross-partial of segregation on the outmarriage effects of increasing Black population share:

$$\frac{\partial^2 \mathrm{Out}_W}{\partial x\,\partial s} = -\,\frac{a_I a_S}{D_W^2} \;+\; \frac{2(1-s)x\, a_I a_S\,(a_I - a_S)}{D_W^3} \;<\; 0,$$

The first partial derivative is negative for interior values (as $a_I < a_S$), meaning that partial effect of increasing Black population share on white outmarriage is decreasing in segregation. This suggests that segregation will mute the effects of increasing the Black population share.

Evolving norms increase interracial marriage and the effects of Black population share: For $g \in \{W, B\}$,

$$\frac{\partial \mathrm{Out}_g}{\partial \kappa} = \frac{q_g(1-q_g)\,a_S\,a_I'(\kappa)}{D_g^2} \;<\; 0.$$

Let stigma evolve via contact-based social learning $\kappa_{t+1} = \kappa_t - \eta\,\mu_o(x, s, \kappa_t)$ with $\eta > 0$. Then

$$\frac{d\mu_o}{dt} = \frac{\partial\mu_o}{\partial\kappa}\cdot\frac{d\kappa_t}{dt} = \frac{\partial\mu_o}{\partial\kappa}\cdot\big(-\eta\,\mu_o\big) \;>\; 0,$$

so interracial marriage rises mechanically over time for fixed (x, s). Moreover, the marginal effect of x grows over time under mild conditions. Using (1), the cross-partial $\frac{\partial^2 \mathrm{Out}_W}{\partial x\,\partial\kappa} = \frac{(1-s)\,a_S\,a_I'(\kappa)\big[(1-q_{WB})a_S - q_{WB}a_I\big]}{D_W^3}$ is negative whenever $(1-q_{WB})a_S \geq q_{WB}a_I$, which holds for small x or when homophily/stigma are sizable so a_S/a_I is large. Historical non-Southern CZs satisfy exactly these conditions.

Increasing the Black population share increases white non-marriage: I can also consider how changes in various parameters might affect the hazard rate of marriage, and by extension, the fraction of the population that remains unmarried. Let $\lambda > 0$ be a Poisson meeting rate. Then, the marriage hazard for whites is

$$h_W(x, s, \kappa_t) = \lambda\,[\,q_{WW}(x, s)\,a_S + q_{WB}(x, s)\,a_I\,] = \lambda\big[\,a_S + (1-s)x\,(a_I - a_S)\big].$$

Then $\frac{\partial h_W}{\partial x} = \lambda(1-s)\big[a_I - a_S\big] < 0$, so the expected waiting time to marriage (and hence nonmarriage in a finite time period) rises with x for white individuals when $a_I < a_S$ (interracial marriage acceptance probabilities are lower than same-race acceptance probabilities).

3 Historical Context

3.1 Interracial Marriage

Interracial marriage has a long history in the Americas, with the earliest recorded unions coming from the Colonial Chesapeake (Hodes, 1999b). Throughout this timeframe, interracial couples have faced a variety of legal constraints, popular reactions, and opportunities for success. Historians and sociologists generally agree that interracial relationships[18] were more common during the Colonial Slavery period, when white indentured servants and Black enslaved people shared similar social class and thus had more opportunity to initiate relationships (Gullickson, 2006; Hodes, 1999b). Additionally, as rigid racial boundaries emerged in response to Reconstruction, the rate of interracial marriage reached a low in the first half of the 20th century, before rapidly increasing in the second half (Fryer, 2007; Gullickson, 2006).

Much of the historical scholarship on interracial marriage has focused on the legal history of anti-"miscegenation" laws, which prohibited interracial marriage (or sexual contact). These laws were often borne of fears of racial mixing and the threat that it could pose to rigid systems of racial hierarchy (Hodes, 1999a,b; Pascoe, 2009). Punishments varied greatly—in some states, interracial marriage was a felony, in others a misdemeanor (Browning, 1951). Additionally, some scholars have used the development and application of these laws to study how the legal system conceptualized and defined the boundaries of race in the 19th and 20th centuries (Pascoe, 2009; Berry, 1991; Johnston, 1970).

Additionally, both qualitative and quantitative evidence from specific geographic areas and time periods allows me to examine portraits of interracial couples in a certain context. Mills (1981), for example, focuses on identifying interracial couples in antebellum Alabama and finds many living in both urban and rural areas. He documents long lengths of interracial relationships and suggests that this points to an absence of community pressure against interracial marriage. Studies of newspapers and other media of the time suggest that there was widespread curiosity about interracial marriages, rather than pure antagonism (Sheffer, 2013; Lemire, 2002). While the legal environment indicates some level of hostility towards interracial couples, popular reactions appear to be more mixed.

As the data landscape for studying interracial couples has improved significantly in the past 20 years, several important quantitative patterns have emerged. First, the prevalence of interracial marriages has increased rapidly since the middle of the 20th century (Fryer, 2007). It appears that interracial relationships are descriptively more likely to divorce, but this pattern may be due to compositional differences (Zhang and Hook, 2009). Additionally, newly available data has enabled the study of historical characteristics of this population—in general, they fell in between Black and white same-race couples on most measures of

[18] These were typically not legally recognized marriages, but often more informal Gullickson (2006)

economic success and health (see (Deal, 2024) for more information). However, a question that remains is what historical forces drove the large changes in interracial marriage over this period.

3.2 Great Migration

The Great Migration was a mass migration of several million Black people from the Southern United States to the North (and West, to a lesser extent) throughout the middle of the 20th Century (Collins, 2021). This movement fundamentally changed the racial composition of the country, leading to large increases in Black urban populations in the North as well as corresponding decreases in the South, especially in rural areas. There is a long tradition of research in economics and history focused on studying the Great Migration (Scroggs, 1917), and topics of interest have ranged from the timing of the Great Migration (Collins, 1997) to the selection of and effects on migrants (Collins and Wanamaker, 2014; Eriksson, 2019; Black et al., 2015) to the effects of the migration on both sending and receiving communities (Boustan, 2010; Black et al., 2015).

One fundamental question relevant to this analysis is the composition of the migrants—who were they and why did they leave the South? With respect to the first question, the most recent evidence suggests that the migration was fairly broad-based, with some minor positive selection on education and human capital (Collins and Wanamaker, 2014). These migrants tended to be both families and individuals, often following the railroad routes that connected their origin community to the North. With respect to the second, explanations tend to emphasize a combination "push" and "pull" factors. Negative shocks to Southern agriculture, including the boll weevil, led to diminished economic opportunity for both Black and white Southerners, which was exacerbated by Jim Crow laws and widespread racial discrimination in the South (Collins, 2021; Lange et al., 2009). These conditions created an environment which many Black Southerners sought to flee. In addition, vibrant labor markets in Northern cities had relatively more economic opportunities, and the widespread distribution of the Northern Black press offered information and encouragement to Black Southerners about the North and their prospects for advancement (Collins, 2021). Finally, previous bouts of migration had created ties that stretched across the North and South, allowing many potential migrants to follow a family member or friend to Northern cities (Derenoncourt, 2022). Importantly, these historical factors will be leveraged to construct an instrument for Black in-migration in Northern cities.

There has been much prior work on the effects these migrants had on their Northern destination communities across fields (and comparatively little about the effects on sending communities). In economics, research has found that the Great Migration caused reductions in wages for some Black workers and upward mobility for Black children in Northern destination cities (Derenoncourt, 2022; Boustan, 2009). Those effects may have emerged in response to the reductions in government investment and increased support for

Civil Rights that accompanied Black in-migration (Calderon et al., 2022; Derenoncourt, 2022; Shi et al., 2022). Beyond this, a large interdisciplinary literature has recorded how coordinated responses to the Great Migration contributed to the rise of residential segregation, shaping the spatial population distribution of many American cities today. Cutler et al. (1999) and Massey and Denton (1993) both discuss how the use of restrictive residential covenants and other forms of racial discrimination emerged in response to the influx of Black people into Northern cities, which then precipitated the rise of residential segregation (Akbar et al., 2022). Additionally, there is evidence of a "white flight" response in which white residents of destination communities moved out of central cities into suburbs when large inflows of Black migrants occurred (Boustan, 2010). These residential and demographic realignments highlight how the Great Migration reshaped the structure of American cities. Beyond these effects, others have studied how the Great Migration shaped music, arts, and culture—many prominent Black artists, musicians, and leaders in the North were descendants of Southern migrants or migrants themselves, and the Great Migration is cited as a contributing factor to the Harlem Renaissance (Wilkerson, 2020).

Comparatively less attention has focused on the social responses to the Great Migration, in part because of a paucity of data on the social interactions of Black and white Americans throughout this time period. However, there is some evidence that the social divide between races was more porous in the North than the South (Grossman, 2011), indicating that migrants might see more social integration with their white neighbors. Importantly, Cook et al. (2023) find that Northern businesses increased their provision of nondiscriminatory services in response to Black in-migration from the Great Migration, suggesting that access to shared public accommodations increased (and potentially offering more opportunities for interracial couples to form). It was not easy by any means—Jackie Robinson's family, for example, moved to an all-white neighborhood in Pasadena, California and faced threats from their new community (Wilkerson, 2020). Despite this, he attended an integrated school and went on to break the color barrier in professional baseball (Wilkerson, 2020). Most related to my study, Fouka et al. (2022) find that the First Great Migration increased intermarriage between European immigrants (who had previously experienced marital segregation from the native population) and native-born US people, because newly arriving Black migrants were defined as a new "other" category, flattening the perceived distance between immigrants and natives. A question that remains to be explored is the Black-white social integration response to the Great Migration—I study this topic using interracial marriage as a proxy for social integration.

I focus on the Second Great Migration (1940-1970) for several reasons. First, the availability of a 5-year residence lookback question in the 1940 full-count Census enables the use of a complete county-level migration matrix to construct my shift-share instrument (Derenoncourt, 2022; Boustan, 2016). This methodological advantage makes it possible to construct a much stronger instrument for Black migration flows, as leveraging pre-1930 migration patterns would require linking across Censuses or coarser definitions of sending and receiving regions. Additionally, interracial marriage rates are relatively flat across this earlier period (see Figure A1a), suggesting that these earlier migration flows did not induce additional interracial marriages. In contrast, the end of the Second Great Migration coincides with large increases in the prevalence of interracial marriage.

3.3 Segregation

Residential segregation is the spatial separation of different groups and their living spaces, and in the US context primarily focuses on differences between Black and white Americans. Residential segregation has deep roots in the United States, and there are historical accounts of many spatial configurations—one, for example, involves white families living on larger, well-maintained streets while (predominantly) Black servants live in alleys and side streets that allow them to locate near their employers' houses (Massey and Denton, 1993). There are many quantitative measures that capture evenness, isolation, and other aspects of spatial population distribution, but the implication remains similar: segregated groups do not interact with each other often, and may face restricted opportunities due to their isolation (Massey and Denton, 1993). Logan and Parman (2017) use the order of enumeration in Census manuscripts to highlight how segregation was not unique to urban areas, and was already present in the late 19th and early 20th century US. Over time, it seems that residential segregation grew through the first two thirds of the 20th century, and peaked in the 1970s before declining since. The Great Migration, combined with a suite of collective actions that white families used to maintain rigid residential separation, contributed to this growth in the middle of the twentieth century (Cutler et al., 1999). Massey and Denton (1993) provide a vivid account of how restrictive residential covenants, reduced access to homeownership, and discrimination by real estate agents and boards allowed white Americans to maintain segregation despite large inflows of Black migrants.

Economists have studied how this historical pattern has contributed to contemporary racial disparities, identifying increases in racial economic inequality as well as decreases in black political efficacy and economic opportunity (Chyn et al., 2022; Ananat and Washington, 2009; Ananat, 2011). The effects on social integration have received less attention within economics, though Massey and Denton (1993) describe how extreme social isolation accompanied extreme spatial isolation, where many Black residents of highly segregated areas (what the authors term "hypersegregated") rarely ventured outside of their communities. One can imagine that rarely interacting with people of different races would greatly reduce the chances that one marries outside their race.

Residential segregation has often followed natural sites of demarcation, such as rivers and railroad tracks, which served as coordination devices for whites seeking to demarcate "Black" and "white" areas of geographic areas (Cutler et al., 1999; Ananat, 2011). Indeed, to quote Massey and Denton (1993), "The expansion of the ghetto generally followed the path of least resistance, slowing or stopping at natural boundaries such as rivers, railroad tracks, or major thoroughfares, and moving toward low status rather than high status areas." Especially as Northern cities went through rapid residential transitions in the latter half of the 20th century, many whites would flee neighborhoods once they reached a certain "tipping point" of Black population share, so the natural demarcation features (e.g., railroads and rivers) of a city allowed for easier coordination and control over the racial population distribution within a city. I will leverage this pattern to identify exogenous variation in residential segregation.

4 Data

My primary data source is the US Decennial Census (Ruggles et al., 2021, 2023). For 1870-1940, I use the full-count data, which allows me to capture a sizeable sample of interracial couples in those years. Additionally, I use the 1950 1%, 1960 5%, 1970 2%, 1980 5%, 1990 5%, and 2000 5% samples to complete a time series from 1870-2000. For my main empirical exercises, I will focus on the 1970-2000 data, though the full series is used for descriptives.[19] The most granular geographic unit with complete coverage identified in public samples in 1970-2000 is the county group (CG) for the 1970 and 1980 censuses and the public-use microdata area (PUMA) for the 1990 and 2000 censuses.[20] Because my independent variables of interest are measured at the commuting zone (CZ) level, I use a probabilistic mapping from these geographic units to CZs. Building on the crosswalk of Autor and Dorn (2013), which records the fraction of a County Group/PUMA that maps to a given CZ, I construct the fraction of a CZ's population that belongs to a certain CG/PUMA. Then, I take the CG/PUMA-level mean of my outcomes of interest (i.e., interracial marriage rates) and collapse them the CZ level, weighting using this fraction. Thus, I map from the Census sample to commuting zones probabilistically and retain full coverage. As long as these allocations are random with respect to my variable of interest, they will simply create measurement error in my outcome, biasing my results towards 0.[21]

I measure my primary variable of interest, the IMR, by linking the race of the spouse and estimating the fraction of Black and white marriages that are between a Black and white spouse. In this project, I focus on Black-white interracial couples, as the historical forces, legal environments, and popular reactions affecting this population were distinct from those of other interracial couples. In practice, I restrict to Black and

[19]While my main time period of interest is the aftermath of the Great Migration and the decline in residential segregation (e.g., the 1970s and 1980s), which coincide with rapid increases in the prevalence of interracial marriage, I provide results for 1990 and 2000 to trace out the time path of these effects.

[20]Counties are identified only above a certain population threshold.

[21]For the analyses using the railroad instrument in the Ananat (2011) sample, I use the metropolitan areas identified in IPUMS as the unit of analysis, requiring no crosswalking.

white respondents that are identified as the head of the household or the spouse of the household head and drop any respondents for whom the race of spouse is a race other than Black or white.[22] More formally, for a geographic unit g and Census year t, the observed IMR (μ^o_{gt}) is simply the fraction of marriages that are interracial:

$$\mu^o_{gt} = \frac{m_i}{m_i + m_w + m_b}$$

where m_i is the count of interracial marriages, m_w is the count of white marriages (both spouses are white), and m_b is the count of Black marriages (both spouses are Black), all within geographic unit g and Census year t.

While this primary outcome variable captures the IMR, it may obscure meaningful variation in social integration. Imagine two geographic units where the observed IMR is 5%, but 10% of the marriage market is Black and 90% is white in the first, while 50% of the marriage market is Black and 50% is white in the second. Evidently, the first is more integrated than the second, but the observed IMR would be the same. To overcome this problem, I create a measure of expected interracial marriages—the IMR expected under random assignment.[23] It takes the form of a weighted average of outmarriage rates for four race-gender groups—that is, the share of Black individuals who marry white spouses (and vice versa), conditioned on gender. For example, to compute the outmarriage rate under randomization for Black men, imagine that I randomly drew a partner among all women. Then, the probability that the marriage is interracial is simply the share of women who are white. I can compute these random outmarriage rate for each of the four race-gender groups, and then weight them by the share of the population in each race-gender cell.

I construct the IMR under random assignment at the level of geographic unit g at time t (μ^r_{gt}) by computing the following shares among married couples (presumably composing the marriage market) in a geographic unit g. $\text{share}^{r,j}_g$ the percent of the marriage market in geographic unit g that are individuals of race r (either Black or white) and gender j (either male or female). Once I have these quantities for each race-gender group, the IMR under random assignment of marriage for a certain group is the fraction of the opposite gender that has the opposite racial identity. For Black men, for example, the predicted IMR would be $\frac{\text{share}^{\text{white, women}}_g}{\text{share}^{\text{white, women}}_g + \text{share}^{\text{black, women}}_g}$. Thus, I calculate this predicted rate for each group and then construct a

[22]Because race is one of the required fields in the Census, and the race of the spouse is imputed using the spouse's location in the Census form, I do not have any missingness in these variables.

[23]Random assignment conditional on geography is the benchmark I chose to use to account for population composition, and this approach has been used in concurrent work (Goldman et al., 2023). However, it is not the only benchmark one could construct—in principle, the relevant marriage market for a given individual is not all married individuals in a geographic area, but there instead a group restricted by similarity in age, class, etc (or not constrained by geography). In principle, one could construct a random assignment conditional on these other characteristics, but it is not clear whether these would serve as better counterfactuals—after a certain point of matching, you would begin excluding levels of dissimilarity on characteristics that are observed in real married couples in the data. Thus, I use a minimal level of conditioning before random assignment, but acknowledge that this benchmark may be unrealistic given the large gaps between Black and white Americans on income, age, and other relevant characteristics for marital matching and the formation of relationships across geography.

weighted average for the geographic unit:

$$\mu^r_{gt} = \text{share}_g^{\text{white, men}} \frac{\text{share}_g^{\text{black, women}}}{\text{share}_g^{\text{black, women}} + \text{share}_g^{\text{white, women}}} + \text{share}_g^{\text{white, women}} \frac{\text{share}_g^{\text{black, men}}}{\text{share}_g^{\text{black, men}} + \text{share}_g^{\text{white, men}}}$$

$$+\text{share}_g^{\text{black, men}} \frac{\text{share}_g^{\text{white, women}}}{\text{share}_g^{\text{white, women}} + \text{share}_g^{\text{black, women}}} + \text{share}_g^{\text{black, women}} \frac{\text{share}_g^{\text{white, men}}}{\text{share}_g^{\text{white, men}} + \text{share}_g^{\text{black, men}}}$$

In the numerical example above, the first scenario would have $\mu^r_{gt} = 0.18$ and the latter would have $\mu^r_{gt} = 0.5$. Finally, I construct a measure of marital integration m_{gt} for a geographic unit g and time t that is the ratio of observed interracial marriage (μ^o_{gt}) and this random benchmark:

$$m_{gt} = \frac{\mu^o_{gt}}{\mu^r_{gt}}$$

This marital integration measure offers a useful supplement to the IMR and helps to 1) show that effects are not driven by population composition; 2) further explore effects when the treatment involves changes in population composition (as it does in the Great Migration analysis).

