Can you move to opportunity? Evidence from the Great Migration

Ellora Derenoncourt*

December 31, 2019

Click here for most recent version.

Abstract

This paper shows that racial composition shocks during the Great Migration (1940-1970) lowered black upward mobility in the northern United States. I identify northern black population increases using a shift-share instrument, interacting pre-1940 black migrants' location choices with predicted southern county out-migration. The Migration's effects on children are driven by locational factors, not negative selection of families. Using data I assembled on destinations from 1920-2015, I show the Migration led to persistent segregation and higher police spending, crime, and incarceration from the 1960s onwards. The changes induced by the Migration explain 27% of the region's racial upward mobility gap today.

*Derenoncourt: Princeton University. Email: ellorad@princeton.edu. An earlier version of this paper was released on January 25, 2019. I thank Joshua Abel, Ran Abramitzky, Joseph Altonji, Kirill Borusyak, Leah Boustan, Raj Chetty, Krishna Dasaratha, Melissa Dell, Rebecca Diamond, James Feigenbaum, Edward Glaeser, Claudia Goldin, Nathaniel Hendren, Lawrence Katz, Maximilian Kasy, Michal Kolesár, Ilyana Kuziemko, Trevon Logan, Robert Margo, Elizabeth Mishkin, Christopher Muller, Suresh Naidu, Nathan Nunn, Thomas Piketty, Emmanuel Saez, Heather Sarsons, Niharika Singh, Isaac Sorkin, Marianne Wanamaker, Gavin Wright, Chenzi Xu, and numerous seminar and conference participants for many helpful comments. Price Fishback, William Collins, Robert Margo, Vicky Fouka, Soumyajit Mazumder, and Marco Tabellini generously shared data. Ariel Gomez, Sergio Gonzales, Julian Duggan, Seung Yong Song, Lukas Althoff, and Will McGrew provided excellent research assistance. This work was supported by the Harvard Lab for Economic Applications and Policy and Russell Sage Foundation award #83-17-19. Any opinions expressed are those of the author alone and should not be construed as representing the opinions of the Foundation.

1 Introduction

The northern United States historically offered black families a pathway to economic mobility. In 1940, black children from similar economic backgrounds fared substantially better in the North than in the South (Card et al., 2018). Today, however, no such northern advantage exists, and racial gaps in upward mobility—defined as children's adult outcomes conditional on parent economic status—are similar across the country (Davis and Mazumder, 2018; Chetty et al., 2018).

The regional shift in black upward mobility coincided with a decisive moment in black geographic mobility. Between 1940 and 1970, four million African Americans left the South and settled in urban areas in the north and west of the country. The Great Migration, as it is known today, radically transformed the racial demographics of destination cities, prompting white flight from urban neighborhoods and potentially altering the policies of local governments (Boustan, 2010).[1] The relationship between these two phenomena—the migration North and declines in upward mobility—is important for assessing the stability of childhood location or neighborhood effects, shown to be consequential in a variety of experimental and quasi-experimental settings, but whose mutability in response to shocks has not been assessed (Chetty et al., 2016; Chetty and Hendren, 2018a,b).

This paper shows that northern cities' responses to the Great Migration (also termed "Migration") ultimately reduced the gains from growing up in destination locations. The effects have been particularly detrimental for black men. Those growing up in former Great Migration destinations today have lower adult income than those from similarly resourced families but in locations less affected by the Migration. The channel appears to be changes in the environment for families, rather than ex-post sorting of negatively selected families into destinations. In response to black migrant arrivals in the

[1] Tabellini (2018) finds that the first wave of the Great Migration lowered city government expenditures on education.

mid-century, white families withdrew from shared urban neighborhoods and public schools. By the late 1960s, riots broke out in Great Migration commuting zones, and in the subsequent decades destination cities increased police spending, suffered from higher murder rates, and incarcerated a greater share of the population. Today, roughly 27% of the gap in upward mobility between black and white families in the urban North can be attributed to changes induced by the Great Migration.

I draw on a large number of data sources to conduct the analysis in this paper. To establish the main results on upward mobility, I use the complete count US censuses from 1920-1940 and contemporary measures from Chetty and Hendren (2018b) and Chetty et al. (2018). To understand mechanisms, I assembled a new database on local government expenditures, private schools, crime, incarceration, and other characteristics of commuting zones spanning the period 1920-2015. I digitized information on local government spending from the Financial Statistics of State and Local Governments and the Census of Local Governments; data on schooling from the Biennial Statistics of Education and the Census; urban murder rates from the Uniform Crime Reports; Census reports on local county jail populations, and the Vera Institute of Justice's In Our Backyards database on the county of commitment of federal and state prisoners for the more recent periods. This newly harmonized database is now available on my website for other researchers to use.

The empirical strategy makes use of the fact that black southern migrants settled in northern cities where previous migrants from their communities had moved, giving rise to highly specific linkages between southern locations and northern destinations (Boustan, 2010; Black et al., 2015; Stuart et al., 2018). To address omitted factors that may codetermine increases in the urban black population during the Great Migration and declines in upward mobility, I use a "shift-share" approach. I combine information on pre-1940 black southern migrants' location choices with supply-side variation in county outmigration from 1940-1970,[2] predicted from southern economic variables. As the set of

[2]One example is variation in the share of agricultural land planted in cotton. Cotton

these variables is potentially large, I use a machine learning technique, Least Absolute Shrinkage and Selection Operator ("LASSO"), to optimize the set of predictors of net-migration rates from the South. Assigning inflows to cities according to historical settlement patterns yields the predicted increase in the black population from southern variation alone, which I normalize by the initial 1940 urban population.[3] Black in-migration is a right-skewed distribution, so I define the Migration shock to a commuting zone to be the percentile of predicted black population increase.[4]

Using this strategy, I show that the Migration led to a reduction in upward mobility in destination commuting zones in the North today. A 1 standard deviation larger increase in the black population, approximately a 30-percentile increase in the shock, lowered adult income rank of children from low-income families by 3 percentiles, approximately a 9% drop in adult income. As a benchmark, a 1 standard deviation increase in residential racial segregation lowers adult income by about 5.2%.[5]

Two potential mechanisms underlie this effect: selection, or changes in the characteristics of the average resident family; and location, or changes in local public goods or neighborhood quality. To disentangle these two channels, I use data on the childhood exposure effects of commuting zones from Chetty and Hendren (2018b). These data contain estimates of each commuting zone's causal effect on children's adult outcomes today. I examine whether the causal effect of a commuting zone varies with exogenous historical increases in the black population. The interpretation is as follows: if an arbitrary child were to spend one additional year in a Great Migration city versus one less affected by the Migration, how does this affect his or her income as an adult? I estimate

mechanization accelerated after World War II, contributing to black outmigration from the South (Whatley, 1985); variation in cotton acreage thus provides plausible variation in southern county migration rates.

[3]Normalizing by the initial urban population accounts for potentially different growth paths in the urban population across CZs.

[4]See Sequeira et al. (2019) for a similar scaling of estimated effects. The authors report the impact of percentile increases in historical European immigration on long-run economic development in US counties.

[5]See Chetty and Hendren (2018b).

a robust negative effect of the Migration on this measure of upward mobility. My estimates suggest that the cumulative effect of spending one's entire childhood in a Great Migration city accounts for all of the negative impact of the Migration on average upward mobility. In other words, I find no evidence that negative selection of families contributes to the association between historical racial composition shocks and declines in upward mobility.

Next I explore which groups of children were affected by the Migration. The largest negative effects manifest for black men. I find no impact of the Migration on the household income of black women, with some evidence of weakly positive effects on their individual earnings. The evidence is consistent with an income effect: black women who formed or would have formed households with black men increased their labor supply to make up for men's reduced income. Nonetheless, the higher individual earnings of black women do not offset overall reductions in black household income in cities that experienced greater inflows during the Migration.

To understand what characteristics of locations changed as a result of the Migration and thus potentially explain the Migration's persistent effect on upward mobility today, I use the data I assembled on local governments, schools, and crime in commuting zones from 1920-2015. I use the same empirical strategy described above to estimate the impact of the Great Migration on potential mechanisms over time. Pre-1940 outcomes serve as placebo checks. My analysis reveals significant and persistent responses in the following areas: decreases in white public school enrollment and urban residence within the commuting zone; higher local government expenditures on police and higher murder rates; and increased rates of incarceration. The late 1960s were a turning point: the race riots that broke out across major American cities were more severe in Great Migration destinations and the racial attitudes of voters aligned more closely with southern segregationist political views. Punitive responses to urban decline and rising crime, including greater investment in police and incarceration appear to have had long-lasting effects, particularly for black men growing up in these locations.

Many black southerners moved to manufacturing centers during the 1950s and 1960s and these may have undergone greater job loss as manufacturing declined. I do not find, however, that the results in this paper merely reflect the consequences of deindustrialization. In all specifications, I control for the share of the labor force in manufacturing in 1940, which largely accounts for variation in manufacturing shares in subsequent decades. Results are also robust to including a Bartik instrument for employment changes using variation in industry composition interacted with national changes in industry-level employment between 1940 and 1970. Finally, I find no effect of the Migration on the outcomes of white men from low-income families, a group likely to have been affected if the findings were driven by deindustrialization alone. Although all destination cities likely underwent manufacturing job loss, the degree of job loss at the commuting zone level does not appear strongly correlated with black migrant inflows during the Great Migration. What is more likely is that a restructuring of economic activity within commuting zones left black families in the urban core without adequate opportunities while white families potentially followed jobs by moving to growing suburban areas. This is in line with the historical and sociological literature on this topic, e.g., Sugrue (1996) and Wilson (1990).

I investigate the extent to which the results reflect responses to southern black migration specifically. White southerners also migrated to northern cities over the 20th century. I instrument for white southern inflows and show that these have no effect on black upward mobility or on the gains to growing up in specific commuting zones. Second, European Mass Migration affected many northern cities in the late 19th and early 20th century. My results are robust to controlling for historical European migration into Great Migration destinations.[6] To determine whether declines in upward mobility reflect fixed characteristics of locations with high black population shares, I show consistent results using first-differenced measures of black men's upward mobility,

[6]I use a measure of historical European migration from Sequeira et al. (2019) that instruments for whether a county was connected to railways during migration booms versus busts during the Age of Mass Migration.

suggesting that changes in the racial composition, not simply the levels of the black population or other immutable destination features, help explain the findings.

A large literature seeks to identify neighborhood effects and the impact of residential segregation and urban poverty on children's outcomes.[7] More recently, both experimental and quasi-experimental studies have shown childhood location to be an important determinant of adult outcomes and that substantial variation in these effects exists across the US (Chetty et al., 2016; Chetty and Hendren, 2018a,b). However, the stability of these effects in response to shocks is much less understood. I show that large mid-century shifts in the racial composition of northern cities altered the effects locations had on children, turning high opportunity locations into opportunity deserts, particularly for black families.

This paper provides a new long-run intergenerational perspective on the Great Migration. Papers studying the contemporaneous effects of the Great Migration found largely positive impacts on migrants themselves, particularly in terms of income (Collins and Wanamaker, 2014; Boustan, 2016). An exception is Black et al. (2015) who find increased mortality and lower longevity of black migrants in the urban North, relative to stayers from the deep South.[8] To my knowledge, this is the first paper to consider the long-run impacts of the Great Migration on outcomes for the third generation living in the North.[9] The results of this study suggest that across the North, responses to the Great

[7]For literature on this topic, see Ananat (2011); Andrews et al. (2017); Cutler and Glaeser (1997); Massey and Denton (1993); Graham (2016); Sampson et al. (2002); Wilson (1990).