I plot a national time series of these three outcome measures in Figure A1. There are marked increases in both the raw interracial marriage rate and the marital integration rate (which adjusts for population composition) in the latter half of the 20th century. Over time, expected interracial marriage declines as the Black population share of the US falls (due in part to mass European emigration). Throughout my analysis, I estimate these quantities at different levels of geography g depending on the level of my independent variable of interest.

Additionally, while the interracial marriage rate offers an intuitive summary measure of the prevalence of Black and white interracial marriages, it is useful to decompose this measure to assess whether the effects I find are coming from increases in interracial marriages or changes on the extensive margin of marriage, for example. To do this, I code four mutually exclusive indicators by race. For Black respondents, these are the outmarriage rate (fraction married to a white spouse), fraction unmarried, fraction married to someone of the same-race (married to a Black spouse), and fraction married to someone of another race (for example, Asian). They are defined symmetrically for white respondents. These measures have the desirable property of summing to 1, so decreases in one must be offset by others, allowing me to study "substitution" effects of in-migration and segregation.

There are some limitations to this data, especially in the historical Censuses (which aren't the focus of my main results). In 1970 and later years, an individual's race was reported by someone in the household or group quarters. Before 1960, the race of respondents was assessed by the Census enumerator rather than self-reported by the respondent or the household head. Thus, these racial categorizations reflect the assessment of the Census enumerator and may differ from how respondents would characterize themselves (though that information still likely connotes something valuable about perception of racial discordance between spouses). Additionally, transcription errors (a mistaken "B" for "W" in the race field) could bias results by creating false interracial couples. While there's no reason to believe that these differences vary systematically across my treatment measures, Deal (2024) conducts several additional checks, including linking to race self-reports from Social Security Applications, and finds that the majority of interracial couples, even in historical Censuses, are not transcription errors.

Additionally, because I rely on place of residence rather than place of marriage for computing my expected measure of interracial marriage, there is likely some error in that the individual's relevant marriage market may be elsewhere if they moved after marriage. However, both historical and contemporary measures of internal migration emphasize that the majority of people settle relatively close to where they grow up (Molloy et al., 2011; Sprung-Keyser et al., 2022; Bernard, 2017). I also replicate the main analyses using only respondents between 16 and 35 to minimize this error and the results are substantively similar.

I use several other datasets to estimate the effects of segregation and the Great Migration. First, I use the historical (1826-1911) railroad placement shapefiles from Atack (2016) to construct novel measures of historical railroad density at the commuting zone level. I also create commuting-zone-level measures of residential dissimilarity to estimate heterogeneity in the effects of the Great Migration using tract-level racial composition data from the NHGIS for 1970-2000.[24] Finally, I use the sample of 130 non-southern commuting zones (CZs) used in Derenoncourt (2022) for which data on the urban Black population in 1940 and 1970 is collected from the census and from the County and City Data Book 1944–1977 (CCDB), which is used to construct the Great Migration shift-share instrument.[25]

5 Empirical Strategy

Internal migration and declining residential segregation may have affected the social integration of Black and white Americans. Thus, I use two approaches to assess the impact of plausibly exogenous variation in Black in-migration and residential segregation on interracial marriage outcomes. I use a shift-share approach with the Second Great Migration to explore the effects of increasing black population share and internal migration. Then, I use historical railroad placement to instrument for contemporary segregation and

[24]For the 1970 Census, I use county-subgroup level racial composition data because the tract data has low coverage.

[25]A list of these commuting zones is available in Appendix Table A24. They are broadly distributed across the non-Southern United States.

evaluate its relationship with interracial marriage.[26] The Great Migration design captures the total effect of increased Black population share, including any segregation responses, while the railroad design captures the partial effect of segregation on interracial marriage.

These exercises offer complementary paths to identifying the effects of historical trends on interracial marriage, as they rely on different sets of assumptions (and thus fail to identify a causal effect under different states of the world). The Great Migration shift-share design would fail if, for example, there were unobserved determinants of the pre-1940 distribution of Black Southern migrants in the North that were correlated with interracial marriage patterns in the latter half of the 20th century. One potential such determinant is racial attitudes—perhaps pre-1940 Black Southern migrants sought destinations that were more accepting of racial minorities, and those places were more conducive for interracial relationships and marriages 30+ years later.[27] In contrast, the segregation design relies on relevance, independence, and exclusion assumptions. While relevance is empirically testable, independence would fail if, for example, railroad density was caused by high local governance capacity and this quality was also correlated with interracial marriage outcomes. Additionally, exclusion would fail if railroad density affected interracial marriage through channels other than residential segregation, such as increased marriage market access. In sum, while both designs rely on inherently untestable assumptions, they are sufficiently differentiated that the failure of an assumption for one design does not annul the other.

5.1 Great Migration

The Second Great Migration, one of the largest instances of internal migration in US history, consisted of more than 4 million African Americans who moved North in search of opportunity outside of the heavily segregated Jim Crow South. It offers an opportunity to examine the effects of exogenous increases in Black population share on marriage markets and social integration in non-Southern areas. In this section I offer a brief overview of my identification strategy, which hews closely to Derenoncourt (2022). For more details, see Appendix C and the text and appendices of Derenoncourt (2022). Then, in Section 5.1.1, I apply recent methodological advances in the shift-share literature to this design to evaluate its validity.

I follow Derenoncourt (2022) by measuring this population change at the Commuting Zone (CZ) level and define the Great Migration Black population change as the 1940 to 1970 increase in urban Black population

[26] Appendix Figure A3 displays the reduced form relationships between these instruments and the interracial marriage rate in 1990.

[27] I provide evidence against these violations by testing the relationship between my instrument and pre-1940 IMRs. Effects are small and insignificant, providing evidence against differences in unobserved variables that impact IMRs.

as a share of initial 1940 urban population:

$$\Delta\text{Black pop}_{CZ}^{1940-1970} = \frac{b_{\text{urban},CZ}^{1970} - b_{\text{urban},CZ}^{1940}}{\text{pop}_{\text{urban},CZ}^{1940}}$$

where $b_{\text{urban},CZ}^{t}$ is the total Black population in all sample cities in commuting zone CZ in year t. As Derenoncourt (2022) outlines, this distribution is highly right-skewed, so I instead use the percentile function of the increase (GM_{CZ}) as the key independent variable for the effects of the Great Migration. Thus, the naive OLS equation takes the following form:

$$\text{OLS: } y_{CZ} = \alpha + \beta GM_{CZ} + X_{CZ}'\rho + \varepsilon_{CZ} \tag{6}$$

The coefficient β from equation 6 represents the OLS estimate of the effect of GM_{CZ}, the commuting zone level percentile of Black population increase 1940-1970, on interracial marriage rates, conditional on the controls outlined above. However, the relationship between this variable and interracial marriage cannot be interpreted as causal because many correlates of Black population change could drive this relationship (for example, the racial attitudes of receiving commuting zones). Thus, I instead use a shift-share approach, which has been used previously in the Great Migration context (Derenoncourt, 2022; Boustan, 2010). The intuition with a migration shift-share is that migration decisions are often due to a combination of "push" and "pull" factors of both origin and destination locations. Additionally, Black southern migrants tended to move where previous migrants from their communities had settled. Thus, when "pushes" from Southern counties cause outmigration, some component of the migration destination can be predicted with the pre-existing loctional distribution of Black Southern migrants. These shocks to "push" factors are plausibly exogenous with respect to shocks to "pull" factors. Interacting exogenous shifts in migration from origin locations with historical migration patterns in destination locations yields a potential instrument for Black population changes in the North.

Shift-share designs can be formulated as a set of assumptions about the exogeneity of the shifts, shares, or both to yield a parameter of interest (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2021; Adão et al., 2019). In my setting, I use shocks to Southern counties (push factors) as my "shifts" and the distribution of pre-1940 Southern migrants as my "shares." Because early Black Southern migrants were not choosing their destinations at random, these shares do not yield a path to identification (Derenoncourt, 2022). However, the shocks to Southern counties are plausibly exogenous to unobserved determinants of interracial marriage rates in Northern cities, and it is this design on which I rely for identification.

More formally, following Derenoncourt (2022) and Borusyak et al. (2021), I rely on two assumptions to identify the effects of the Second Great Migration on interracial marriage rates:

1. Conditional on baseline characteristics, the instrument for Black population increases[28] is orthogonal to omitted characteristics that are correlated with changes in interracial marriage after 1940:

$$\mathbb{E}[\widehat{GM}_{CZ} \times \tilde{\varepsilon}_{CZ} | X'_{CZ}] = 0$$

2. A shock-level law of large numbers applies—there are sufficient independent shocks, each with sufficiently small average exposure (weight in the shift-share design).

While Assumption (1) is inherently untestable, I provide corroborating evidence by testing whether the instrument is associated with pre-Great Migration interracial marriage rates (in the spirit of testing pretrends in a difference-in-differences design). Appendix Table A5 displays results from these regressions—in all cases the coefficients are insignificant and small in magnitude. Additionally, they can be seen as a figure in Appendix Figure A4. Thus, the Migration does not predict interracial marriage rates 1900-1940.[29] Assumption (2) is supported by using shocks to over 1200 origin counties rather than a state-level analysis, enabled by Derenoncourt (2022)'s use of complete-count Census data. Thus, I rely on these two assumptions to identify the effect of the Second Great Migration on interracial marriage outcomes in the North with a shift-share approach.

The estimating equations are as follows:

$$\text{First Stage: } GM_{CZ} = \gamma + \delta \widehat{GM}_{CZ} + X'_{CZ}\rho + \varepsilon_{CZ} \tag{7}$$

$$\text{Reduced Form: } y_{CZ} = \tilde{\alpha} + \tilde{\beta} \widehat{GM}_{CZ} + X'_{CZ}\tilde{\rho} + \tilde{\varepsilon}_{CZ} \tag{8}$$

The first stage equation 7 estimates the first stage relationship between the instrument, the percentile of predicted Black population change $\widehat{GM}_{CZ}$, and the percentile of actual Black population change GM_{CZ}. In equation 8, $\tilde{\beta}$ represents the reduced form impact of the Great Migration instrument on observed IMR, marital integration, and IMR expected under randomization. All specifications include the control vector X'_{CZ}, which consists of census region fixed effects and the share of the urban population made up of 1935–1940 Black southern migrants. For my main results, I report the reduced form ($\tilde{\beta}$) and 2SLS ($\tilde{\beta}/\delta$) coefficients outlined above, as well as OLS (β) coefficients from Equation 6. Additionally, I estimate heterogeneity in the 2SLS effects using commuting-zone-level measures of residential dissimilarity.

[28] For more information on the construction of $\widehat{GM}_{CZ}$, see Appendix C.

[29] Additionally, Derenoncourt (2022) highlights that a key component of this assumption is that shocks to the South are uncorrelated with shocks to the North. She provides evidence that identification with this instrument is not the result of correlated shocks to origin and destination locations by constructing alternative instruments and conducting an overidentification test.

5.1.1 Shift-Share Checks

While this shift-share design (and its variants) have been extensively used in previous literature to study the effects of the Great Migration (Derenoncourt, 2022; Boustan, 2009, 2010; Chyn et al., 2022; Fouka et al., 2022), many recent methodological advances allow further study of this design and its validity (Borusyak et al., 2021, 2024; Goldsmith-Pinkham et al., 2020). First, given that the path to identification runs through exogenous shocks in this context, I can consider the ideal shock-level experiment in this setting and then evaluate the variation relative to that benchmark. I described above several potential confounders that might bias the relationship between observed Black population change and interracial marriage outcomes. To account for these confounders, the ideal experiment would involve randomly assigning Black outmigration to Southern counties, which would flow to Northern cities based on previous migrant networks and then result in a predicted change in Black population share for each Northern receiving area by summing over all Southern counties. Because this outmigration would be randomly assigned, it would be unrelated to the interracial marriage outcomes of Northern receiving cities, even if the shares are endogenous. Thus I instrument for observed population increases with predicted population increases. Additionally, the large number of Southern counties (over 1,000 in my context) yields sufficiently small average exposure for each shock so the pooled results are not driven too much by any one outmigration event.

In reality, of course, outmigration is not randomly assigned. Instead, I use outmigration from each Southern county predicted purely based on Southern economic factors (share of workforce in manufacturing, WWII spending per capita, etc.). Because this exercise isolates variation in outmigration that is purely due to economic variables in the South, it should remain exogenous to Northern interracial marriage outcomes and confounders that might bias the results. Additionally, I include two unit-level controls to further isolate exogenous variation: census region fixed effects and the total share of the 1940 urban population made up of recent Black migrants from any southern county (Derenoncourt, 2022). These controls isolate the variation coming from Southern county shocks rather than Northern city characteristics by making the instrument only leverage residual variation conditional on region and pre-1940 migration patterns. Additionally, the instrument uses shares measured from 1935-1940 migrants to avoid concerns that the shares are shifting endogenously with the shocks (Borusyak et al., 2024). Because these shares were measured during a period of relatively low migration between the First and Second Great Migrations, the assumption that the shocks are not affecting the shares is plausible.[30]

[30] It is possible that this first stage is misspecified because individuals are making bilateral location choices and shocks to alternative locations are omitted (Borusyak et al., 2022). This would require shocks to potential destinations to be correlated to origin shocks, but if this correlation is positive, it could bias the first stage coefficient towards 0. Derenoncourt (2022) conducts an over-identification test that accounts for correlated shocks to southern states and northern destinations and finds similar evidence in her study of the effect of the Great Migration on upward mobility, so this concern appears somewhat minor in this setting.

Recent work has emphasized that additional checks may be necessary in scenarios where the unit-level weights do not sum to 1 (Borusyak et al., 2021, 2024). In these cases, the shift-share instrument is no longer the weighted *average* of the shocks, but instead the weighted *sum*. Thus, if certain units have systematically higher weights, then even with random shocks, they would get systematically higher levels of the instrument, introducing endogeneity. In my setting, this would take the form of network centrality—if certain Northern cities, say Detroit and Chicago, tended to receive large shares of previous migrants from many Southern counties, then even randomly assigned outmigration would systematically propagate to those places. Indeed, my weights ω_{jc} do not sum to 1, and instead can be any scalar ≥ 0 because their sum is over ~ 1200 Southern counties. To give an intuition for what this captures, say that every Southern county sent 10% of its migrants to Detroit. If there were exactly 1200 counties, then $S_c = \sum_{j=1}^{1200} \omega_{jc} = 1200(0.1) = 120$, capturing that Detroit is very exposed to Southern outmigration shocks through its migrant network. Borusyak et al. (2024) highlight the solution of controlling for the sum of incomplete shares, as this will force the comparison to only exploit residual variation from Southern shocks that is unrelated to the network centrality of the receiving area. While I already control for the total share of the 1940 urban population made up of recent Black migrants from any southern county following Derenoncourt (2022), which offers a coarser measure of this share-related average exposure, I also control for the sum of shares S_c as a robustness check. Appendix Table A18 displays these results and shows that the coefficients for the effect of Black population change on interracial marriage rates 1980-2000 are substantively unchanged and remain statistically significant when controlling for the sum of shares.

Additionally, Borusyak et al. (2024) highlight the value of describing variation in the shocks. I present the distribution of predicted Black Southern outmigration by county in Appendix Figure A7. The mean predicted Black outmigration per county is approximately 3,000 people, and the standard deviation is approximately 12,000. These predicted values (based on a model using only Southern economic factors) represent the shocks or shifts in my shift-share design. Additionally, I describe the distribution of importance weights (or mean average exposure for each Southern county shock—computed as $s_j = \frac{1}{N} \sum_{c=1}^{130} \omega_{jc}$) in Appendix Figure A8. These importance weights are all small, demonstrating that the design is not driven by a small number of highly weighted shocks. Additionally, the inverse Herfindahl index of these importance weights ($\frac{1}{\sum_j s_j^2}$) gives an "effective number of shocks"—in my setting this quantity is 60.39, offering reassurance that the findings are not driven by a few counties.

Finally, Adão et al. (2019); Borusyak et al. (2021) highlight that valid inference in shift-share settings requires adjusting for the "exposure" design, which must account for the fact that units with similar shares will mechanically have similar shocks and may have correlated error terms given the exposure to common shocks. This is especially difficult in my setting given that the independent variable (percentiles of Black in-migration) is a nonlinear transformation of the standard shift-share. Derenoncourt (2022) tackles this issue with the Great Migration design using a placebo permutation test and finds strong evidence that effects are unlikely to be driven by noise and remain statistically significant (Borusyak et al., 2021).

It is also worthwhile to assess what variation is composing my shift-share instrument and driving my results. Recent work by Goldsmith-Pinkham et al. (2020) formulates shift-share designs as a "pooled exposure design" in which the "shares measure differential exogenous exposure to shocks," which in my case are the shares of Black southern migrants living in a southern county in 1935 that report residing in a northern commuting zone in 1940. Following Goldsmith-Pinkham et al. (2020), I decompose my instrument into Rotemberg weights to assess what variation in the data is driving the estimates. Tables A19, A20 detail summary statistics about the Rotemberg weights. A majority of the weights (99%, Table A19) are positive and the weights are highly correlated with predicted migration flows from southern counties. The correlation between Rotemberg weights and predicted migration flows is 0.793 (Table A20) which means that the migration flows predicted by southern "push-factors" explain about 63% of the variation in the weights. Conversely, the weights are weakly correlated with variation in historical migration shares ($Var(z_k)$) with a low correlation coefficient of 0.158 (indicating they explain less than 3% of my variation) . This indicates that my identification is primarily driven by the shocks to Southern counties as opposed to historical migration shares, which is desirable because historical migration patterns to the North were much more likely to endogenous to potential interracial marriage outcomes in the North than economic shocks to Southern counties.