[8]Papers focusing on the earlier period of the Migration (1910-1930) have shown that the Migration increased residential racial segregation (Shertzer and Walsh, 2016), lowered city government expenditures (Tabellini, 2018), and aided the assimilation of European immigrants (Fouka et al., 2018). Two studies examine the effects of the first wave of the Great Migration on incarceration. Muller (2012) finds that the Migration increased racial disparities in incarceration in the North, and Eriksson (2018) shows that migrating North increased black men's likelihood of incarceration.

[9]Leibbrand et al. (2019) consider the differences in neighborhood of residence at older ages between children of migrants in the North and those of non-migrants in the South. The study concludes that the children of migrants live in better neighborhoods but that some of this difference can be explained by positive selection of the migrants.

Migration worsened neighborhood environments. These changes were so dramatic that outcomes for the third generation in the North look no better today than for black children growing up in the South.

An important component of the relationship between the Great Migration and intergenerational mobility that this paper does not speak to, however, is the causal effect of the Migration on the descendants of migrants themselves. The best estimates suggest that moving North nearly doubled the wages of migrants compared to those who stayed behind in the South (Boustan, 2016). Thus the children and grandchildren of migrants living in the North likely benefited from their parents and grandparents moving up in the national income distribution. Losses incurred through northern cities' responses to the Migration must be placed in context with overall improvements in black economic status from moving North.

The paper is structured as follows. Section 2 gives an overview of the historical context. Section 3 describes the data sources, including on upward mobility and black population change in northern cities and provides some descriptive evidence on the relationship between the two. Section 4 describes my empirical strategy for identifying the causal impact of the Migration. In Section 5, I present the main results on upward mobility and on the contribution of selection versus location to these findings. In Section 6, I present results on local mechanisms that may explain the persistent effects of the Migration. Section 7 concludes.

2 Historical background

"My mother was my inspiration... she was one of those 6,000,000 black people who left the South so that her children wouldn't have to grow up and put up with what she had to grow up and put up with." - Helen Singleton, Civil Rights activist from Los Angeles

Starting in the 1910s, black Americans migrated in large numbers from southern states to northern states, a phenomenon known as the Great Migration.[10] By the middle of the 20th century, the Migration was so great that the share of the black population in the South fell to just over 50% by 1970, from 90% in 1910.

Under Jim Crow laws in the South, black Americans faced significant limitations on their political, social, and economic freedoms. Declining labor demand in southern agriculture gradually loosened the largely rural black population's ties to the land. Further, job opportunities for black workers opened up in many northern cities. As a result of these changes, black migrants increasingly undertook the journey north.[11] In doing so, they sought better lives for themselves and their children, and for many decades, the North appeared to deliver on this promise.

Helen Singleton, the daughter of a migrant and later an activist in the Civil Rights Movement, recalled her surprise hearing about *Brown v. Board of Education*, the US Supreme Court ruling that rendered segregated schooling unconstitutional. Having attended high school in Los Angeles, California, the concept of a segregated school was foreign to her. By contrast, for many black children in the South, even those from educated families, the paucity of public black high schools made secondary schooling very costly (Margo, 1990, 1991a). Singleton's experience was reflected more broadly in educational patterns for black children across the US in 1940.

Figure 1a shows the fraction of black teenagers from median-educated households who obtained 9 or more years of schooling. The map illustrates stark differences in upward mobility for black children in the North compared

[10]For a comprehensive study of the Great Migration and its contemporaneous economic impacts on destination cities, see Boustan (2016).

[11]See Whatley (1985); Collins (1997); Hornbeck and Naidu (2014) for further discussion of the economic and political determinants of the Great Migration. For example, Collins (1997) shows how northern industrialists' hiring and recruiting black workers hinged on reduced presence of and access to European immigrant labor due to World War I and immigration controls put in place in the 1920s.

to the South. A major shift in the geography of upward mobility for black Americans appears to have taken place in the decades after 1940.

Figure 1b illustrates the current geographic distribution of black upward mobility in the US. Illustrated in the map is average income rank for black men and women who grew up in low income families in each commuting zone in the 2000s. Several northern locations that exhibited high outcomes for black children in 1940 exhibit some of the worst outcomes for black children today. The fact that the peak of the Great Migration took place in between motivates an empirical investigation of the Migration's role in the decline in black upward mobility in the North.

3 Data

3.1 Upward mobility

Educational upward mobility in 1940 To measure upward mobility in commuting zones prior to the 1940-1970 wave of the Great Migration, I use the complete count 1940 census.[12] Following Card et al. (2018), I define educational upward mobility as the fraction of 14-18 year-old boys and 14-16 year-old girls in each commuting zone with 9 or more years of schooling from households where parents have between 5 and 8 years of schooling, approximately the median for adults in the US at the time.[13] In addition, I use complete count censuses from 1920 and 1930 to develop pre-1940 measures of educational upward mobility, specifically, the school attendance rates of teenagers with low occupation score fathers.

[12] I use the Integrated Public Use Microdata Series ("IPUMS") version for all complete count census data used in this paper.

[13] Here, parent education is defined as the maximum of the mother's or father's education. Card et al. (2018) show that up to age 18 for boys and up to age 16 for girls, there is little selection on observable characteristics into living with one or more parent (Card et al., 2018, p. 14).