5.2 Segregation

Following prior literature, I measure segregation using the index of dissimilarity (Chyn et al., 2022; Ananat, 2011):

$$\text{Seg}_{CZ} = \frac{1}{2} \sum_{n \in CZ} \left| \frac{\text{Black}_n}{\text{Black}_{CZ}} - \frac{\text{White}_n}{\text{White}_{CZ}} \right|$$

where Black_n is the Black population in tract n, Black_{CZ} is the Black population in commuting zone CZ, and White_n and White_{CZ} are defined analogously for white population. This measure can be thought of as the share of the Black population that must relocate to achieve complete integration. It ranges between 0 and 1, indicating complete integration and segregation, respectively. To estimate the effects of segregation on interracial marriage, I could use a naive OLS regression of the following form:

$$\text{OLS: } y_{CZ} = \alpha + \beta \text{Seg}_{CZ} + X'_{CZ}\rho + \varepsilon_c \tag{9}$$

The coefficient β from equation 9 represents the OLS estimate of the effect of Seg_{CZ}, the commuting zone level dissimilarity index, on two interracial marriage outcomes, conditional on the controls outlined above. However, interpreting the direct effects of segregation on interracial marriage is difficult, as there are many factors that might simultaneously cause segregation and have effects on IMRs—for example, local government policies, labor market geography, or racial attitudes. To address this potential endogeneity, I build on prior work by Ananat (2011), which constructs an instrumental variable (IV) for contemporary segregation in Northern cities using the historical placement of railroads. The basic intuition is that when Black migrants arrived in an area, preexisting railroad networks facilitated the division of the area into predominantly single-race subareas through coordinated behaviors by white residents. To quote Massey and Denton (1993), "The expansion of the ghetto generally followed the path of least resistance, slowing or stopping at natural boundaries such as rivers, railroad tracks, or major thoroughfares, and moving toward low status rather than high status areas."[31]

To measure this activity, I use a railroad division index (RDI) at the commuting zone level:

$$\text{RDI}_{CZ} = 1 - \sum_{r \in CZ} \left(\frac{\text{area}_r}{\text{area}_{CZ}} \right)^2$$

where r indexes "railroad neighborhoods" (the regions constructed by the intersection of historical railroad lines), area_r is the area of land in railroad neighborhood r, and area_{CZ} is the total area of land in commuting zone CZ. This captures the amount of subdivision generated by railroad track placement, so commuting zones that had a greater number of delineated areas had more potential for segregation. This measure ranges from 0 (representing a commuting zone with 1 railroad neighborhood) to 1 (representing a commuting zone with a nearly infinite number of railroad neighborhoods).

I follow Ananat (2011) and Chyn et al. (2022) in using RDI_{CZ} as an instrument for residential segregation and estimate using two-stage least-squares (2SLS). Additionally, I test the robustness of my results to controlling for historical railroad track length, a measure that is correlated with RDI and could affect outcomes independently. Chyn et al. (2022) choose not to control for this, noting issues with the interpretation of linear IV estimates when controlling for covariates and an outlier in their data that causes substantial uncertainty across estimates. In general, my results are substantively unchanged whether controlling for this variable or not.

The main estimating equations take the following form:

[31] Methodological concerns about use of rivers and other topographic figures motivates the use of railroad tracks in place of those sources of variation (Rothstein, 2007).

$$\text{First Stage: } \text{Seg}_{CZ} = \gamma + \delta \text{RDI}_{CZ} + X'_{CZ}\rho + \varepsilon_{CZ} \tag{10}$$

$$\text{Reduced Form: } y_{CZ} = \tilde{\alpha} + \tilde{\beta}\text{RDI}_{CZ} + X'_{CZ}\tilde{\rho} + \tilde{\varepsilon}_{CZ} \tag{11}$$

The first stage equation 10 estimates the relationship between the instrument, the RDI index RDI_{CZ}, and the contemporary segregation of an area, measured with the dissimilarity index, Seg_{CZ}. In equation 11, $\tilde{\beta}$ represents the reduced form impact of the RDI instrument on the observed IMR and marital integration, respectively. Some specifications include the control vector X'_{CZ}, which consists of historical railroad track length. For my main results, I report the reduced form ($\tilde{\beta}$) and 2SLS ($\tilde{\beta}/\delta$) coefficients outlined above, as well as OLS (β) coefficients from Equation 9. See Appendix Appendix D for a more in-depth discussion of the railroad IV assumption.

I compute the railroad density index across entire commuting zones to harmonize the sample with the Great Migration analysis sample and to include non-urban areas rather than focusing on the small areas around historical downtowns (as in Ananat (2011)). I also replicate these results using the original sample of 121 non-Southern cities from Ananat (2011) in Appendix D. The results are similar, reinforcing the validity of the design.

6 Results

6.1 Great Migration

The Great Migration was a mass movement of millions of African Americans who fled restrictive racial hierarchies and a lack of economic opportunities in the South. When they migrated to their new communities, they provided an opportunity to see how marriage markets respond to an influx of racial minorities—does the IMR of these communities increase in response? Or do endogenous responses and segregation frictions reduce social integration?

Figure 2a displays the relationship between commuting-zone-level Black population change from 1940-1970 and 2000 IMRs. There is a linear and positive relationship between the two variables—it appears that higher Black population change is associated with higher rates of interracial marriage (Slope=0.15 (0.02)). However, this variation is not necessarily exogenous—there may be factors that would bias the relationship between Black population change and 2000 interracial marriage outcomes. As a result, I use the shift-share approach introduced in Derenoncourt (2022) to instrument for Black population change. Table 1 displays the first-stage relationship between the predicted Black population change and actual Black population change—there is a strong relationship between the two when controlling for region dummies and Black Southern Migration from 1935-1940 (F-statistic=43.53).[32]

Then, Table 2 displays the 2SLS relationship between instrumented Black population change and interracial marriage outcomes across four decades (1970-2000), allowing me to trace out the emergence and growth of these effects over time. Each column reports results for a different Census year. Panel A displays results for the observed IMRs. In 2000, a 10-percentile increase in Black population change caused an additional 1.32 interracial marriages per 1000 marriages. There is a positive and significant effect on IMRs across time. In line with model predictions, it appears that the effects on white outmarriage are large enough to increase interracial marriage in the aggregate. Indeed, this effect grows across decades, doubling or more every 10 years. One possibility is that the emergence of some interracial marriages may have spillover effects on future interracial marriages. For example, if younger people observe interracial couples, that could erode social norms against interracial marriage and make them more likely to consider a partner of a different race (Bursztyn et al., 2020). The stylized model formalizes this intuition, and my findings are consistent with eroding norms amplifying the effects of Black population change and effects growing over time.

In Figure 3, I plot the reduced form relationship between percentile of predicted Black migration and IMR over time. This figure, in the spirit of an event study, confirms that there are not differential trends in interracial marriage outcomes between places that received more or less Great Migration migrants prior to the Second Great Migration.[33] Additionally, it shows how the effect size grows over time after 1970.

These results confirm that the relationship in Figure 2a is indeed causal—cities that experienced higher Black population growth due to the Great Migration had higher IMRs as a result. In the aggregate, minority in-migration increased one salient measure of social integration. However, does this suggest the absence of endogenous responses or mediating frictions that would reduce social integration in response to minority in-migration? I explore this question further by 1) evaluating the relative magnitudes of observed and predicted increases; 2) examining heterogeneity in effects across the residential segregation of receiving CZs.

First, I examine marital integration, which accounts for changing population composition. My stylized model suggests that the effects of in-migration on this measure are ambiguous, as mechanical increases in the interracial marriage rate under randomization can overwhelm observed increases. Figure 4, Panel B displays the relationship between ventile of Black population change and marital integration—which deflates

[32] My first stage is stronger than that reported in Derenoncourt (2022) because I omit two controls that she includes in her specifications—education upward mobility in 1940 and the share of the labor force that is employed in manufacturing. In her setting, these controls are relevant for the primary outcome of upward mobility. In my setting, they are unnecessary for identifying effects on interracial marriage, and may instead serve as bad controls, if (for example) Black people who are more likely to interracially marry choose their destination due in part to its labor market or educational opportunities.

[33] One critique of this test might be that IMRs in the pre-periods is quite low, and thus the placebo analysis is limited. While the levels of IMR are relatively low in 1900-1940, I show in Appendix Table A6 that even in 1910 and 1920 there is a strong and significant gradient between IMR and other CZ-level characteristics (mixed-race/black population shares), highlighting that the CZ-level variation in IMR in this time period was meaningful.

the observed IMR by the IMR expected under randomization. Strikingly, it appears that this outcome is negatively associated with the ventile of Black population change. Panel C offers a potential explanation, showing that the IMR under randomization is increasing in ventile of Black population change—and that the slope of this relationship is much higher than that observed in Panel A. As a result, while the observed IMR increases in Black population change, the random rate increases at a greater magnitude, and thus the marital integration of those communities that had higher Black population change is lower.[34] Appendix Tables A12, A13, and A14 display the relationship between observed Black population increase and each of these outcomes, confirming the direction of the relationship established in Figure 4 across four decades. In 1990, a 10 percentile increase in observed black population increase is associated with approximately 0.7 additional interracial marriages per 1000 marriages. However, this increase is also associated with 12.1 additional expected interracial marriages per 1000 marriages. As a result, this level of Black population increases is associated with 0.04 points less marital integration.

When I examine these effects using the shift-share instrument, I find confirmation of the OLS relationships. Table 2, Panel C displays the effects of Black population change on the IMR expected under random assignment. Again, there is a significant positive effect across decades. The magnitude of the change in the prevalence of interracial marriage (Panel A) is much smaller than that expected under randomization to spouses (Panel C). For example, in 1990, a 1-percentile increase in Black population changes causes an increase in IMR expected under randomization that is 27 times larger than the increase in observed IMR. In terms of the standard deviations of both variables, this increase is still 70% larger. Panel B displays the effects on marital integration—there is a significant negative relationship between Great Migration induced Black population change and the marital integration of those communities, corresponding to the discrepancy in magnitudes above. In 1990, a 10-percentile increase in Black population change caused a 0.15 point decrease in marital integration. This discrepancy in magnitudes suggests there may have been mediating frictions or endogenous responses that limited the social integration response to the Great Migration. [35]

Next, I examine whether the effects of the Great Migration differed across cities with differing levels of pre-existing integration barriers, testing the model prediction that segregation frictions would dampen the effects of Black population change. Figure A6 displays the CZ-level distribution of the residential dissimilarity index in 1990. There is significant variation between 0.2 and 0.8, and I leverage this variation to examine heterogeneity in Great Migration effects across the median of this distribution.[36] Table 3 displays 2SLS

[34]The discontinuity in Panel (c) can be attributed to the right skew of the distribution of black population change in destination cities.

[35]Another possible interpretation is that because Black Southern migrants follow primarily Black migration networks (and my instrument relies on this fact for identification), they may have lower interracial marriage rates on arrival because they have more within-race ties in their destination community than unattached migrants.

[36]Appendix Table A23 reports means of demographic and economic characteristics for above/below median segregation commuting zones. There are no significant differences in white share, share married, age, or occupational income across these

coefficients for high segregation CZs in Panel A and low segregation CZs in Panel B, highlighting how in 1970-2000, the effects are driven by low-segregation destination CZs, with large and statistically significant coefficients in low-segregation cities and insignificant results among high-segregation cities.[37] This pattern suggests that some mediating frictions limited the social integration effects of the Great Migration.[38]

Additionally, I can further decompose these effects by coding outcome variables from four mutually exclusive indicators for each race—unmarried, same-race married, opposite race married, and other-race married. Table 4 displays IV coefficients on these outcomes. I find that higher rates of GM-induced Black population change caused increases in outmarriage rates for white individuals, in line with model predictions. However, when I divide the outmarriage rate by the rate expected under random assignment to spouses, these increases in outmarriage are dwarfed by those expected, resulting in a negative effect on the integration of these marriage markets. This mirrors the earlier pattern observed in the marital integration outcome. Interestingly, I also find increases in nonmarriage among the white population, suggesting that some white people are induced into remaining unmarried rather than marrying across race when the Black population in their commuting zone increases. This also accords with a model prediction that increasing the Black population share will decrease white marriage hazard. One explanation for this pattern might be that married white couples flee the region in response to black in-migration, in line with Boustan (2010). This response is possible, but given that the unit of analysis is a commuting zone, rather than central city, any moves from central city to suburb within the same commuting zone would not appear as exits, likely dampening this effect. These increases in outmarriage and nonmarriage are offset by decreases in same-race marriages.

For Black populations, I find increases in same-race marriage that are offset by decreases in outmarriage and marriage to other races (non white or Black). The decreases in outmarriage are consistent with the predictions of my stylized model. These effects are perhaps expected given the increase in Black population. However, in the aggregate, the prevalence of interracial marriage increases because the white population is much larger. Additionally, I find no effects on the extensive margin of marriage for Black individuals, suggesting the increase in marriage market thickness did not induce additional participation in the marriage market among Black indviduals. I additionally decompose these results further by gender in Table A17, and find very little evidence of heterogeneity by gender in the effects of the Great Migration. The integration effects are a bit more muted for white women, in accordance with the historical gendered opposition to interracial marriage on purity grounds.

groups. However, there is a significant difference in Black share.

[37] I allow CZ-level segregation to vary across year and in Table A21, I show that these measures are highly correlated across time.

[38] I also examine effects for each quartile of the CZ segregation distribution in Appendix Figure A5 and find consistent results—strong effects for CZs in the bottom quartiles of the segregation distribution and no effects for the top two quartiles.

Comparing the results from IV and OLS analyses, the magnitudes and direction of coefficients are broadly similar. In 1990, the instrumented effect of an increase in the Black population change is 11% smaller than the OLS relationship. Additionally, effects on marital integration are 7% smaller when instrumented. Finally, the effects on expected interracial marriage are 35% larger when instrumented. However, none of the point estimates are statistically distinct from each other. These differences might emerge from accounting for endogeneity that biases the OLS estimates, but the concordance of sign and (broadly) magnitude across OLS and IV approaches lends confidence that the true causal effect of black population accords with my results. I also estimate results among 16-35 year olds (presented in Appendix Table A15) to address potential issues with differences between place of enumeration and the relevant marriage market and proxy for the flow of new marriages rather than the stock of marriages. These results are qualitatively and statistically similar to the main results, allaying concerns about these issues.

Overall, these results reinforce that the observed patterns in Figure 4 correspond to significant causal effects on interracial marriage outcomes when instrumenting for Black population change. I find that Great Migration induced Black population change increased the interracial marriage rates in destination communities. Additionally, I find that these effects are substantive and significant from 1970-2000. It should be noted that it is not necessarily the migrants themselves who are marrying in the "extra" interracial marriages formed—indeed, due to the timing of the Great Migration and when my interracial marriage outcomes are measured, it seems unlikely that this is the case. Instead, the effect could be driven by the children of these new migrants or less measurable changes in racial attitudes and social norms that result from Black population increases. For example, Calderon et al. (2022) find that the Great Migration increased support for Civil Rights and the Democratic Party in destination cities—they also find improvements in racial attitudes among whites, which may have increased the probability of entering an interracial relationship.

Additionally, I decompose the effects of the Great Migration on interracial marriage by birth cohort, allowing me to consider dynamics across age and over time for the same cohorts. First, I define birth cohorts by decade of birth: pre-1920 (denoted 1910 cohort), 1920s, 1930s, 1940s, 1950s, 1960s, 1970s, and 1980s. I construct a measure of observed interracial marriage for each cohort-census year-CZ cell and estimate equation 8 for each cohort-census year cell. I plot these estimates by cohort in Figure 5. Several patterns are apparent. First, each successive birth cohort has generally larger interracial marriage effects, with the 1980 cohort seeing effects that are 30 times larger than those on the 1910 cohort in 2000 (though both are imprecisely estimated). Additionally, for earlier cohorts, there is a leveling off of effects, consistent with limited marriage market movement past age 40. Finally, for the middle cohorts (1930-1950), I see consistent growth across time within cohort, consistent with evolving social norms. These patterns again accord with the model outline of how evolving norms can amplify the effects of Black population increases over time.

One further question is the composition of these marginally induced interracial couples. Do they consist of Southern migrants? Are they younger? While I am unable to identify the "extra" interracial couples induced by the Great Migration, I take a step in this direction by regressing the predicted migration instrument on several demographic characteristics of interracial couples in receiving CZs. If, for example, the "extra" interracial couples tended to be younger, I might find a negative relationship between average age of interracial couples in CZ and the predicted migration to that CZ, as they would pull down the average. Appendix Table A22 displays the results from running this regression using indicators for being born in the South, having a birthplace that differs from reported state of residence, age, occupational income score, and number of children among the population of interracial couples in 1990. All of these regressions return insignificant and substantively small results, suggesting that the "extra" interracial couples did not differ demographically from other interracial couples on these dimensions. Alternatively, because the effect size on the prevalence of interracial marriage is not very large, it is possible that this method is not well-powered to detect changing composition of interracial couples.

6.2 Segregation

One factor that may impact the IMR is the opportunity for people of different races to interact. Residential segregation could restrict this opportunity, and has been decreasing over the past 50 years (Chyn et al., 2022). Beyond segregation in people's residences, residential segregation is also closely related with the geographic location (and segregation) of spaces like churches, schools, and social locations—where one might meet their partner. As US residential segregation has declined dramatically in the past 50 years—from a dissimilarity index of 0.73 in 1950 to 0.49 in 2000—it is possible that this shift may have affected interracial marriage rates (see Appendix Figure A2). I investigate whether the large increase in IMRs might be related to the decrease in residential segregation using plausibly exogenous cross-sectional variation. My stylized model predicts that segregation should reduce all measures of interracial marriage and social integration (holding population composition fixed).

The OLS relationship between residential segregation (measured by the dissimilarity index) and observed IMR—presented in Figure 2—suggests that a negative, though statistically insignificant, relationship exists in 1980 and 1990, and this estimate becomes much larger and more precise in 2000.[39] Moving from a completely unsegregated to a completely segregated city (or from 0 to 1 on the dissimilarity index) is associated with 21.9 less interracial marriages per 1000 marriages.

There are many factors that might bias the relationship between residential segregation and interracial marriage, so I use the strategy introduced in Ananat (2011) to isolate exogenous variation in residential

[39] See Appendix Tables A8, D10, and D11.

segregation using the placement of railroad tracks. Table 5 displays the first stage relationships between the RDI instrument and the dissimilarity index (a measure of residential segregation) from 1970-2000. This relationship is strongly positive and statistically significant, with similar magnitude across the four decades. The 1990 first-stage F-statistic is 27.1. Additionally, Appendix Table A1 reports these first stage results with an additional control for track length, and these results are very similar.

Table 6 displays the instrumented 2SLS relationship between the residential dissimilarity index and the observed IMR from 1970-2000. Each column reports results from a different regression equation. I find that residential segregation causes lower IMRs in 2000. In 2000, a standard deviation increase in the dissimilarity index caused 6.77 less interracial marriages per 1000 marriages. Panel B displays the effects of residential segregation on marital integration. I find that increased residential segregation decreases the marital integration of that community in 1980-2000, and that this effect grows over time. In 2000, a standard deviation increase in the dissimilarity index causes a 0.47 point decrease in marital integration. Appendix Table A2 displays these results when controlling for track length, and they are very similar. A back of the envelope calculation suggests that in a counterfactual perfectly integrated world, the black-white interracial marriage rate in the 2000 Census for my sample of 130 non-Southern commuting zones would be 43.5 per 1000 black and white marriages, as opposed to the 14.1 per 1000 that is observed. Comparing the effects from the OLS specifications to the IV estimates, I find that the IV estimates are generally larger, though not statistically distinct from the OLS estimates. This broad agreement aligns with the findings of Cutler et al. (1999), who find that OLS and IV estimates of the effects of segregation are very similar using a related (but distinct) topographical instrument. Any discrepancies may reflect that cities that were more segregated differed in unobservable ways that were conducive to interracial relationships, or that there is some measurement error in segregation that is accounted for with the instrument. I also estimate results among 16-35 year olds (presented in Appendix Table A3) to address potential issues with differences between place of enumeration and the relevant marriage market and proxy for the flow of new marriages rather than the stock of marriages. These results are qualitatively and quantitatively similar to the main results, allaying concerns about these issues.

Additionally, I can further decompose these effects by assessing four mutually exclusive outcome variables for each race—fraction unmarried, same-race married, opposite race married, and other-race married. Table 7 displays IV coefficients on these outcomes. I find that higher residential segregation had no effect on white outmarriage rates. However, when I divide the outmarriage rate by the rate expected under random assignment to spouses, I find a negative effect on the integration of these marriage markets. Interestingly, I also find decreases in nonmarriage among the white population and in marriage to other races, which are offset by increases in same-race marriage. These results suggest that in more highly segregated places, white individuals are much more likely to marry within race. For Black populations, I find large increases in same-race marriage that are offset by decreases in outmarriage and marriage to other races (non white or Black). Both white and Black individuals seem to respond to residential segregation by becoming more endogamous.

Additionally, I replicate these results for the sample of 121 non-Southern cities used in Ananat (2011), displayed in Appendix D. The qualitative pattern of results is the same. Using this sample, in 2000, a standard deviation increase in the dissimilarity index caused 3.47 less interracial marriages per 1000 marriages. Additionally, in 2000, a standard deviation increase in the dissimilarity index causes a 0.61 point decrease in marital integration. See Table D2 for the full set of main results.

Thus, I find that residential segregation does have a negative causal effect on interracial marriage and social integration using the placement of railroads as an instrument and cross-sectional variation, in accordance with the stylized model. The national decrease in residential segregation from 1970 onwards may play a role in the large contemporaneous increase in interracial marriage and marital integration. There are several potential mechanisms that might explain this effect. Perhaps the most immediate is that residential segregation determines the degree of interaction between people of different races—where one goes to church, walks their block, etc. Thus, more segregated cities may have lower interracial marriage rates given that Black and white residents do not interact as often. Indeed, Massey and Denton (1993) famously emphasize the extreme social isolation and lack of contact that occurred in cities with high segregation in the latter half of the 20th century (what they termed "hypersegregation"). However, there may also be less direct pathways from residential segregation to interracial marriage. Ananat (2011), for example, found that residential segregation increased racial economic inequality, which may have also decreased interracial marriage rates given patterns of assortative matching in the marriage market. Ananat and Washington (2009) found that segregation decreased Black political efficacy, which may also have had downstream impacts on interracial marriage and social integration.

These results suggest that one expected effect of the Great Migration would be a reduction of interracial marriage in receiving cities—as previous literature has recorded increases in residential segregation in response to the Great Migration, a second order effect of in-migration would be to reduce IMRs. The fact that the aggregate results for the Great Migration are positive suggest that exposure effects may have overwhelmed any negative effects on social integration stemming from residential segregation or other coordinated actions to reduce integration.

7 Conclusion

Between 1950 and 2000, the Black-white IMR in the United States increased from 1.5 per 1000 marriages to 26.5 per 1000 marriages. Several concurrent trends—the decrease in regional segregation spurred by the Great Migration and declining residential segregation—may have played a role in this increase. While previous studies have documented this increase or used structural models to study the determinants of interracial marriage, none have evaluated the historical forces impacting the prevalence of interracial marriage and social integration (Wong, 2003; Fryer, 2007). As a result, I evaluate the roles of migration and segregation using a stylized model of the marriage market and two historical quasi-experiments.

I find that minority in-migration increases interracial marriage, largely via increases in white outmarriage, while Black outmarriage falls in response to increased same-race contact. Additionally, residential segregation serves as both a direct friction to interracial marriage, decreasing rates when holding composition fixed, and a moderator, dampening the effects of minority in-migration. I find that the effects of Black population change on interracial marriage grow over time, consistent with eroding social norms that induce spillovers that amplify the effect of minority in-migration. Finally, I show that when population shifts dramatically, increases in observed interracial marriage are overwhelmed by increases in random interracial marriage rates, producing aggregate decreases in marital integration measures. These patterns largely accord with the model's predictions, suggesting that it captures the relevant forces in this setting.

Further work should seek to explore the mechanisms driving these effects across domains—the stylized model predicts that spheres with limited racial segregation (e.g., schools and workplaces) can serve as conduits for increasing social integration. One question that remains is: why did it take so long for this marriage market integration to occur? Early 20th century European immigrants offer a contrasting case—while they were differentiated from native-born Americans (sometimes racially) and often lived in ethnic enclaves that would similarly limit contact, their integration in the marriage market occurred far quicker and much earlier than Black individuals (Guterl, 2002; Hatton and Williamson, 1998; Fouka et al., 2022). Although a more restrictive legal environment is a potential contributing factor, Deal (2024) finds evidence against this. One possibility is that the uniquely severe and persistent residential segregation of African Americans in the 20th century served as a persistent barrier to interaction (Massey and Denton, 1993). Fouka et al. (2022) document a distinct contributing factor: Black migrants themselves served to help European immigrants assimilate and intermarry natives, by creating a new "other" category. Finally, the unique history of racism and economic suppression against African Americans is likely also responsible. While there has been significant progress on this measure of social integration in the last 50 years as internal migration and residential desegregation have occurred, my results also serve as a reminder that substantial racial sorting in marriage persists.

References

Adão, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010.

Ager, P., Boustan, L., and Eriksson, K. (2021). The intergenerational effects of a large wealth shock: White southerners after the civil war. *American Economic Review*, 111(11):3767–3794.

Akbar, P. A., Hickly, S. L., Shertzer, A., and Walsh, R. P. (2022). Racial segregation in housing markets and the erosion of black wealth. *Review of Economics and Statistics*, pages 1–45.

Allport, G. W. (1954). The nature of prejudice. *Reading/Addison-Wesley*.

Ananat, E. O. (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.

Ananat, E. O. and Washington, E. (2009). Segregation and black political efficacy. *Journal of Public Economics*, 93(5-6):807–822.

Atack, J. (2016). Historical geographic information systems (gis) database of u.s. railroads for 1826-1911. Revised October 2023.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American economic review*, 103(5):1553–1597.

Becker, G. S. (1973). A theory of marriage: Part i. *Journal of Political Economy*.

Bernard, A. (2017). Cohort measures of internal migration: Understanding long-term trends. *Demography*, 54(6):2201–2221.

Berry, M. F. (1991). Judging morality: Sexual behavior and legal consequences in the late nineteenth-century south. *The Journal of American History*, 78(3):835.

Black, D. A., Sanders, S. G., Taylor, E. J., and Taylor, L. J. (2015). The impact of the great migration on mortality of african americans: Evidence from the deep south. *American Economic Review*, 105(2):477–503.

Borusyak, K., Dix-Carneiro, R., and Kovak, B. (2022). Understanding migration responses to local shocks. *Available at SSRN 4086847*.

Borusyak, K., Hull, P., and Jaravel, X. (2021). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213.

Borusyak, K., Hull, P., and Jaravel, X. (2024). A practical guide to shift-share iv. Technical report.

Boustan, L. P. (2009). Competition in the promised land: Black migration and racial wage convergence in the north, 1940–1970. *The Journal of Economic History*, 69(3):755–782.

Boustan, L. P. (2010). Was postwar suburbanization "white flight"? evidence from the black migration. *Quarterly Journal of Economics*, 125(1):417–443.

Boustan, L. P. (2016). *Competition in the promised land: Black migrants in northern cities and labor markets*, volume 39. Princeton University Press.

Browning, J. R. (1951). Anti-miscegenation laws in the united states. *Duke Bar Journal*.

Bursztyn, L., Egorov, G., and Fiorin, S. (2020). From extreme to mainstream: The erosion of social norms. *American economic review*, 110(11):3522–3548.

Calderon, A., Fouka, V., and Tabellini, M. (2022). Racial diversity and racial policy preferences: The great migration and civil rights. *The Review of Economic Studies*, 90(1):165–200.

Carlana, M. and Tabellini, M. (2018). Happily ever after: Immigration, natives' marriage, and fertility. Technical report, Harvard Business School.

Chiappori, P.-A., Oreffice, S., and Quintana-Domeque, C. (2016). Black-white marital matching. *Journal of Demographic Economics*, 82(4):399–421.

Chiappori, P.-A., Salanié, B., and Weiss, Y. (2017). Partner choice, investment in children, and the marital college premium. *American Economic Review*, 107(8):2109–2167.

Choo, E. and Siow, A. (2006). Who marries whom and why. *Journal of political Economy*, 114(1):175–201.

Chyn, E., Haggag, K., and Stuart, B. (2022). The effects of racial segregation on intergenerational mobility: Evidence from historical railroad placement. Technical report, NBER.

Collins, W. J. (1997). When the tide turned: Immigration and the delay of the great black migration. *The Journal of Economic History*, 57(3):607–632.

Collins, W. J. (2021). The great migration of black americans from the US south: A guide and interpretation. *Explorations in Economic History*, 80:101382.

Collins, W. J. and Margo, R. A. (2000). Residential segregation and socioeconomic outcomes. *Economics Letters*, 69(2):239–243.

Collins, W. J. and Wanamaker, M. H. (2014). Selection and economic gains in the great migration of african americans: new evidence from linked census data. *American Economic Journal: Applied Economics*, 6(1):220–252.

Cook, L. D., Jones, M. E., Logan, T. D., and Rosé, D. (2023). The evolution of access to public accommodations in the united states. *The Quarterly Journal of Economics*, 138(1):37–102.

Cox, R., Cunningham, J., Ortega, A., and Whaley, K. (2022). Black lives: The high cost of segregation. Technical report, Center for Equitable Growth.

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (1999). The rise and decline of the american ghetto. *Journal of political economy*, 107(3):455–506.

Deal, C. (2024). Interracial couples in the united states 1850-2000: Evidence from census data. Technical report, Vanderbilt University.

Derenoncourt, E. (2022). Can you move to opportunity? evidence from the great migration. *American Economic Review*, 112(2):369–408.

Derenoncourt, E., Kim, C. H., Kuhn, M., and Schularick, M. (2023). Wealth of two nations: The u.s. racial wealth gap, 1860–2020. *The Quarterly Journal of Economics*.

Dillender, M. (2014). The death of marriage? the effects of new forms of legal recognition on marriage rates in the united states. *Demography*, 51(2):563–585.

Eriksson, K. (2019). Moving north and into jail? the great migration and black incarceration. *Journal of Economic Behavior & Organization*, 159:526–538.

Fouka, V., Mazumder, S., and Tabellini, M. (2022). From immigrants to americans: Race and assimilation during the great migration. *The Review of Economic Studies*, 89(2):811–842.

Fryer, R. 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.

Goldman, B., Porter, S., and Gracie, J. (2023). Who marries whom? the role of segregation by race and class. Technical report, Harvard University.

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gordon, N. and Reber, S. (2017). The effects of school desegregation on mixed-race births. *Journal of Population Economics*, 31(2):561–596.

Grossman, J. R. (2011). *Land of hope: Chicago, black southerners, and the great migration*. University of Chicago Press.

Gullickson, A. (2006). Black/white interracial marriage trends, 1850–2000. *Journal of Family History*, 31(3):289–312.

Guterl, M. P. (2002). *The color of race in America, 1900-1940*. Harvard University Press.

Hatton, T. J. and Williamson, J. G. (1998). *The age of mass migration: Causes and economic impact*. Oxford University Press, USA.

Hodes, M. (1999a). *Sex, Love, Race: Crossing Boundaries in North American History*. New York University Press.

Hodes, M. (1999b). *White Women, Black Men: Illicit Sex in the Nineteenth-Century South*. Yale University Press.

Johnston, J. H. (1970). *Race Relations in Virginia and Miscegenation in the South 1776-1860*. University of Massachusetts Press.

Kennedy, S. and Fitch, C. A. (2012). Measuring cohabitation and family structure in the united states: Assessing the impact of new data from the current population survey. *Demography*, 49(4):1479–1498.

Koh, Y. K. (2024). Racial sorting in the us marriage market: Evolution and welfare implications. Technical report, Columbia University.

Lange, F., Olmstead, A. L., and Rhode, P. W. (2009). The impact of the boll weevil, 1892–1932. *The Journal of Economic History*, 69(3):685–718.

Lemire, E. (2002). *"Miscegenation": Making Race in America*. University of Pennsylvania Press.

Logan, T. D. and Parman, J. M. (2017). The national rise in residential segregation. *The Journal of Economic History*, 77(1):127–170.

Manning, W. D. and Smock, P. J. (2005). Measuring and modeling cohabitation: New perspectives from qualitative data. *Journal of marriage and family*, 67(4):989–1002.

Margo, R. A. (2016). Obama, katrina, and the persistence of racial inequality. *The Journal of Economic History*, 76(2):301–341.

Massey, D. S. and Denton, N. A. (1993). *American Apartheid: segregation and the making of the underclass*. Democracy and urban landscapes. Harvard University Press, Cambridge, Mass.

Merlino, L. P., Steinhardt, M. F., and Wren-Lewis, L. (2019). More than just friends? school peers and adult interracial relationships. *Journal of Labor Economics*, 37(3):663–713.

Mills, G. B. (1981). Miscegenation and the free negro in antebellum "anglo" alabama: A reexamination of southern race relations. *The Journal of American History*, 68(1):16.

Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal migration in the united states. *Journal of Economic perspectives*, 25(3):173–196.

Moran, R. F. (2003). *Interracial Intimacy: The Regulation of Race and Romance*. The University of Chicago Press.

Newbeck, P. (2008). *Virginia Hasn't Always Been for Lovers: Interracial Marriage Bans and the Case of Richard and Mildred Loving*. Southern Illinois University Press.

Pascoe, P. (2009). *What Comes Naturally: Miscegenation Law and the Making of Race in America*. Oxford University Press.

Rothstein, J. (2007). Does competition among public schools benefit students and taxpayers? comment. *American Economic Review*, 97(5):2026–2037.

Ruggles, S., Fitch, C. A., Goeken, R., Hacker, J. D., Nelson, M. A., Roberts, E., Schouwiler, M., and Sobek, M. (2021). Ipums ancestry full count data: Version 3.0.

Ruggles, S., Flood, S., Sobek, M., Brockman, D., Cooper, G., Richards, S., and Schouwiler, M. (2023). Ipums usa: Version 13.0.

Schwandt, H., Currie, J., Bär, M., Banks, J., Bertoli, P., Bütikofer, A., Cattan, S., Chao, B. Z.-Y., Costa, C., González, L., Grembi, V., Huttunen, K., Karadakic, R., Kraftman, L., Krutikova, S., Lombardi, S., Redler, P., Riumallo-Herl, C., Rodríguez-González, A., Salvanes, K. G., Santana, P., Thuilliez, J., van Doorslaer, E., Ourti, T. V., Winter, J. K., Wouterse, B., and Wuppermann, A. (2021). Inequality in mortality between black and white americans by age, place, and cause and in comparison to europe, 1990 to 2018. *Proceedings of the National Academy of Sciences*, 118(40).

Scroggs, W. O. (1917). Interstate migration of negro population. *Journal of Political Economy*, 25(10):1034–1043.

Sheffer, J. A. (2013). *The Romance of Race: Incest, Miscegenation, and Multiculturalism in the United States, 1880-1930*. Rutgers University Press.

Shen, M. (2018). How i met your mother: The effect of school desegregation on birth outcomes. *Economics of Education Review*, 63:31–50.

Shi, Y., Hartley, D., Mazumder, B., and Rajan, A. (2022). The effects of the great migration on urban renewal. *Journal of Public Economics*, 209:104647.

Sprung-Keyser, B., Hendren, N., Porter, S., et al. (2022). *The radius of economic opportunity: Evidence from migration and local labor markets*. US Census Bureau, Center for Economic Studies.

Wilkerson, I. (2020). *The warmth of other suns: The epic story of America's great migration*. Penguin UK.

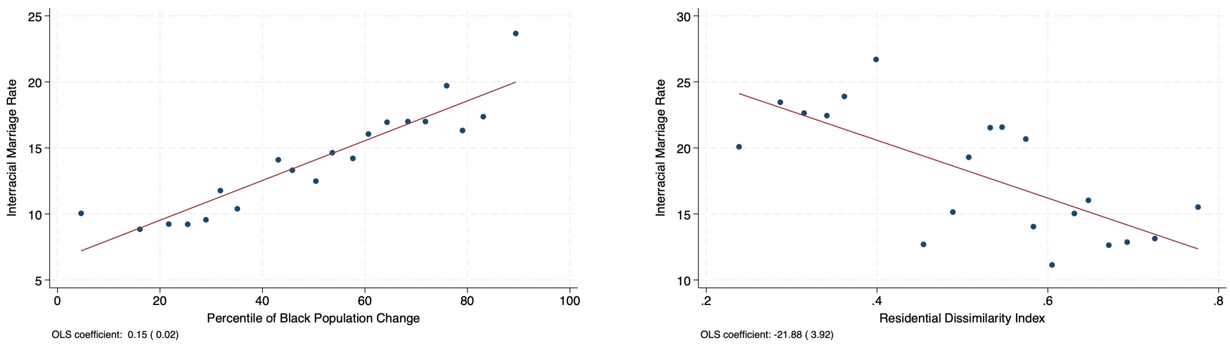
Wong, L. Y. (2003). Why so only 5.5% of black men marry white women? *International Economic Review*, 44(3):803–826.

Zhang, Y. and Hook, J. V. (2009). Marital dissolution among interracial couples. *Journal of Marriage and Family*, 71(1):95–107.