Teenagers typically reside in the same households as their parents, obviating the need to match them across censuses to observe parent economic status. At the same time, teenagers are old enough that their educational attainment is likely predictive of their adult educational attainment and future labor market outcomes. Observing outcomes for the near universe of enumerated teenagers reduces the scope for sampling bias in constructing upward mobility measures at fine geographies. Finally, teenager upward mobility can be constructed separately by race without differential selection bias across groups arising from lower match rates for African Americans.[14]

Income upward mobility for 1980s birth cohorts For contemporary measures of upward mobility in commuting zones, I use data made available by Chetty and Hendren (2018b) and Chetty et al. (2018). Based on the universe of federal income tax records from 1996-2012, the data contain measures of income upward mobility by childhood commuting zone for individuals born between 1980 and 1986. Parent and children were linked via dependent claiming. The key measure of upward mobility is estimated mean individual or household income rank, conditional on parent household income rank.[15]

Income for individuals in the sample is income at age 26, during the years 2006-2012, and income rank is rank in the national income distribution for individuals from the same birth cohort. Parent income is measured using returns filed when individuals were between the ages of 14 and 20, and parent income rank is rank in the national parent income distribution by child birth cohort. Separate upward mobility estimates are available for individuals from the 25th and 75th percentile of the parent income distribution. Estimates are also available separately by gender.

[14]Matching methods, which typically rely on first and last name to link individuals across historical censuses, are not well suited to linking African Americans who have fewer unique surnames as a result of slavery.

[15]Household income measures for parents and children are drawn from Adjusted Gross Income on 1040 tax returns, and individual income rank is measured using income reported on W-2 forms, UI and SSDI benefits, or half of household self-employment income where relevant.

How comparable are educational upward mobility in 1940 and income upward mobility in the 2000s? Across US CZs where both measures are available, the two measures are strongly correlated, with a correlation coefficient of 0.49. Additionally, income upward mobility is strongly correlated with high school graduation rates in low income families today, with a correlation coefficient of 0.65.

Childhood exposure effects of commuting zones I use an alternative measure of upward mobility in the 2000s from Chetty and Hendren (2018b): the childhood exposure effects of commuting zones. Starting from the universe of tax filers described above, the authors restricted the sample to individuals whose parents moved once across commuting zones during their childhood. They then compare the outcomes of children exposed for more or less time to a given commuting zone based on children's ages at the time their families moved. Precisely, the data contain estimates of the causal effect of one additional year of childhood in a given commuting zone relative to an average commuting zone, for an arbitrary child. The outcome of interest is adult income rank at age 26. The estimates and assumptions behind them are discussed in greater detail in Section 5.2.

Race-specific measures of upward mobility Race-specific measures of upward mobility come from Chetty et al. (2018). These data are based on the same universe of federal income tax records as the measure described above; however, they cover a slightly different set of birth cohorts: 1978-1983. Individual federal income tax records were linked to the US Census in order to retrieve information on race as well as additional outcomes measured by the Census. The data contain the estimated mean individual or household income rank, conditional on parent household income rank, of black and white men and women at the 25th and the 75th percentiles of the parent income distribution by childhood commuting zone. In this dataset, outcomes are measured in 2015 when individuals were between the ages of 32 and 37.

3.2 Local public finance and neighborhood quality measures

I assembled a new database of local public finance and neighborhood quality measures for commuting zones spanning the years 1920-2015. The database covers statistics on schooling, demographics, racial tension and voting behavior, local government expenditures, incarceration, and crime, among other characteristics. I harmonized data from a variety of sources, including historical reports that I digitized from the US Department of Education's Biennial Statistics of Education reports, the FBI's Uniform Crime Reporting series, and the US Census Bureau's Financial Statistics of States and Local Governments. William Collins and Robert Margo generously shared data on race riots in US cities in the 1960s, based on the work of Gregg Carter (Collins and Margo, 2007; Carter, 1986). Finally, political outcomes come from Clubb et al. "Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972" available from the Inter-University Consortium for Political and Social Research ("ICPSR").

Additionally I used the 1920-1940 complete count censuses to construct additional measures of local government investments and incarceration rates and digitized special reports and tables from the 1940 and 1960 censuses on local county jail populations. I supplemented these data sources with the City and County Data Books series and several county-level tabulated measures from 1970-2010 US Censuses. Finally, I used a pre-release of rich new data on incarceration from the Vera Institute of Justice, locating for each federal and state prisoner the county jail that committed them to federal and state prison. I provide details on the construction of this database, including data sources, and detailed definitions of key measures in Appendix D.

3.3 City demographic data, 1940-1970

I draw on two main sources of data to construct historical black population measures for cities in northern commuting zones in 1940 and 1970: the complete count 1940 US census and the City and County Data Books 1944-1977 series ("CCDB"),[16] which contains information on cities with a population of 25,000 or more. I measure urban black populations in 1940 using the complete count census, as the CCDB only report information on the number of whites and non-whites in cities that year. I collect information on the black population in 1970 from the CCDB.[17] I further restrict the sample to cities that are not missing 1940 population data in the CCDB and to those cities that had at least one recent black southern migrant.[18] The total number of cities that meet these criteria is 294.[19] My final sample of commuting zones is the 130 commuting zones containing these cities.

I define black population change in a commuting zone during the Great Migration as the 1940 to 1970 increases in the urban black population as a share of the initial 1940 urban population:

$$\Delta \text{Black pop}_{CZ}^{1940-1970} = \frac{b_{\text{urban},CZ}^{1970} - b_{\text{urban},CZ}^{1940}}{\text{pop}_{\text{urban},CZ}^{1940}} \tag{1}$$

where $b_{\text{urban},CZ}^{t}$ is the total black population in all sample cities in commuting zone CZ in year t.

[16] Available from the Inter-university Consortium for Political and Social Research ("ICPSR").