Tables and Figures

Figure 1: Time Series of Interracial Marriage, Residential Segregation, and Regional Segregation

Source: IPUMS-USA Full Counts and Samples, 1870-2000; NHGIS state-race population counts; Cutler et al. (1999); Author's calculations. Interracial marriages as a fraction of all Black and white marriages. Residential dissimilarity index measures the percentage of a group's population that would have to change residence for each neighborhood to have the same percentage of that group as the metropolitan area overall. Regional dissimilarity is defined analogously for the four Census regions.



(a) IMR vs. Black Population Change

(b) IMR vs. Residential Segregation

Figure 2: OLS Relationship between IMR and Independent Variables

Source: IPUMS-USA 2000 5% Sample; Author's calculations. Panel (a) is a binned scatter plot depicting the relationship between interracial marriage rates and the percentile of actual Black population increase during the Great Migration (1940-1970) for northern CZs. The unit of observation is a CZ. The black population change variable is grouped into 20 bins (5 percentiles each). Interracial marriages as a fraction of all Black and white marriages, dependent variable is IMR per 1000 marriages. Panel (b) is a binned scatter plot depicting the relationship between interracial marriage rates and the residential dissimilarity index for non-Southern metro areas. The unit of observation is an MSA.

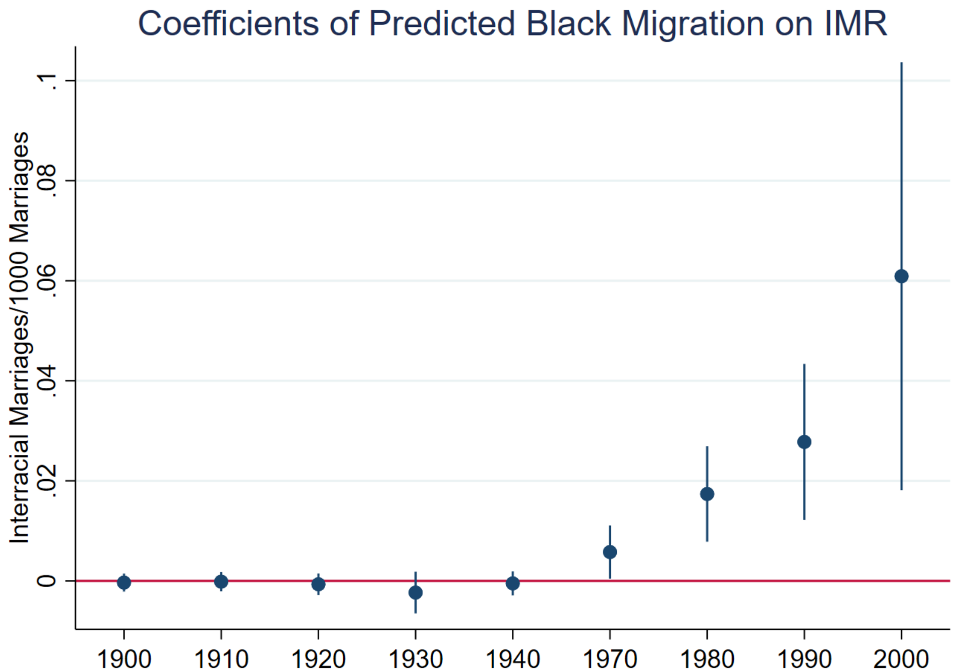
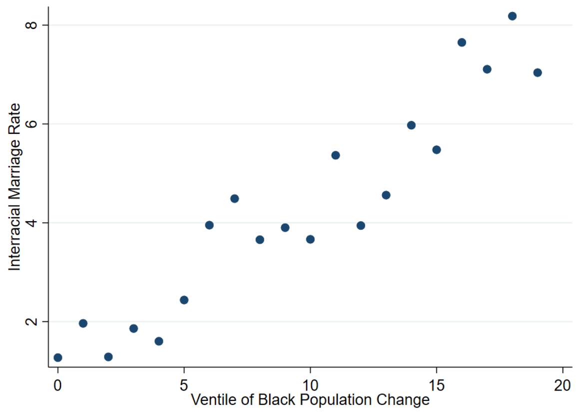
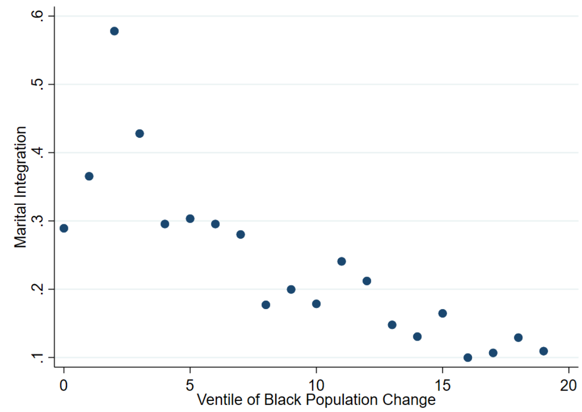


Figure 3: Regression of Great Migration Instrument on 1900-2000 IMR

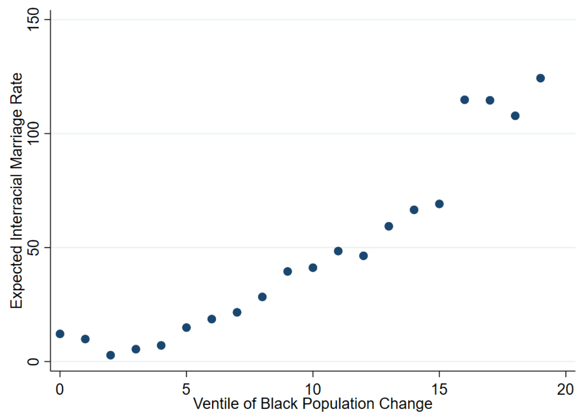
Source: Data from Derenoncourt (2022), 1900-1940 Full-Count Censuses, and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. This figure reports the estimated impact of the Great Migration on observed interracial marriage rates in 1900-1940 (a placebo outcome) and then 1970-2000 after the Great Migration has occurred. The unit of observation is a CZ. The dependent variable is the observed rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. The independent variable is the instrument for Black population increase (percentile of predicted Black population increase), defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects.



(a) Observed Interracial Marriage



(b) Marital Integration



(c) Expected Interracial Marriage

Figure 4: Relationship between Black Population Change and Outcomes

Source: IPUMS-USA 1990 5% Sample; Author's calculations. These binned scatter plots depict the relationship between interracial marriage outcomes and the percentile of actual Black population increase during the Great Migration (1940-1970) for northern CZs. The unit of observation is a CZ. The right-hand-side variable is grouped into 20 bins (5 percentiles each). Interracial marriages as a fraction of all Black and white marriages, dependent variable is IMR per 1000 marriages. Marital integration is the observed interracial marriage rate scaled by the expected rate of interracial marriage.

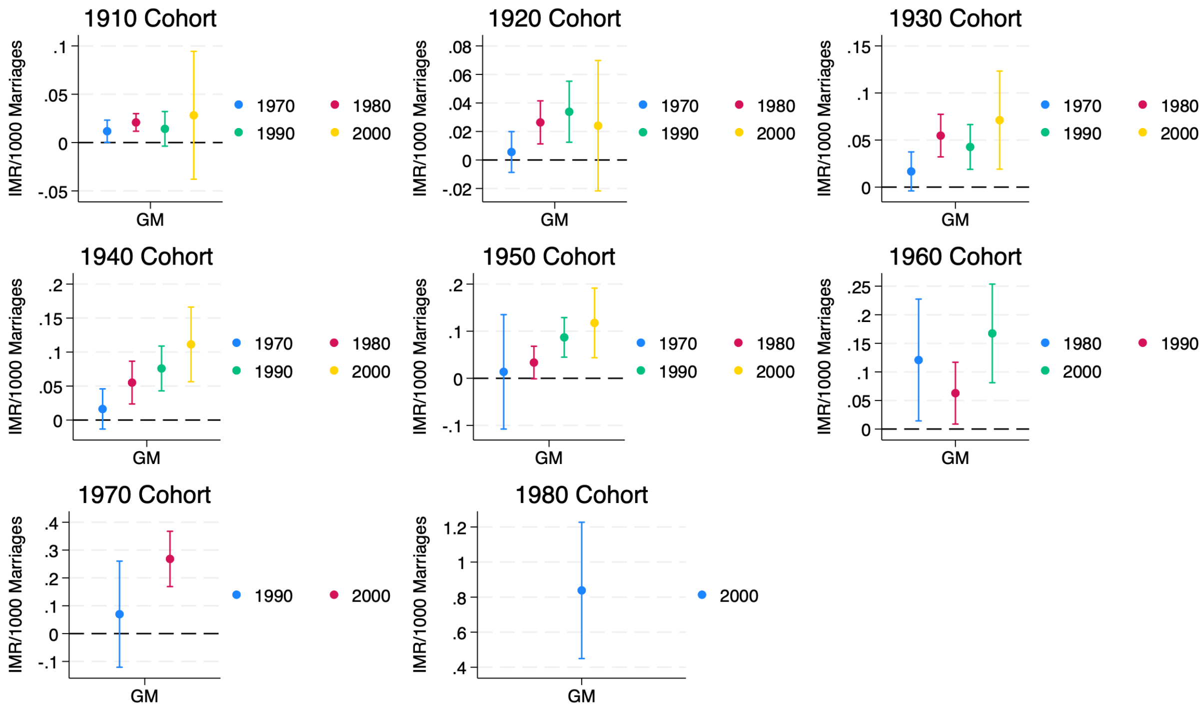


Figure 5: Great Migration Effects by Cohort and Year

Source: Data from Derenoncourt (2022) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. Each figure reports the estimated impact of the Great Migration on observed interracial marriage rates in 1970-2000 for a given birth cohort. 1910 cohort refers to any individuals born before 1920, while the rest of the cohorts are defined as individuals born in the 10 years following the outlined year (for example, the 1920 cohort is any individuals born in 1920-1929). The unit of observation is a CZ. The dependent variable is the observed rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. The independent variable is percentile of Black population increase, instrumented by the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects.

Table 1: First Stage on Black Population Change

	Percentile of Black Population Change
$\hat{GM}$	0.461***
	(0.0699)
F	43.53

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Derenoncourt (2022), Author's calculations. This table reports the first stage relationship (coefficients and heteroskedasticity-robust standard errors) between the percentile of predicted Black population change and the actual Black population change 1940-1970, conditional on 1935-1940 Black Southern migration and region indicators. Unit of observation is a commuting zone ($N = 130$).

Table 2: Effects of Great Migration on Interracial Marriage Outcomes

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
GM	0.0125**	0.0377***	0.0602***	0.132***
	(0.00614)	(0.00852)	(0.0138)	(0.0301)
ymean	1.26	2.43	4.28	14.1
N	130	130	130	130
Panel B: Marital Integration				
GM	-0.00287***	-0.00339***	-0.00439***	-0.0155***
	(0.00103)	(0.000927)	(0.00109)	(0.00375)
ymean	.0778	.109	.183	.667
N	128	129	130	130
Panel C: Expected IMR Under Randomization				
GM	1.880***	2.018***	1.633***	1.654***
	(0.248)	(0.234)	(0.176)	(0.173)
ymean	51.7	52.7	47.9	50.1
N	130	130	130	130

Source: Data from Derenoncourt (2022) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. This table reports the estimated impact of the Great Migration on observed interracial marriage rates, expected interracial marriage rates, and marital integration. The unit of observation is a CZ. The dependent variable in Panel A is of the observed rate of interracial marriage per 1000 marriages. The dependent variable in Panel B is marital integration. The dependent variable in Panel C is the expected rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. Marital integration is the observed interracial marriage rate scaled by the expected rate of interracial marriage. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table 3: Heterogeneity of Great Migration Effects by Segregation

	1970	1980	1990	2000
Panel A: High Segregation Cities				
GM	0.00892	0.0241	0.0319	0.0405
	(0.00605)	(0.0164)	(0.0217)	(0.0497)
Observations	65	65	65	65
Panel B: Low Segregation Cities				
GM	0.0178	0.0477**	0.0903***	0.515***
	(0.0114)	(0.0195)	(0.0261)	(0.199)
Observations	65	65	65	65

Source: Data from Derenoncourt (2022) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Additionally, IPUMS NHGIS Extracts from 1970-2000 to calculate commuting-zone-level dissimilarity indices. Author's calculations. This table reports the estimated impact of the Great Migration on observed interracial marriage rates by above/below median segregation cities. The unit of observation is a CZ. The dependent variable in Table 3 is of the observed rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table 4: Decomposing Great Migration Marriage Effects

	Outmarriage	Integration	Unmarried	Same Race	Other Race
Panel A: White Marriage Outcomes					
GM	0.00335***	-0.116***	0.0615***	-0.0687***	0.00391
	(0.000643)	(0.0266)	(0.0165)	(0.0180)	(0.00444)
ymean	.182	7.24	39.7	59	1.11
N	130	130	130	130	130
Panel B: Black Marriage Outcomes					
GM	-0.0869***	-0.0782***	-0.0349	0.138***	-0.0166*
	(0.0296)	(0.0299)	(0.0348)	(0.0330)	(0.00914)
ymean	9.46	9.76	64.6	24.1	1.84
N	130	130	130	130	130

Source: Data from Derenoncourt (2022) and the 2000 5% state IPUMS-USA sample. Author's calculations. This table reports the estimated impact of the Great Migration on five marriage outcomes for each race. The unit of observation is a CZ. Panel A reports marriage outcomes for white respondents, while Panel B reports marriage outcomes for Black respondents. The dependent variable in Column 1 is the number of respondents who are married to the opposite race (Black spouse for white respondents and vice versa) per 100 respondents. The dependent variable in Column 2 is the outmarriage rate (displayed in Column 1) divided by the outmarriage rate under random assignment to spouses. The dependent variable in Column 3 is the number of respondents per 100 respondents who are unmarried. The dependent variable in Column 4 is the number of respondents per 100 respondents who are married to someone of the same race. The dependent variable in Column 5 is the number of respondents per 100 respondents who are married to someone whose race is neither white nor Black. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table 5: First Stage on Residential Segregation

	1970	1980	1990	2000
RDI Instrument	0.296***	0.261***	0.375***	0.384***
	(0.0550)	(0.0804)	(0.0721)	(0.0779)
Observations	130	130	130	130
F	29.00	10.55	27.10	24.30

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Atack (2016) and IPUMS NHGIS Extracts from 1970-2000, Author's calculations. This table reports the first stage relationship (coefficients and heteroskedasticity-robust standard errors) between the railroad density instrument and the dissimilarity index segregation measure by decade. Column 1 reports the results for 1970, 2 reports 1980, 3 reports 1990, and Column 4 reports 2000. The unit of observation is a CZ.

Table 6: Effects of Residential Segregation on Interracial Marriage Outcomes

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
Dissimilarity Index	-1.021	-9.227	-9.798	-53.70***
	(1.710)	(7.005)	(6.422)	(19.06)
Outcome Mean	1.26	2.43	4.28	14.1
Observations	130	130	130	130
Panel B: Marital Integration				
Dissimilarity Index	-0.0955	-0.969***	-0.782***	-3.692***
	(0.181)	(0.366)	(0.270)	(1.038)
Outcome Mean	.0778	.109	.183	.667
Observations	128	129	130	130

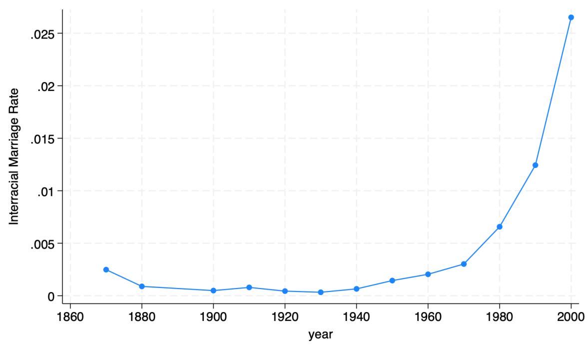
Source: Data from Atack (2016), IPUMS NHGIS Extracts from 1970-2000, and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from 2SLS models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. In Panel A the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Panel B the dependent variable is the marital integration in that Census year. Sample contains 130 non-Southern CZs.

Table 7: Decomposing Segregation Effects

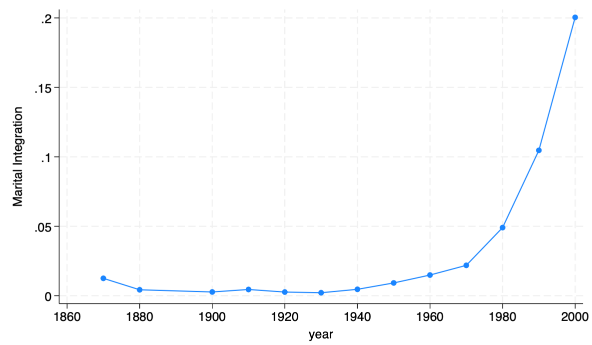
	Outmarriage	Integration	Unmarried	Same Race	Other Race
Panel A: White Marriage Outcomes					
Dissimilarity Index	-0.333	-32.11***	-16.95***	25.21***	-7.932***
	(0.261)	(9.130)	(5.985)	(7.815)	(2.159)
ymean	.182	7.24	39.7	59	1.11
N	130	130	130	130	130
Panel B: Black Marriage Outcomes					
Dissimilarity Index	-26.50***	-26.55***	17.28	19.60*	-10.38***
	(9.158)	(9.053)	(10.91)	(11.49)	(2.998)
ymean	9.46	9.76	64.6	24.1	1.84
N	130	130	130	130	130

Source: Data from Atack (2016), IPUMS NHGIS Extracts from 1970-2000, and the 2000 5% state IPUMS-USA sample. Author's calculations. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from 2SLS models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. Panel A reports marriage outcomes for white respondents, while Panel B reports marriage outcomes for Black respondents. The dependent variable in Column 1 is the number of respondents who are married to the opposite race (Black spouse for white respondents and vice versa) per 100 respondents. The dependent variable in Column 2 is the outmarriage rate (displayed in Column 1) divided by the outmarriage rate under random assignment to spouses. The dependent variable in Column 3 is the number of respondents per 100 respondents who are unmarried. The dependent variable in Column 4 is the number of respondents per 100 respondents who are married to someone of the same race. The dependent variable in Column 5 is the number of respondents per 100 respondents who are married to someone whose race is neither white nor Black. Sample contains 130 non-Southern CZs.

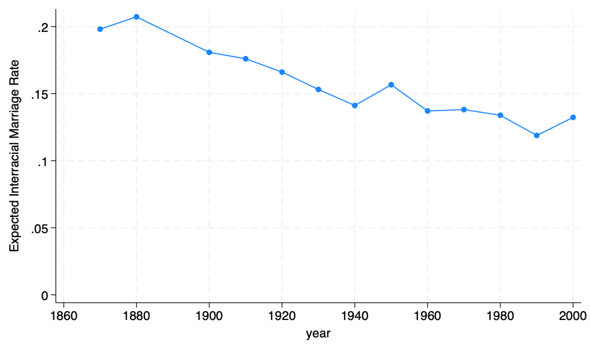
Appendix A: Supplementary Figures and Tables



(a) Observed Interracial Marriage



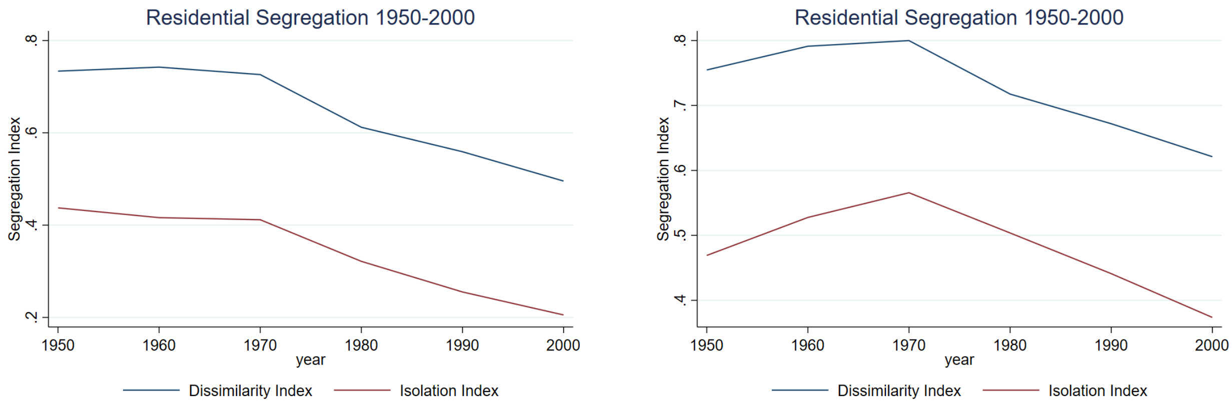
(b) Marital Integration



(c) Expected Interracial Marriage

Figure A1: Time Series of Interracial Marriage, Marital Integration, Expected Interracial Marriages

Source: IPUMS-USA Full Counts and Samples, 1870-2000; Author's calculations. Interracial marriages as a fraction of all Black and white marriages. Marital integration is the observed interracial marriage rate scaled by the expected rate of interracial marriage.

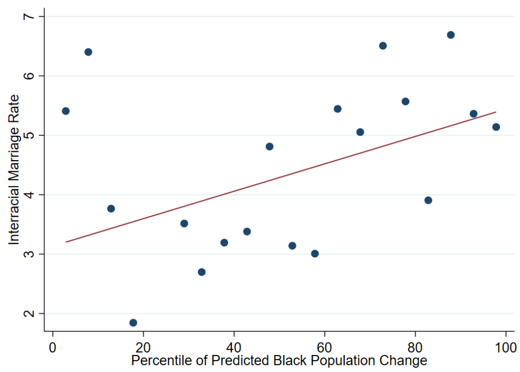


(a) Full Sample

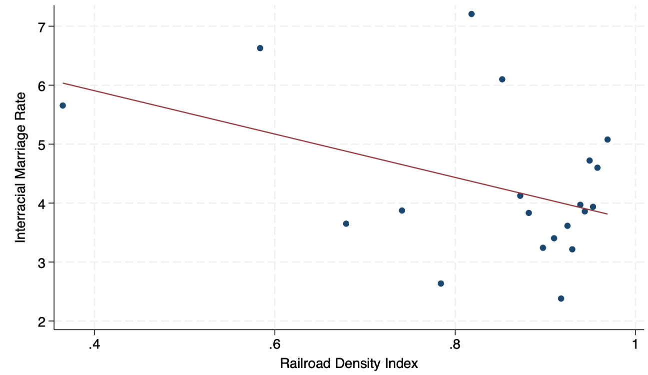
(b) Consistent Panel 1950-2000

Figure A2: Time Series of Residential Segregation, 1950-2000

Source: Cutler et al. (1999); Author's calculations. Dissimilarity index measures the percentage of a group's population that would have to change residence for each neighborhood to have the same percentage of that group as the metropolitan area overall. Isolation index measures minority-weighted average of the minority proportion in each area.



(a) Predicted Black Population Increase



(b) Railroad Density Index

Figure A3: Reduced Form Relationship between Instruments and IMR

Source: IPUMS-USA 1990 5% Sample; Author's calculations. Panel A is a binned scatter plot that displays the relationship between IMR and the percentile of predicted Black population increase during the Great Migration (1940-1970) for northern CZs. Panel B is a a binned scatter plot that displays the relationship between IMR and the Railroad Density Index (RDI). Interracial marriages as a fraction of all Black and white marriages, dependent variable is IMR per 1000 marriages.

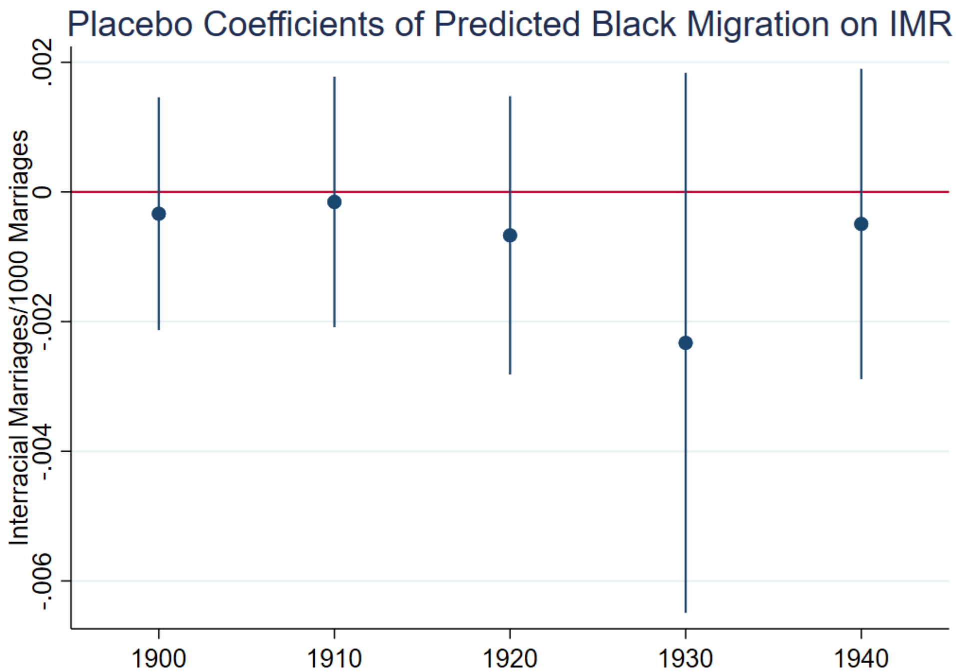


Figure A4: Placebo Test of Great Migration Instrument on 1900-1940 IMR

Source: Data from Derenoncourt (2022) and 1900-1940 Full-Count Censuses. Author's calculations. This figure reports the estimated impact of the Great Migration on observed interracial marriage rates in 1900-1940 (a placebo outcome). The unit of observation is a CZ. The dependent variable is the observed rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects.

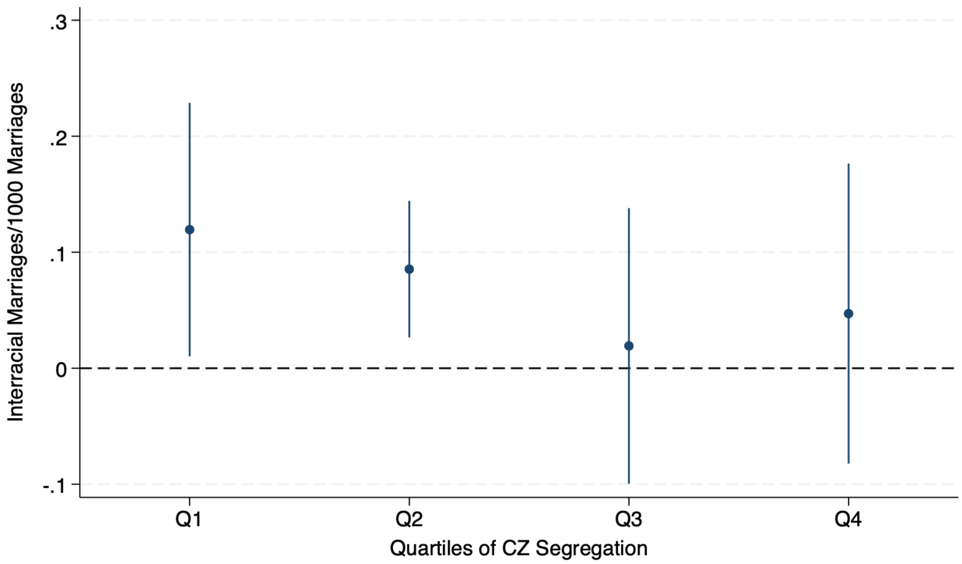


Figure A5: Heterogeneity of Great Migration Effects by Segregation Quartile

Source: Data from Derenoncourt (2022) and the 1990 5% state IPUMS-USA sample. Additionally, IPUMS NHGIS Extracts from 1990 to calculate commuting-zone-level dissimilarity indices. Author's calculations. This figure reports the estimated impact of the Great Migration on observed interracial marriage rates by CZ segregation quartiles in 1990. The unit of observation is a CZ. Q1 denotes the lowest segregation cities, while Q4 is the highest. The dependent variable is of the observed rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

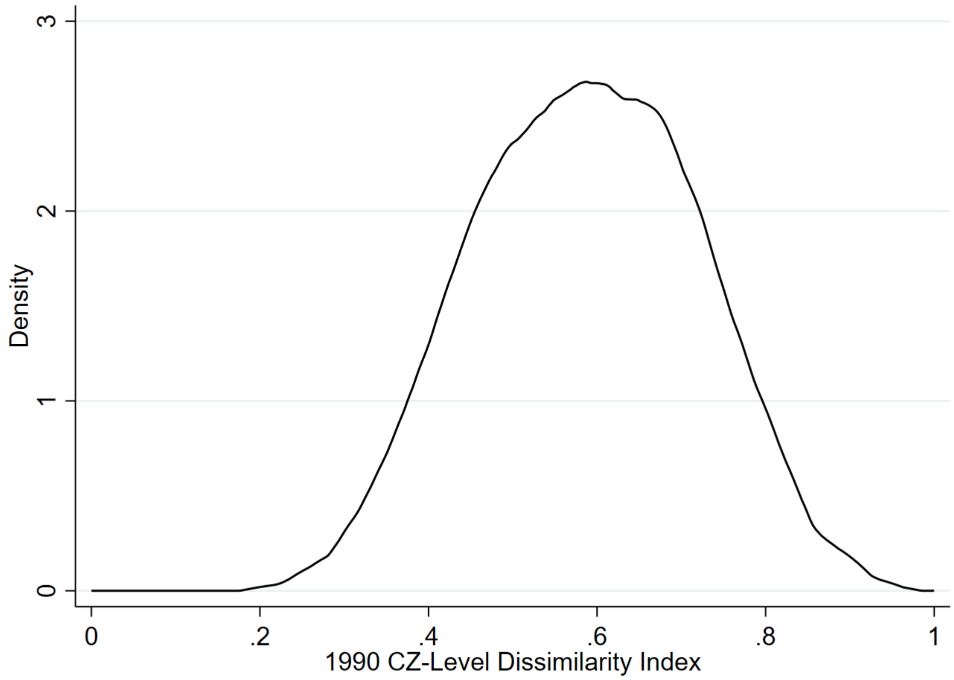


Figure A6: Distribution of Commuting Zone Residential Segregation

Source: IPUMS NHGIS Extracts from 1990. Author's calculations. This figure reports a kernel density plot of the black-white residential dissimilarity index. The unit of observation is a CZ. Dissimilarity index calculated over Census tracts. Kernel is Epanechnikov, and bandwidth set to 0.05.

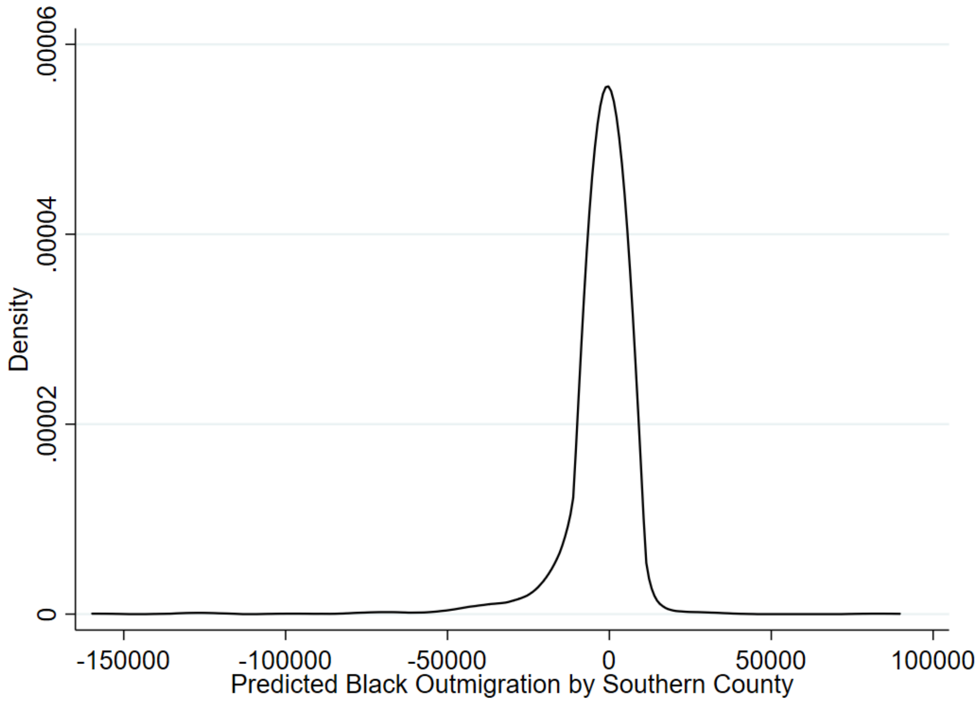


Figure A7: Distribution of Predicted Southern Outmigration (Shocks)

Source: Data from Derenoncourt (2022) Author's calculations. This figure reports a kernel density plot of the predicted county-level Black outmigration 1940-1970. The unit of observation is a Southern county. Kernel is Epanechnikov, and bandwidth set to 10,000.

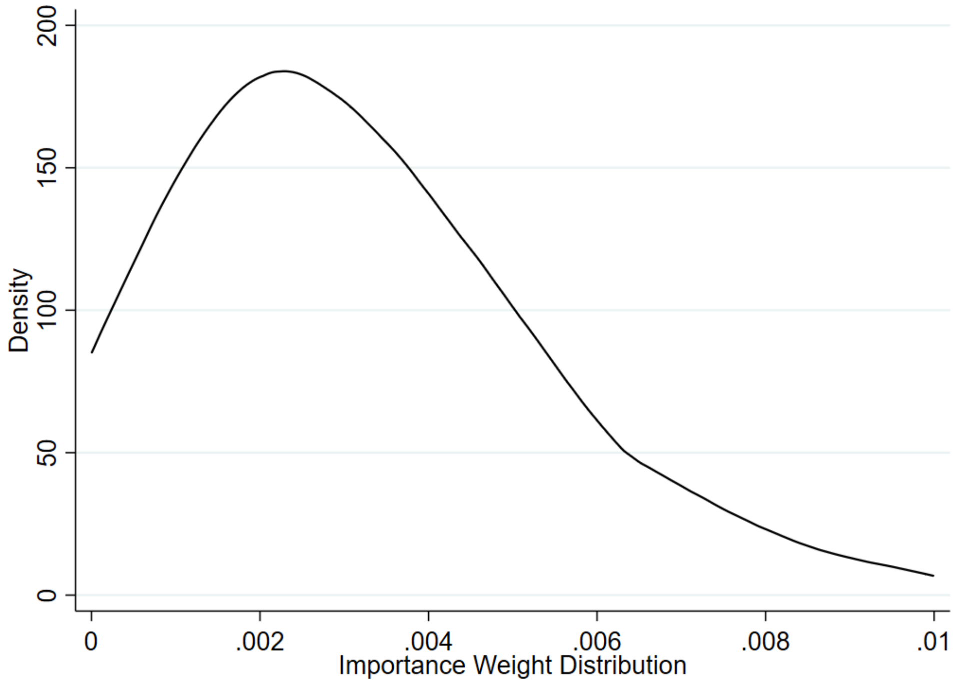


Figure A8: Distribution of Importance Weights (Average Exposure of Each Shock)

Source: Data from Derenoncourt (2022) Author's calculations. This figure reports a kernel density plot of importance weights (average exposure) for each shock. The unit of observation is a Southern county. Kernel is Epanechnikov, and bandwidth set to 0.001.

Table A1: First Stage on Residential Segregation (+Track Length)

	1970	1980	1990	2000
RDI Instrument	0.277***	0.160*	0.321***	0.334***
	(0.0579)	(0.0844)	(0.0758)	(0.0819)
Track Length	0.00831	0.0430***	0.0230*	0.0210*
	(0.0102)	(0.0135)	(0.0120)	(0.0108)
Observations	130	130	130	130
F	14.37	11.91	16.42	15.50

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Atack (2016) and IPUMS NHGIS Extracts from 1970-2000, Author's calculations. This table reports the first stage relationship (coefficients and heteroskedasticity-robust standard errors) between the railroad density instrument and the dissimilarity index segregation measure by decade, controlling for railroad track length. Column 1 reports the results for 1970, 2 reports 1980, 3 reports 1990, and Column 4 reports 2000. The unit of observation is non-Southern Commuting Zones.

Table A2: Effects of Residential Segregation on Interracial Marriage Outcomes (+Track Length)

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
Dissimilarity Index	-2.341	-21.08	-15.45*	-72.75***
	(2.152)	(17.50)	(8.841)	(26.09)
Outcome Mean	1.26	2.43	4.28	14.1
Observations	130	130	130	130
Panel B: Marital Integration				
Dissimilarity Index	-0.0897	-1.733*	-1.020***	-4.306***
	(0.249)	(0.899)	(0.332)	(1.365)
Outcome Mean	.0778	.109	.183	.667
Observations	128	129	130	130

Source: Data from Atack (2016), IPUMS NHGIS Extracts from 1970-2000, and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables present point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. These tables also include a control for railroad track length. In Panel A, the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Panel B, the dependent variable is the marital integration in that Census year. Sample contains 130 non-Southern CZs.