[17] I'm unable to locate the following cities from the CCDB in the 1940 census: Boise City, ID; East Providence, RI; Huntington Park, CA; West Haven, CT; and Warwick, RI. I drop these cities from the analysis due to missing data.

[18] Defined as an individual who listed a southern county of residence in 1935, but who resided in a northern city in 1940.

[19] I manually record black population data for two cities in the published 1940 US census: Butte, MT and Amsterdam, NY. Both cities received black southern migrants between 1935 and 1940, but data on their black population in 1970 was not available in the CCDB. Including these two cities brings the total number of commuting zones in the sample to 130 from 128. Finally, the city of New Albany, IN is in the Louisville, KY commuting zone, which is included in the sample. Results are robust to excluding this commuting zone.

Functional form Because the distribution of black population increases is highly right-skewed, I define the quantile function GM_{CZ}, or the percentile of the increase, to be the key independent variable in the empirical analysis.

Figure 2 depicts GM_{CZ} across northern commuting zones during the Great Migration. Plotted on the y-axis is the measure in equation 1, multiplied by 100 so that the units are percentage points. The x-axis measures GM_{CZ}, the quantile function or the percentile of urban black population increase.

The median increase across commuting zones in the sample was 5.5 percentage points. As the figure demonstrates, however, historical black share increases were very unevenly distributed across the North, even among commuting zones in the same region. Take for example, two commuting zones in the midwest—Pittsburgh, PA and Detroit, MI. Both were major manufacturing centers in the 1940s. Pittsburgh's urban black population share increased by 6.6 percentage points (53rd percentile) while Detroit's increased by 29.3 percentage points (97th percentile). Salt Lake City, UT saw almost no increase in its black population while Washington, DC saw an increase of roughly 50 percentage points.

The descriptive relationship between black population change during the Great Migration and average income upward mobility today can be seen in Figure 3. The relationship is strikingly negative and linear.[20] A 1-percentile greater black population increase between 1940 and 1970 is associated with a .07-percentile reduction in adult income rank for individuals with lower income parents. However, as discussed in sections 3.4 and 4, this relationship cannot be interpreted as causal as correlates of black population change may drive this relationship. Moving towards a causal framework requires understanding

[20]The linearity of the relationship suggests that very large increases in the black population share at the tail end of the distribution in Figure 2 had similar effects as smaller increases at the bottom and middle of the distribution. This may in part be due to the positive relationship between levels of the black population share and changes in the black population between 1940 and 1970. Small absolute increases which nevertheless took place in locations with small black population shares may still have prompted large responses. As I discuss in Section 5, my results are robust to flexibly controlling for the level of the black population share in 1940.

the historical forces behind migration during this period.

3.4 Descriptive characteristics of Great Migration CZs

Why did urban black populations in the North increase so dramatically between 1940 and 1970? After a period of reduced mobility during the Great Depression, black outmigration from the South resumed at an accelerated pace after 1940. War-time jobs in the defense industry and in naval shipyards led to substantial black migration to California and other Pacific states for the first time since the Migration began. Migration continued apace to midwestern cities in the 1950s and 1960s, as the booming automobile industry attracted millions more black southerners to the North, particularly to cities like Detroit or Cleveland. Of the six million black migrants who left the South during the Great Migration, four million of them migrated between 1940 and 1970 alone.[21]

As is clear from the discussion above, mid-century economic conditions in northern cities influenced where migrants moved and are thus likely correlated with increases in the black population during this period. These underlying characteristics may also determine the dynamics of upward mobility in destination cities.

Black urban populations increased more in places with higher levels of educational upward mobility (correlation: 0.27). If higher educational upward mobility reflects higher quality educational institutions and this factor persists over time, then OLS measures of the relationship between the Great Migration and upward mobility will be biased towards zero. At the same time, black population increases are positively correlated with the share of the labor force in manufacturing in 1940 (correlation: 0.18). Former manufacturing centers define the Rust Belt today, an area of low upward mobility. Thus, deindustrialization in former manufacturing centers could be a confounding factor in

[21] After 1970, black migration reversed course, with individuals on net relocating to the South, though in much smaller numbers than the migration north.

estimating the effects of the Great Migration. Finally, Great Migration inflows were larger in locations that already had a substantial population of recent black southern migrants (correlation: 0.56),[22] raising questions about the characteristics of destinations that led them to be hubs for black southerners prior to 1940. Given that these destination-level factors may influence both black population increases and future levels of upward mobility, I construct an instrument for the former that is plausibly exogenous with respect to pre-1940 destination characteristics.

4 Empirical Strategy

The intuition behind the empirical strategy is well captured by the migration histories of Detroit and Baltimore. Both were major destinations for black migrants during the Great Migration, and both were major industrial centers in 1940. However, black migrants arriving in these locations in 1940 came from parts of the South that experienced very different patterns of outmigration between 1940 and 1970. Figure 4 depicts variation in black migration for these two cities. Detroit drew the plurality of its migrants from Alabama while Baltimore drew the plurality from Virginia. Migrants from Alabama tended to come from counties specialized in cotton production, and negative shocks to cotton spurred outmigration from these areas. Virginia, by contrast, was a major recipient of war production spending during World War II. War production jobs attracted black workers and consequently lowered outmigration rates.

The empirical strategy generalizes from the example above and builds on

[22] Data on recent black southern migrants come from the 1940 complete count census. The 1940 census was the first census to systematically record internal migration. Enumerators asked individuals about their prior residence (city, county, and state) in 1935. I define recent southern black migrants as those who reported a southern county of residence in 1935 and lived in an northern city as of 1940. Here, southern is defined as being from the following states: Alabama, Arkansas, Florida, Georgia, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia.

a standard shift-share approach used to estimate local labor market impacts of migration (Altonji and Card, 1991). The technique was first adapted to the Great Migration context by Boustan (2010). Black southern migrants tended to move where previous migrants from their communities had settled, thus generating correlated origin-destination flows similar to those observed in the international migration context. Shocks to migrants' origin locations ("push factors") are plausibly orthogonal to shocks to the destinations ("pull factors") that could also influence the location choices of future migrants. Interacting exogenous swings in migration at the origin level with historical migration patterns in the destinations yields a potential instrument for black population changes in the North.

The validity of this type of instrument has been extensively explored in recent papers (Goldsmith-Pinkham et al., 2018; Adao et al., 2019; Borusyak et al., 2019). After describing how I construct the instrument and how it differs from previous papers studying the Great Migration, I discuss the relevant identification and inference concerns raised by the shift-share literature, and how I address them, in Section 5.4.

To construct my instrument for black population change in northern cities, I interact variation in the cities' pre-1940 migrant composition with variation in outmigration from southern counties driven by push factors alone. These push factors include defense facility spending in southern counties during World War II and shocks to cotton and other economic sectors in the South, e.g., tobacco and mining. More precisely, I replace the numerator in Equation 1 with the predicted, as opposed to actual, increase in the black population:

$$\text{Predicted black pop}_{CZ}^{1940-1970} = \frac{\hat{\Delta} b_{\text{urban},CZ}^{1940-1970}}{\text{pop}_{\text{urban},CZ}^{1940}} \tag{2}$$

where $\hat{\Delta} b_{\text{urban},CZ}^{1940-1970}$ denotes the predicted increase, which I define as follows:

$$\hat{\Delta} b_{\text{urban},CZ}^{1940-1970} = \sum_{j \in S} \sum_{c \in CZ} \omega_{jc}^{1935-1940} \times \hat{m}_{j}^{1940-1970}. \tag{3}$$

The term $\hat{m}_j$ is predicted black migration from southern county j over the decades 1940 to 1970; ω_{jc} is the share of recently migrated pre-1940 black southern migrants from county j living in city c in 1940. The term $\hat{m}_j^{1940-1970}$ consists of the sum of fitted values of decadal predictions of southern county net migration (from 1940-1950, 1950-1960, and 1960-1970) using lagged southern economic predictors of migration:

$$\hat{m}_j^{1940-1970} = \sum_{t=1950}^{1970} \widehat{\text{mig rate}}_{jt} \cdot \text{black pop}_{jt}$$

where fitted values, $\widehat{\text{mig rate}}_{jt} = \text{mig rate}_{jt} - \varepsilon_{jt}$, come from the following prediction of net-migration rates:

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

Appendix B describes the construction of ω_{jc} and $\hat{m}_j^{1940-1970}$, and the procedure for choosing predictors Z'_{jt-10} in detail.

Functional form My instrument for the percentile of black population increase during the Great Migration, GM_{CZ}, is $\hat{GM}_{CZ}$, the percentile of the predicted black population increase defined above. I use the percentile of the predicted increase as the key independent variable because the distribution of predicted black population increases mirrors that of the actual increases—both are heavily right skewed. In reporting the effects of percentile changes in the black population, I follow Sequeira et al. (2019) who report the impact of a zero to 50th percentile increase in European immigration during the Age of Mass Migration on the long-run economic development of US counties.

My empirical strategy builds off the identification strategy developed by Boustan (2010) and used in subsequent papers to estimate impacts of the Great Migration on destination cities (Tabellini, 2018; Fouka et al., 2018). I introduce two innovations to this empirical strategy. First, I use the complete count 1940 census, which contains microdata on the universe of recent black

southern migrants into northern cities, including their county of residence in 1935. Using county of residence in 1935 and city of residence in 1940, I construct a matrix of southern-county-to-northern-city linkages containing the share of each southern county's outmigrants who settled in each northern city. This detailed linkage contrasts with the state-level linkage used in the prior literature.[23] Using the complete count census data, I am able to leverage shocks to over 1200 origin counties as opposed to just 14 southern states. As I explain in Section 5.4, a large number of shocks is important for the validity of the empirical strategy when identification relies on shocks to origin locations being orthogonal to shocks to the destinations (Goldsmith-Pinkham et al., 2018; Adao et al., 2019; Borusyak et al., 2019).

The second innovation is that I use machine learning to improve the prediction of net migration from southern counties. The motivation for this approach is that the set of potential predictors from southern county variables is large. Given that the first stage prediction of an endogenous variable by an instrument can be viewed as a pure prediction problem (Belloni et al., 2011), I select among the predictors for migration used by Boustan (2010) using a Post-LASSO estimation procedure. In this procedure, for each decade of migration between 1940 and 1970, I use LASSO to select predictors among county characteristics in the previous decade with a penalty on the absolute number of predictors, where the tuning parameter has been chosen by 5-fold cross-validation. I then use the variables chosen by this procedure to estimate their relationship with county net-migration rates using OLS.[24]

To ensure that I am leveraging variation from specific southern-county shocks, I control for the total share of the 1940 urban population made up

[23]Boustan (2010) uses tabulated census reports that recorded 1935 state of residence to construct southern-state-to-northern-city weights to assign migrants. The complete count 1940 census was declassified in 2012, so the empirical strategy used in this paper was not feasible at the time. One exception to the state-of-origin variation used in analyzing the effects of the Great Migration is the work of Black et al. (2015); Stuart and Taylor (2017); Stuart et al. (2018), which uses the Duke SSA/Medicare dataset, which no longer accepts applications from new researchers.