Table A3: Effects of Residential Segregation on Interracial Marriage Outcomes (16 to 35 Year Olds)

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
Dissimilarity Index	-2.729	-15.59	-17.83*	-77.35***
	(3.149)	(12.14)	(10.01)	(28.67)
ymean	1.96	4.23	7.09	23
N	130	130	130	130
Panel B: Scaled Interracial Marriage				
Dissimilarity Index	-0.238	-0.842**	-0.909***	-3.405***
	(0.286)	(0.337)	(0.338)	(1.120)
ymean	.101	.155	.253	.799
N	128	127	128	128

Source: Data from Atack (2016), IPUMS NHGIS Extracts from 1970-2000, and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. Sample restricted to 16 to 35 year olds. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. In Panel A the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Panel B the dependent variable is the marital integration in that Census year. Sample contains 130 non-Southern CZs.

Table A4: Placebo Test of RDI on 1900-1940 IMR

	1900	1910	1920	1930	1940
RDI Instrument	-0.0451	-0.0447	0.0140	-0.283	-0.285
	(0.156)	(0.193)	(0.127)	(0.239)	(0.173)
ymean	.377	.564	.321	.254	.294
N	130	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Source: Data from Atack (2016) and 1900-1940 Full-Count Census. Author's calculations. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the RDI variable. In Columns 1-4, the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year. Sample contains 130 non-Southern CZs.

Table A5: Placebo Test of Great Migration on 1900-1940 IMR

	1900	1910	1920	1930	1940
$\hat{GM}$	-0.000336	-0.000155	-0.000669	-0.00233	-0.000494
	(0.000907)	(0.000976)	(0.00108)	(0.00210)	(0.00121)
ymean	.377	.564	.321	.254	.294
N	130	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Derenoncourt (2022) and 1900-1940 Full-Count Censuses. Author's calculations. This table reports the estimated impact of the Great Migration on observed interracial marriage rates in 1900-1940 (a placebo outcome). The unit of observation is a CZ. The dependent variable is the observed rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A6: Predictiveness of CZ-level Black/Mixed-Race Share for IMR

	1910 IMR	1910 IMR	1920 IMR	1920 IMR
Share Mixed Race	15.38***		14.25***	
	(3.197)		(3.392)	
Share Black		3.692***		2.651***
		(0.918)		(0.571)
ymean	.573	.573	.348	.348
N	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data 1910-1920 Full-Count Censuses. Author's calculations. This table reports bivariate relationship between interracial marriage rates and black/mixed-race population shares. The unit of observation is a CZ. The dependent variable is the observed rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. Standard errors are in parentheses.

Table A7: Reduced Form of RDI on Interracial Marriage Outcomes

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
RDI Instrument	-0.303	-2.411*	-3.677*	-20.61***
	(0.509)	(1.233)	(1.993)	(5.477)
Outcome Mean	1.26	2.43	4.28	14.1
Observations	130	130	130	130
Panel B: Marital Integration				
RDI Instrument	-0.0293	-0.222**	-0.294**	-1.417**
	(0.0563)	(0.0997)	(0.137)	(0.574)
Outcome Mean	.0778	.109	.183	.667
Observations	128	129	130	130

Source: Data from Atack (2016), and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables present point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the RDI variable. In Panel A, the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Panel B, the dependent variable is the marital integration in that Census year. Sample contains 130 non-Southern CZs.

Table A8: OLS of Segregation on Interracial Marriage Outcomes

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
Dissimilarity Index	-0.0226	4.158***	3.772**	-7.593
	(0.727)	(0.873)	(1.893)	(5.051)
Outcome Mean	1.26	2.43	4.28	14.1
Observations	130	130	130	130
Panel B: Marital Integration				
Dissimilarity Index	-0.180**	-0.445***	-0.916***	-3.718***
	(0.0870)	(0.0971)	(0.150)	(0.601)
Outcome Mean	.0778	.109	.183	.667
Observations	128	129	130	130

Source: Data from Atack (2016), IPUMS NHGIS Extracts from 1970-2000, and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables present point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the dissimilarity index in that Census year, Seg_{CZ}. In Panel A, the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Panel B, the dependent variable is the marital integration in that Census year. Sample contains 130 non-Southern CZs.

Table A9: Reduced Form of Predicted Migration on Observed IMR

	1970	1980	1990	2000
$\hat{GM}$	0.00577**	0.0174***	0.0278***	0.0609***
	(0.00269)	(0.00482)	(0.00788)	(0.0216)
Observations	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A10: Reduced Form of Predicted Migration on Expected IMR

	1970	1980	1990	2000
$\hat{GM}$	0.868***	0.931***	0.754***	0.763***
	(0.156)	(0.145)	(0.119)	(0.124)
Observations	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A11: Reduced Form of Predicted Migration on Marital Integration

	1970	1980	1990	2000
$\hat{GM}$	-0.00129***	-0.00156***	-0.00203***	-0.00714***
	(0.000454)	(0.000469)	(0.000507)	(0.00191)
Observations	128	129	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Derenoncourt (2022) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables report the estimated impact of predicted migration on observed interracial marriage rates, expected interracial marriage rates, and marital integration. The unit of observation is a CZ. The dependent variable in Table A9 is the observed rate of interracial marriage per 1000 marriages. The dependent variable in Table A10 is the expected rate of interracial marriage per 1000 marriages. The dependent variable in Table A11 is marital integration. Interracial marriages as a fraction of all Black and white marriages. Marital integration is the observed interracial marriage rate scaled by the expected rate of interracial marriage. The independent variable is the percentile of predicted Black population increase during the Great Migration, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A12: OLS of Great Migration on Observed IMR

	1970	1980	1990	2000
GM	0.00996***	0.0394***	0.0676***	0.151***
	(0.00366)	(0.00501)	(0.00755)	(0.0219)
Observations	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A13: OLS of Great Migration on Expected IMR

	1970	1980	1990	2000
GM	1.299***	1.354***	1.208***	1.254***
	(0.144)	(0.134)	(0.0999)	(0.0978)
Observations	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A14: OLS of Great Migration on Marital Integration

	1970	1980	1990	2000
GM	-0.00214***	-0.00286***	-0.00409***	-0.0185***
	(0.000537)	(0.000676)	(0.000713)	(0.00252)
Observations	128	129	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Derenoncourt (2022) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables report the OLS impact of the Great Migration on observed interracial marriage rates, expected interracial marriage rates, and marital integration. The unit of observation is a CZ. The dependent variable in Table A12 is the observed rate of interracial marriage per 1000 marriages. The dependent variable in Table A13 is the expected rate of interracial marriage per 1000 marriages. The dependent variable in Table A14 is marital integration. Interracial marriages as a fraction of all Black and white marriages. Marital integration is the observed interracial marriage rate scaled by the expected rate of interracial marriage. The independent variable is the percentile of Black population increase during the Great Migration. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A15: Effects of Great Migration on Interracial Marriage Outcomes (16 to 35 Year Olds)

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
GM	0.0140	0.0468***	0.0719***	0.238***
	(0.0123)	(0.0155)	(0.0234)	(0.0448)
ymean	1.96	4.23	7.09	23
N	130	130	130	130
Panel B: Scaled Interracial Marriage				
GM	-0.00400***	-0.00489***	-0.00266*	-0.0183***
	(0.00139)	(0.00109)	(0.00153)	(0.00511)
ymean	.101	.155	.316	.799
N	128	127	128	128
Panel C: Expected Interracial Marriage				
GM	2.163***	2.103***	1.586***	1.769***
	(0.280)	(0.244)	(0.188)	(0.193)
ymean	57.8	56.8	50.9	57.5
N	130	130	130	130

Source: Data from Derenoncourt (2022) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. Sample restricted to 16 to 35 year olds. This table reports the estimated impact of the Great Migration on observed interracial marriage rates, expected interracial marriage rates, and marital integration. The unit of observation is a CZ. The dependent variable in Panel A is of the observed rate of interracial marriage per 1000 marriages. The dependent variable in Panel B is marital integration. The dependent variable in Panel C is the expected rate of interracial marriage per 1000 marriages. Interracial marriages as a fraction of all Black and white marriages. Marital integration is the observed interracial marriage rate scaled by the expected rate of interracial marriage. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A16: Great Migration Outmarriage Effect Heterogeneity

	White Outmarriage	Black Outmarriage
Panel A: High Segregation Cities		
GM	0.00243**	-0.0135
	(0.000968)	(0.0337)
ymean	.211	6.33
N	65	65
Panel B: Low Segregation Cities		
GM	0.00636***	-0.223
	(0.00235)	(0.150)
ymean	.153	12.6
N	65	65

Source: Data from Derenoncourt (2022) and the 2000 5% state IPUMS-USA sample. Author's calculations. This table reports the estimated impact of the Great Migration on outmarriage rates for Black and white respondents (e.g., the share of white married people who are married to a Black person, etc.). Panel A presents results for above-median segregation cities and Panel B presents results for below-median segregation cities. The unit of observation is a CZ. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A17: Great Migration Marriage Effects by Gender and Race

	Outmarriage	Integration	Unmarried	Same Race	Other Race
			Panel A: Black Men		
GM	0.00374***	0.00474***	0.103***	-0.108***	0.000581
	(0.000861)	(0.000935)	(0.0180)	(0.0190)	(0.00379)
ymean	.364	.389	38.9	59.5	1.25
N	130	130	130	130	130
			Panel B: Black Women		
GM	0.00429***	0.00539***	0.102***	-0.109***	0.00321
	(0.00122)	(0.00130)	(0.0191)	(0.0200)	(0.00279)
ymean	.476	.508	43.3	55.3	.93
N	130	130	130	130	130
			Panel C: White Men		
GM	0.00370***	-0.341***	0.0701***	-0.0767***	0.00294
	(0.000714)	(0.0625)	(0.0166)	(0.0176)	(0.00331)
ymean	.284	14.5	40.1	58.5	1.1
N	130	130	130	130	130
			Panel D: White Women		
GM	0.00272***	-0.0601	0.0861***	-0.0912***	0.00234
	(0.000491)	(0.120)	(0.0177)	(0.0185)	(0.00319)
ymean	.198	8.94	40.3	58.4	1.07
N	130	130	130	130	130

Source: Data from Derenoncourt (2022) and the 2000 5% state IPUMS-USA sample. Author's calculations. This table reports the estimated impact of the Great Migration on five marriage outcomes for each race. The unit of observation is a CZ. Panel A reports marriage outcomes for Black men, Panel B for Black women, Panel C for white men, and Panel D for white women. The dependent variable in Column 1 is the number of respondents who are married to the opposite race (Black spouse for white respondents and vice versa) per 100 respondents. The dependent variable in Column 2 is the outmarriage rate (displayed in Column 1) divided by the outmarriage rate under random assignment to spouses. The dependent variable in Column 3 is the number of respondents per 100 respondents who are unmarried. The dependent variable in Column 4 is the number of respondents per 100 respondents who are married to someone of the same race. The dependent variable in Column 5 is the number of respondents per 100 respondents who are married to someone whose race is neither white nor Black. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A18: Effects of Great Migration on Interracial Marriage (Controlling for Sum of Shares)

	1970	1980	1990	2000
GM	0.00496	0.0253***	0.0447***	0.104***
	(0.00649)	(0.00908)	(0.0151)	(0.0335)
Sum of Shares	Yes	Yes	Yes	Yes
Observations	130	130	130	130

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Derenoncourt (2022) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. This table reports the estimated impact of the Great Migration on interracial marriage rates while controlling for the sum of shares in the shift-share design. The unit of observation is a CZ. The independent variable is the percentile of Black population increase during the Great Migration. The instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A19: Rotemberg Negative and Positive Weights

	Sum	**Mean**	**Share**
Negative	-0.001	-0.000	0.001
Positive	1.001	0.001	0.999

Table A20: Rotemberg Correlations of Predicted Migration Aggregates

	α_k	g_k	β_k	F_k	Var(z_k)
α_k	1				
g_k	0.793	1			
β_k	-0.016	-0.010	1		
F_k	-0.036	-0.066	0.004	1	
Var(z_k)	0.158	-0.075	0.040	0.273	1

Notes: These tables summarize statistics about Rotemberg weights, where k indexes counties, following Goldsmith-Pinkham et al. (2020). Table A19 reports share of positive and negative Rotemberg weights. Table A20 reports correlation between the weights (α_k), predicted migration inflows into commuting zones (g_k), the just identified coefficient estimates (β_k), the first stage F-statistic of the historical settlement patterns of Black southern migrants (F_k), and the variation in the shares of Black southern migrants ($Var(z_k)$) residing in the north in 1940.

Table A21: Correlation of CZ-level Segregation Across Years

Variables	1970 Segregation	1980 Segregation	1990 Segregation	2000 Segregation
1970 Segregation	1.000			
1980 Segregation	0.541	1.000		
1990 Segregation	0.634	0.871	1.000	
2000 Segregation	0.624	0.812	0.957	1.000

Notes: This correlation table displays the correlation in CZ-level segregation across years in my sample of 130 commuting zones. Note that 1970 segregation is calculated using county subgroups rather than Census tracts due to data availability.

Table A22: Interracial Couple Characteristics and Predicted Migration

	Born in South	Birthplace Different	Occupational Score	Age	Number of Children
$\hat{GM}$	0.0000494	-0.000754	0.000207	0.0157	-0.00159
	(0.000320)	(0.000489)	(0.0144)	(0.0128)	(0.00131)
ymean	.185	.5	23.4	37.8	1.31
N	129	129	129	129	129

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Derenoncourt (2022) and the 1990 5% state IPUMS-USA sample. Author's calculations. Sample restricted to interracial couples. This table reports the reduced form relationship between predicted migration and the demographic characteristics of interracial couples. The unit of observation is a CZ. The dependent variable in Column 1 is an indicator for whether the respondent was born in the South. The dependent variable in Column 2 is whether the respondent's birthplace differs from their state of residence. The dependent variable in Column 3 is the occupational income score of the respondent. The dependent variable in Column 4 is age of the respondent. The dependent variable in Column 5 is the number of own children in the household for the respondent. The independent variable is the instrument for Black population increase is the percentile of predicted Black population increase, defined as the interaction between pre-1940 Black southern migration patterns and post-1940 outflows of migrants as predicted by southern economic factors alone. Baseline 1940 controls include the share of the urban population made up of 1935–1940 Black southern migrants and census region fixed effects. Standard errors are in parentheses.

Table A23: Observable Characteristics Across CZ-level Segregation

	Below Median Segregation	Above Median Segregation	Total
Black Share	0.0186	0.0610	0.0398
	(0.0200)	(0.0559)	(0.0469)
White Share	0.935	0.910	0.923
	(0.0693)	(0.0793)	(0.0752)
Married Share	0.692	0.688	0.690
	(0.0379)	(0.0376)	(0.0376)
Age	46.34	46.76	46.55
	(1.872)	(1.527)	(1.715)
Occupational Income Score	20.69	20.89	20.79
	(1.456)	(1.720)	(1.591)
Observations	130		

Source: IPUMS NHGIS Extracts from 1990 and 1990 Census. Sample is restricted to population 15 and older. Author's calculations. This table reports means and standard deviations of demographic and economic characteristics across the median of CZ-level segregation (measured by the residential dissimilarity index). Median of commuting zone residential dissimilarity index is 0.59. The unit of observation is a CZ.

Table A24: Commuting zones in sample

Phoenix, AZ	Rockford, IL	Joplin, MO	Youngstown, OH
Tucson, AZ	Springfield, IL	Kansas City, MO	Zanesville, OH
Bakersfield, CA	Center, IN	Springfield, MO	Eugene, OR
Fresno, CA	Concord, IN	St. Joseph, MO	Portland, OR
Los Angeles, CA	Evansville, IN	St. Louis, MO	Allentown, PA
Sacramento, CA	Fort Wayne, IN	Butte–Silver Bow, MT	Altoona, PA
San Diego, CA	Gary, IN	Great Falls, MT	Erie, PA
San Francisco, CA	Indianapolis, IN	Fargo, ND	Hagerstown, PA
San Jose, CA	Lafayette, IN	Lincoln, NE	Harrisburg, PA
Santa Barbara, CA	Muncie, IN	Omaha, NE	Philadelphia, PA
Colorado Springs, CO	South Bend, IN	Manchester, NH	Pittsburgh, PA
Denver, CO	Terre Haute, IN	Newark, NJ	Reading, PA
Pueblo, CO	Wayne, IN	Albuquerque, NM	Scranton, PA
Bridgeport, CT	Hutchinson, KS	Albany, NY	Williamsport, PA
Washington, DC	Topeka, KS	Amsterdam, NY	Providence, RI
Wilmington, DE	Wichita, KS	Buffalo, NY	Sioux Falls, SD
Burlington, IA	Louisville, KY	Elmira, NY	Salt Lake City, UT
Cedar Rapids, IA	Boston, MA	New York, NY	Burlington, VT
Clinton, IA	Pittsfield, MA	Poughkeepsie, NY	Bellingham, WA
Des Moines, IA	Springfield, MA	Syracuse, NY	Seattle, WA
Dubuque, IA	Baltimore, MD	Union, NY	Spokane, WA
Mason City, IA	Cumberland, MD	Watertown, NY	Yakima, WA
Ottumwa, IA	Bangor, ME	Canton, OH	Eau Claire, WI
Sioux City, IA	Portland, ME	Cincinnati, OH	Green Bay, WI
Waterloo, IA	Detroit, MI	Cleveland, OH	Kenosha, WI
Bloomington, IL	Grand Rapids, MI	Columbus, OH	La Crosse, WI
Chicago, IL	Jackson, MI	Dayton, OH	Madison, WI
Davenport, IL	Kalamazoo, MI	Lima, OH	Milwaukee, WI
Decatur, IL	Lansing, MI	Lorain, OH	Oshkosh, WI
Edwardsville, IL	Saginaw, MI	Mansfield, OH	Sheboygan, WI
Galesburg, IL	Duluth, MN	Scioto, OH	Wausau, WI
Peoria, IL	Minneapolis, MN	Steubenville, OH	
Quincy, IL	Rochester, MN	Toledo, OH	

Notes: Name refers to largest city in the commuting zone.