[24]Southern county-net migration rates are taken from Boustan (2016).

of recent black migrants from any southern county. My preferred specification also includes the following baseline 1940 characteristics for robustness: educational upward mobility and the share of the labor force in manufacturing. These regressions can be interpreted as estimating the effect of historical change in the black population on the change in upward mobility in the sample commuting zones, where I allow for dynamics in upward mobility. If upward mobility changed in the treated commuting zones for reasons other than the Great Migration, forcing the coefficient on historical upward mobility to be 1 may be a mis-specification of the true relationship between the Migration and upward mobility.[25] Finally, I include census region fixed effects. The inclusion of these controls does not significantly alter the point estimates, and I report key results with and without this baseline set of controls.[26]

Estimating equation I estimate the relationship between the Great Migration and upward mobility using the following empirical framework:

$$\bar{y}_{p,CZ} = \alpha + \beta GM_{CZ} + \mathbb{X}'_{CZ}\Gamma + \varepsilon_{CZ} \tag{4}$$

$$\text{First Stage: } GM_{CZ} = \gamma + \delta \hat{GM}_{CZ} + \mathbb{X}'_{CZ}\mu + \epsilon_{CZ} \tag{5}$$

[25] My main results are robust to alternative specifications where I estimate the impact of the Great Migration directly on the change in upward mobility for black Americans between 1940 and 2015. See a discussion of this result in Section 5.4.

[26] Including census region fixed effects leads to more precise and larger IV estimates of the impact of the Great Migration on upward mobility. However, the point estimate without controls is not statistically different from the point estimate with census division fixed effects or with the full set of baseline controls. See columns 1, 2, and 3 of Table 6. A potential reason for the difference in the point estimates between columns 1 and 2 is that the instrument for black population increases leverages linkages between southern origin locations and northern destinations made between 1935 and 1940. Relatively few black southern migrants had settled in the West by 1940, thus, relative to the endogenous variable, the instrument reallocates migrants towards the Midwest as opposed to the West. It would be ideal to use the 1950 census to establish the migrant network for the West as many African Americans moved west for the first time during World War II. The required micro data from the 1950 census will be available in 2022. Given these data constraints, inclusion of census division fixed effects reduces the noise introduced by the pre-1940 migrant networks.

In equation 4, the coefficient β represents the OLS estimate of the effect of GM_{CZ}, the percentile of a commuting zone's 1940-1970 black population increase, on $\bar{y}_{p,CZ}$, the average adult income rank of children with parents at income rank p, conditional on baseline characteristics and census division fixed effects represented by the control vector X_{CZ}. Equation 5 estimates the first stage relationship between the percentile of predicted black population change $\hat{GM}_{CZ}$ and the percentile of actual black population change, GM_{CZ}. The reduced form effect of my instrument for the Great Migration on upward mobility can be written as follows:

$$\bar{y}_{p,CZ} = \tilde{\alpha} + \tilde{\beta}\hat{GM}_{CZ} + \mathbb{X}'_{CZ}\tilde{\Gamma} + \tilde{\varepsilon}_{CZ} \tag{6}$$

where $\tilde{\beta}$ represents the reduced form impact of the percentile of predicted black population change on upward mobility. For all main results, I report the estimated OLS (β), reduced form ($\tilde{\beta}$), and two-stage least squares ($\frac{\tilde{\beta}}{\delta}$) coefficients.

Identifying assumption In order for the approach above to identify the causal impact of the Great Migration, conditional on the specified baseline 1940 characteristics, my instrument for black population increases must be orthogonal to omitted characteristics that are correlated with changes in upward mobility after 1940.

$$\mathbb{E}[\hat{GM}_{CZ} \cdot \tilde{\varepsilon}_{CZ} | \mathbb{X}_{CZ}] = 0 \tag{7}$$

Although this identifying assumption cannot be directly tested, I show that using this empirical strategy, the Great Migration has no effect on pre-1940 measures of educational upward mobility, defined as the school attendance rate of teens with low occupation score fathers, or on median adult educational attainment in 1940.[27]

[27]Defined as the population-weighted average median educational attainment of adults by county.

Table 1 reports the results from these placebo checks. The results show that conditional on baseline controls, the instrument for the Great Migration is uncorrelated with educational upward mobility prior to 1940. The coefficients on $\hat{GM}$ are very small in magnitude, statistically insignificant, and similar across the decades 1920 to 1940. The Migration also does not predict any differences in adult median educational attainment in 1940.

Validity of shift-share instrument Recent analyses of shift-share instruments commonly used to estimate the causal impact of immigration or of demand shocks to labor markets have generated new intuitions on exogeneity and inference of shift share instruments (Goldsmith-Pinkham et al., 2018; Adao et al., 2019; Borusyak et al., 2019).

Borusyak et al. (2019) show that in the case where shares are endogenous, shifters can provide exogeneity provided that shocks to industry, or in my setting, southern counties, are not correlated with shocks to northern destinations. To demonstrate that the results are not simply generated by specific origin-level shocks correlated with shocks to destinations, I construct alternative instruments and conduct an over-identification test.

In addition to the baseline instrument, I construct a second instrument using the southern county outmigration rates that are first residualized on state fixed effects. This version of the instrument accounts for correlated shocks to southern states and northern destinations (e.g., Virginia and Baltimore, which both have a substantial defense industry).[28] A third version of the instrument uses variation in state of birth across the southern born black population in northern cities in 1940 interacted with variation in state-level net migration. This instrument leverages different variation in historical migration patterns thus varying northern cities' exposure to southern origin shocks. Results using each of these versions of the instrument are extremely similar—Appendix Figure C9 reports these results. A formal over-identification test fails to re-

[28]It's worth noting that shocks must be negatively correlated to generate both outmigration from the southern location and endogenous in-migration into the northern destination.

ject the null that the estimated effects on upward mobility are statistically indistinguishable from each other.[29]

Adao et al. (2019) note that in the case of shift-share instruments, standard inference procedures, just as geographic clustering, may result in standard errors that are too small. This will be the case if, in this setting a set of counties bears similar importance across multiple destination locations, generating correlation at the origin county level across destinations. I run a placebo analysis, interacting recent black southern migrant location choices with a normally distributed random variable with mean 0 and variance 5.[30] I iterate this procedure 1,000 times and document the fraction of times results show significant effects at the 5% and 1% level. These results are reported in Appendix Figure C10. The coefficient on the Great Migration was significant in both the positive or negative directions, 16.1% of the time at the 5% level and 6.1% of the time at the 1% level. Given the precision of the findings, the results from this placebo analysis suggest that while the standard errors likely warrant adjusting, the impact of the Great Migration on upward mobility is unlikely to be driven by noise and would remain highly significant.