Appendix B

B.1 Derivation of Outmarriage Rates

Fix a group $g \in \{W, B\}$ and write its cross-race meeting probability as q:

$$q = \begin{cases} q_{WB}(x, s) = (1-s)\, x & \text{if } g = W, \\ q_{BW}(x, s) = (1-s)\,(1-x) & \text{if } g = B. \end{cases}$$

Let $a_S = 1 - F(-v_S)$ and $a_I = 1 - F(h + \kappa_t - v_S)$ be the acceptance probabilities for same-race and interracial proposals. In any meeting, define the per-meeting "acceptance hazards"

$$p_I \equiv q\, a_I, \qquad p_S \equiv (1-q)\, a_S, \qquad r \equiv 1 - p_I - p_S.$$

A marriage occurs at the first accepted proposal (I assume no divorce/exit). The probability the first acceptance is interracial is

$$\Pr(\text{first acceptance is interracial}) = p_I + r\, p_I + r^2\, p_I + \cdots = \Big(\sum_{k=0}^{\infty} r^k\Big)\, p_I = \frac{p_I}{1-r} = \frac{p_I}{p_I + p_S},$$

where the geometric sum is valid since $r = 1 - p_I - p_S < 1$ whenever $p_I + p_S > 0$. Therefore, individual out-marriage propensities are

$$\begin{aligned} \text{Out}_W(x, s, \kappa_t) &= \frac{q_{WB}(x, s)\, a_I(h, \kappa_t)}{q_{WB}(x, s)\, a_I(h, \kappa_t) + [1 - q_{WB}(x, s)]\, a_S}, \\ \text{Out}_B(x, s, \kappa_t) &= \frac{q_{BW}(x, s)\, a_I(h, \kappa_t)}{q_{BW}(x, s)\, a_I(h, \kappa_t) + [1 - q_{BW}(x, s)]\, a_S}. \end{aligned}$$

B.2: Endogenous Segregation

I can also rederive my comparative statics when I allow segregation to respond to Black population shares. Then, the expressions for meeting probabilities become:

$$q_{WB}(x, s) \equiv (1 - s(x))\, x, \qquad q_{BW}(x, s) \equiv (1 - s(x))\,(1 - x),$$

Thus, when I rederive my comparative statics for the effects of increasing the Black population share, I see that they take a slightly different form:

$$\frac{\partial \text{Out}_W}{\partial x} = \frac{a_I a_S}{D_W^2}(1 - s(x) - xs'(x)); \quad \frac{\partial \text{Out}_B}{\partial x} = \frac{a_I a_S}{D_B^2}(-1 + s(x) - (1-x)s'(x))$$

Intuitively, when segregation does not respond to Black population increases ($s'(x) = 0$), these simplify to the same expressions outlined in the main text. However, if segregation rises in response to Black population increases ($s'(x) > 0$), then these expressions suggest that this response will dampen the increase in white outmarriage, perhaps turning the effect negative (if $s'(x)$ is large enough), and amplify the decline in Black outmarriage. Thus, segregation responses can serve as a countervailing force that limits the social integration gains to minority in-migration.

Appendix C

C.1 Great Migration Instrument Construction

The instrument is constructed as follows: I replace the numerator in the black population change measure with the predicted, instead of observed, increase in the Black population:

$$\text{Predicted Black pop}_{CZ}^{1940-1970} = \frac{\hat{\Delta b}_{\text{urban},CZ}^{1940-1970}}{\text{pop}_{\text{urban},CZ}^{1940}}$$

where $\hat{\Delta b}_{\text{urban},CZ}^{1940-1970}$ is the predicted increase, defined as follows:

$$\hat{\Delta b}_{\text{urban},CZ}^{1940-1970} = \sum_{j \in S} \sum_{c \in CZ} \omega_{jc}^{1935-1940} \cdot \hat{m}_j^{1940-1970}$$

and $\omega_{jc}^{1935-1940}$ is the share of recently migrated pre-1940 Black southern migrants from county j living in city c in 1940. These shares are computed using the IPUMS version of the complete count 1940 Census (Ruggles et al., 2021; Derenoncourt, 2022). The 1940 Census required respondents to report their 1935 place of residence, so I classify Black Southern migrants as all Black Southerners whose place of enumeration in 1940 (whether in the South or not) does not match the Southern county of residence (j) reported in 1935. Among this population, the shares are defined as the ratio of Black Southern migrants from county j who are enumerated in a northern city c in 1940 over all Black Southern migrants from county j.

The term $\hat{m}_j^{1940-1970}$ is the predicted Black migration from southern county j, which comes from the sum of fitted values of decadal predictions of southern county net migration using lagged southern economic predictors of migration. This prediction stage uses economic characteristics of southern characteristics (such as reliance on cotton, or WWII spending per capita) to predict how many Black people leave each Southern county j each decade, which are then summed to compute $\hat{m}_j^{1940-1970}$, the aggregate predicted Black outmigration for Southern county j. More formally, $\hat{m}_j^{1940-1970} = \sum_{t=1950}^{1970} \widehat{\text{mig rate}}_{jt} \times \text{Black pop}_{jt}$, where decadal predictions of net migration rates from each Southern county ($\widehat{\text{mig rate}}_{jt}$) come from the following regression:

$$\text{mig rate}_{jt} = \beta_0 + Z_{jt-10}^{'}\beta_1 + \varepsilon_{jt}$$

As Derenoncourt (2022) discusses, under the assumption that economic shocks to Southern counties are exogenous to receiving city characteristics, predicting outmigration using these factors is a pure prediction problem. Thus, she employs least absolute shrinkage and selection operator (LASSO) methods to select the set of lagged predictors, $Z_{jt-10}^{'}$. The initial set of predictors that is selected from comes from Boustan

(2010): percent acreage in cotton, percent tenant farms, share of labor force in agriculture, indicator for being in a tobacco-growing state, interaction between tobacco growing indicator and share of labor force in agriculture, WWII spending per capita, share of the labor force in mining, indicator for being a mining state, and interaction between mining state indicator and share of labor force in mining. Using LASSO to select the predictors for each decade, the regression predicts Southern county-level net migration figures from Boustan (2010, 2016). Then, the predicted values from these regressions are used for each decade and summed to construct the predicted outmigration from each Southern county j over the course of the Second Great Migration: $\hat{m}_j^{1940-1970}$. The main text and appendices of Derenoncourt (2022) contain more detail about this procedure, including which predictors were selected for each decade.

These predicted values are the "push" component of the shift-share instrument, and are interacted with the distribution of pre-1940 Black Southern migrants (the "pull") to generate a predicted increase in Black population for destination cities in the North. Then, after computing predicted increases in the northern CZ-level Black population, I use the percentile of predicted increases, $\widehat{GM}_{CZ}$, to instrument for observed increases in the Black population. More formally:

$$\widehat{GM}_{CZ} \equiv 100 \times \frac{1}{|\mathcal{CZ}|} \sum_{CZ' \in \mathcal{CZ}} \mathbf{1}\Big\{\widehat{\Delta B}_{CZ'} \leq \widehat{\Delta B}_{CZ}\Big\}.$$

Appendix D

D.1 Railroad IV Assumptions

More formally, this design relies on three crucial assumptions to identify the causal effect of segregation on interracial marriage outcomes:

1. The RDI serves as a valid instrument for the contemporary segregation, meaning it has a strong relationship with the causal variable of interest.

2. The RDI instrument is independent of potential outcomes (in this case potential rates of interracial marriage at the city level).

3. The RDI instrument only affects interracial marriage outcomes through segregation—the exclusion restriction.

Assumption (1) is equivalent to the strength of the first stage—I provide evidence in the results that the RDI instrument is strongly predictive of residential segregation, and the first-stage F-statistic for 1990 is 27.1. Other work has also verified the strength of this relationship (Ananat and Washington, 2009; Ananat, 2011; Chyn et al., 2022; Cox et al., 2022) in a sample of 121 non-Southern cities. Assumption (2) is inherently untestable, but I provide some evidence by examining the relationship between the instrument and placebo interracial marriage outcomes in 1900-1940, before the segregation differences between high RDI and low RDI commuting zones emerge. Appendix Table A4 displays these results for interracial marriage. These coefficients are all insignificant and small in magnitude, suggesting that prior to segregation differences, interracial marriage outcomes across places with high and low RDIs were similar. In terms of magnitude, going from a city with 1 railroad neighborhood to infinite railroad neighborhoods is associated with 0.28 fewer interracial marriages per 1000 marriages. Given, the prevalence of interracial marriage at the time is rather low so this effect would be large in percent terms, but the coefficients are an order of magnitude smaller than those I find when examining outcomes later in the 20th century.

Assumption (3) states formally that historical railroad placement measured via RDI_{CZ} is only related to interracial marriage outcomes through segregation and is untestable. This identification rises in part from geographic factors like hill placement and distance that may have determined both the extent of railroad track development in a city and the layout of that track (Chyn et al., 2022). I present evidence that my results are robust to the inclusion of historical railroad track length as a control as support for this assumption.

D.2 Replication of Segregation Results using Ananat (2011) Sample

For supplementary segregation analyses, I use the sample of 121 non-Southern metropolitan areas for which Ananat (2011) located 19th-century maps needed to construct the railroad placement instrumental variable. This sample of cities was constrained in part by whether maps of railroad placement were available, but Ananat (2011) shows that this sample is similar to the full sample of cities for which segregation measures are estimable in their Table A. Additionally, data from Cutler et al. (1999) is used to measure metropolitan residential segregation in the years 1970-2000.[40] I combine these two datasets and match them to IPUMS extracts using MSA codes to calculate interracial marriage outcomes for the sample from 1970-2000. Then, I replicate the empirical strategy outlined in Section 4.2 using this sample of cities to assess the robustness of the results. Overall, the results are robust to this sample change, displaying similar patterns of effect sign, size, and statistical significance.

[40] They provide data on the residential dissimilarity index, which is calculated as follows for city c:

$$\text{Seg}_c = \frac{1}{2}\sum_{n \in c}\left|\frac{\text{Black}_n}{\text{Black}_c} - \frac{\text{White}_n}{\text{White}_c}\right|$$

where Black_n and White_n are the population counts of Black and white people, respectively, in sub-city geographic area n (in this case Census tracts), while Black_c and White_c are population counts for Black and white people over the entire city c.

Table D1: First Stage on Residential Segregation

	1970	1980	1990	2000
RDI Instrument	0.355***	0.420***	0.406***	0.429***
	(0.0918)	(0.0794)	(0.0820)	(0.0956)
Observations	69	87	104	96
F	14.97	27.96	24.50	20.13

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Ananat (2011); Cutler et al. (1999), Author's calculations. This table reports the first stage relationship (coefficients and heteroskedasticity-robust standard errors) between the railroad density instrument and the dissimilarity index segregation measure by decade. Column 1 reports the results for 1970, 2 reports 1980, 3 reports 1990, and Column 4 reports 2000. The unit of observation is non-Southern metro areas for which both segregation and RDI are available.

Table D2: Effects of Residential Segregation on Interracial Marriage Outcomes

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
Dissimilarity Index	1.196	-4.539	-8.425*	-24.82**
	(3.132)	(4.626)	(5.039)	(11.73)
Outcome Mean	1.49	3.45	5.46	18
Observations	48	80	104	95
Panel B: Marital Integration				
Dissimilarity Index	-0.0409	-0.487***	-0.933***	-4.420***
	(0.0749)	(0.158)	(0.333)	(1.429)
Outcome Mean	.0306	.0762	.16	.567
Observations	48	80	104	95

Source: Data from Ananat (2011); Cutler et al. (1999) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from 2SLS models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. In Panel A the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Panel B the dependent variable is the marital integration in that Census year. Sample contains those non-Southern metro areas which can be matched to the Census MSA codes and are present in the Ananat (2011); Cutler et al. (1999) data.

Table D3: Decomposing Residential Segregation Marriage Effects

	Outmarriage	Integration	Unmarried	Same Race	Other Race
Panel A: White Marriage Outcomes					
Dissimilarity Index	0.108	-25.12***	-12.26*	17.11**	-4.960***
	(0.158)	(7.394)	(6.294)	(6.857)	(1.460)
ymean, %	.215	5.98	42.1	56.3	1.45
N	95	95	95	95	95
Panel B: Black Marriage Outcomes					
Dissimilarity Index	-25.62***	-25.51***	12.38	24.86**	-11.62***
	(7.547)	(7.498)	(8.687)	(11.41)	(2.653)
ymean, %	5.82	6.03	66.3	25.9	1.94
N	95	95	95	95	95

Source: Data from Ananat (2011) and the 2000 5% state IPUMS-USA sample. Author's calculations. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from 2SLS models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. Panel A reports marriage outcomes for white respondents, while Panel B reports marriage outcomes for Black respondents. The dependent variable in Column 1 is the number of respondents who are married to the opposite race (Black spouse for white respondents and vice versa) per 100 respondents. The dependent variable in Column 2 is the outmarriage rate (displayed in Column 1) divided by the outmarriage rate under random assignment to spouses. The dependent variable in Column 3 is the number of respondents per 100 respondents who are unmarried. The dependent variable in Column 4 is the number of respondents per 100 respondents who are married to someone of the same race. The dependent variable in Column 5 is the number of respondents per 100 respondents who are married to someone whose race is neither white nor Black. Sample contains those non-Southern metro areas which can be matched to the Census MSA codes and are present in the Ananat (2011); Cutler et al. (1999) data.

Table D4: Placebo Test of Segregation on 1930 IMR

	IMR	IMR	Integration	Integration
RDI Instrument	-0.255	-0.243	-0.0149	0.00261
	(0.252)	(0.252)	(0.0322)	(0.0410)
Track Length		-5.739		-8.079
		(6.034)		(6.274)
ymean	.346	.346	.0344	.0344
N	113	113	113	113

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Source: Data from Ananat (2011) and 1930 Full-Count Census. Author's calculations. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the RDI variable. Columns 2 and 4 also include a control for historical railroad track length per square kilometer. In Columns 1 and 2, the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Columns 3 and 4, the dependent variable is the marital integration in that Census year. Sample contains those non-Southern metro areas which can be matched to the Census city codes in 1940 and are present in the Ananat (2011) data.

Table D5: First Stage on Residential Segregation (+Track Length)

	1970	1980	1990	2000
RDI Instrument	0.343***	0.386***	0.366***	0.386***
	(0.0922)	(0.0793)	(0.0832)	(0.0962)
Track Length	5.875**	14.21***	17.84*	19.08*
	(2.541)	(5.322)	(10.31)	(11.16)
Observations	69	87	104	96

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Ananat (2011); Cutler et al. (1999), Author's calculations. This table reports the first stage relationship (coefficients and heteroskedasticity-robust standard errors) between the railroad density instrument and the dissimilarity index segregation measure by decade, controlling for railroad track length. Column 1 reports the results for 1970, 2 reports 1980, 3 reports 1990, and Column 4 reports 2000. The unit of observation is non-Southern metro areas for which both segregation and RDI are available.

Table D6: Effect of Segregation on Observed IMR (+Track Length)

	1970	1980	1990	2000
Dissimilarity Index	1.026	-5.913	-9.568*	-25.76**
	(3.863)	(5.070)	(5.645)	(13.02)
ymean	1.49	3.45	5.46	18
N	48	80	104	95

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D7: Effect of Segregation on Marital Integration (+Track Length)

	1970	1980	1990	2000
Dissimilarity Index	0.0408	-0.490***	-0.957**	-4.605***
	(0.124)	(0.168)	(0.381)	(1.635)
ymean	.0306	.0762	.16	.567
N	48	80	104	95

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Ananat (2011); Cutler et al. (1999) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables present point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. These tables also include a control for railroad track length. In Table D6, the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Table D7, the dependent variable is the marital integration in that Census year. Sample contains those non-Southern metro areas which can be matched to the Census MSA codes and are present in the Ananat (2011); Cutler et al. (1999) data.

Table D8: Reduced Form of RDI on Observed IMR

	1970	1980	1990	2000
RDI Instrument	0.885	-0.315	-3.419	-10.70*
	(1.022)	(1.773)	(2.082)	(5.557)
Observations	50	88	104	95

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D9: Reduced Form of RDI on Marital Integration

	1970	1980	1990	2000
RDI Instrument	0.000744	-0.155**	-0.379***	-1.905***
	(0.0269)	(0.0767)	(0.116)	(0.543)
Observations	50	88	104	95

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Ananat (2011) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables present point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the RDI variable. In Table D8, the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Table D9, the dependent variable is the marital integration in that Census year. Sample contains those non-Southern metro areas which can be matched to the Census MSA codes and are present in the Ananat (2011); Cutler et al. (1999) data.

Table D10: OLS of Segregation on Observed IMR

	1970	1980	1990	2000
Dissimilarity Index	0.895	-1.205	-2.364	-21.88***
	(1.141)	(1.765)	(2.130)	(3.924)
Observations	48	80	104	95

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table D11: OLS of Segregation on Marital Integration

	1970	1980	1990	2000
Dissimilarity Index	-0.0802*	-0.329***	-0.555***	-2.825***
	(0.0465)	(0.0826)	(0.0969)	(0.520)
Observations	48	80	104	95

Standard errors in parentheses
* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Source: Data from Cutler et al. (1999); Ananat (2011) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. These tables present point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the dissimilarity index in that Census year, Seg_c. In Table D10, the dependent variable is the observed interracial marriage rate per 1000 marriagesin that Census year, and in Table D11, the dependent variable is the marital integration in that Census year. Sample contains those non-Southern metro areas which can be matched to the Census MSA codes and are present in the Ananat (2011); Cutler et al. (1999) data.

Table D12: Effects of Residential Segregation on Interracial Marriage Outcomes (16 to 35 Year Olds)

	1970	1980	1990	2000
Panel A: Observed Interracial Marriage				
Dissimilarity Index	-2.511	-11.78	-13.77*	-31.24
	(6.269)	(8.200)	(7.773)	(19.12)
ymean	2.12	5.99	9.29	30
N	48	80	104	95
Panel B: Scaled Interracial Marriage				
Dissimilarity Index	-0.119*	-0.797***	-2.080**	-4.114***
	(0.0713)	(0.301)	(0.954)	(1.483)
ymean	.0288	.117	.255	.693
N	48	80	104	94

Source: Data from Ananat (2011); Cutler et al. (1999) and the following IPUMS-USA samples: 1970 2% metro, 1980 5% state, 1990 5% state, and 2000 5% state. Author's calculations. Sample restricted to 16 to 35 year olds. This table presents point estimates and heteroskedasticity-robust standard errors (in parentheses) from regression models in which the key independent variable is the dissimilarity index in that Census year, instrumented by the RDI variable. In Panel A the dependent variable is the observed interracial marriage rate per 1000 marriages in that Census year, and in Panel B the dependent variable is the marital integration in that Census year. Sample contains those non-Southern metro areas which can be matched to the Census MSA codes and are present in the Ananat (2011); Cutler et al. (1999) data.