First-stage results Figure 5 shows a binned scatterplot of the relationship between GM, the percentile of actual black population increase, and $G\hat{M}$, the percentile of predicted black population increase, where both measures have been residualized on census division fixed effects and the set of 1940 baseline controls: educational upward mobility, the share of the labor force in manufacturing, and the share of the 1940 urban population made up of recent southern black migrants from any southern county. The y-axis plots mean percentile of black population change within each 5-percentile bin of predicted black population change. The slope of the regression line is equivalent to the coefficient $\hat{\delta}$ from equation 5. A one-percentile larger predicted black

[29] Hansen J statistic p-value of .20.

[30] For simplicity, I use the mean and variance used by Adao et al. (2019). Because I use a rank transformation of black migration, the ranking of migration inflows is invariant to the specific variance chosen.

population increase is associated with a 0.3 percentile greater actual black population increase over the time period. The F-statistic on the first stage is 15.3.

5 Results on upward mobility

The Great Migration represented a large-scale movement to opportunity for black Americans. In the North, jobs were far better paying, black children could attend high school, and racial equality was taken for granted in many facets of northern life.[31] From the vantage point of 1940, there was every reason to believe future generations of black children would continue to reap the benefits of their parents and grandparents having migrated. The results from the empirical analysis in this paper suggests otherwise.

Many of the locations where black migrants moved in large numbers are now among the worst places to grow up, in stark contrast with geographic patterns in upward mobility in the northern US in 1940. I show that this transformation appears causally related to the Migration. Using exogenous variation in where the black population increased the most during the period of the Migration, I find that mid-century shocks to the racial composition of northern cities lowered the average outcomes of children growing up in the 1990s and 2000s. The driver of this effect appears to be changes in location characteristics, not shifts in the composition of families living in Great Migration cities, which could mechanically give rise to lower average upward mobility. Analysis of which groups of children were affected by these changes suggest that black men were the most negatively affected sub-group. This section describes the key results on upward mobility in detail, before exploring local mechanisms in Section 6.

To focus on the more plausibly causal estimates of the impact of the Great

[31] See Wilkerson (2011) for accounts and experiences of individual migrants arriving in and navigating new lives in the North.

Migration, I primarily discuss reduced form and IV results in what follows. For all main results, however, I report first stage, OLS, reduced form, and two-stage least squares ("2SLS") results and briefly discuss differences between OLS and 2SLS estimates.

5.1 Impact on average upward mobility

I first estimate the impact of predicted black population increases during the Great Migration, or $\hat{GM}$, on average upward mobility at the commuting zone level (the model in equation 6). The outcome variable is mean expected household income rank of individuals from the 1980-1986 birth cohorts with parents at income rank p by their childhood commuting zone, where individuals' income is measured at age 26. Figure 6 shows a binned scatterplot of the relationship between $\hat{GM}$ and upward mobility for individuals with low income parents (at the 25th percentile of the parent income distribution). Both the outcome and $\hat{GM}$ have been residualized on the baseline set of controls discussed in Section 4. Each dot represents average outcomes across commuting zones within 5-percentile bins of the shock. The figure shows a stark negative relationship between historical black migrant inflows and average outcomes for individuals from low income families in the destination CZs today.

Table 2 reports 2SLS estimates of the relationship. A 1-percentile increase in the historical black population lowered household income rank by -0.125 percentile points (s.e. = 0.0328). OLS estimates are reported in Table 2 as well.[32] The 2SLS coefficients are larger in magnitude than the OLS. One potential explanation for this is that omitted characteristics are positively correlated with both black population change and upward mobility. For example, the black population grew more in places with higher levels of median educa-

[32] Appendix Table C3 reports the results for individuals with high income parents. I find more modest effects of $\hat{GM}$ on the outcomes of individuals with high income parents (at the 75th percentile of the parent income distribution). For this group, a 1-percentile increase in the historical black population lowered household income rank by -0.0541 percentile points (s.e. = 0.0232).

tional attainment in 1940. To the extent that higher education levels reflect better school quality, which may persist over time, OLS estimates of the relationship between the Great Migration and upward mobility today would be biased towards zero.

How should one interpret the negative effect of the Migration on average upward mobility? In a simple framework where the adult outcomes of children conditional on parent economic status are a function of childhood location and an unobservable family component, the Migration may influence mean outcomes either by changing aspects of the location or changing the characteristics of the average child. More formally, let the outcome for a child i with parent household income rank p living in CZ be the sum of a pure location component and an idiosyncratic family component:

$$y_{ip,CZ} = \mu_{p,CZ} + \theta_{ip,CZ} \tag{8}$$

Recall, I observe mean outcomes in a location at a given parent rank p:

$$\bar{y}_{p,CZ} = \mu_{p,CZ} + \bar{\theta}_{p,CZ} \tag{9}$$

Because any migration event changes the composition of families in a destination location, there is a potential mechanical effect of the Great Migration on $\bar{\theta}_{p,CZ}$.[33] Alternatively, the Migration may affect behavior of incumbents within a commuting zone, for example, altering the equilibrium bundle of public goods voted on by local residents or their residential choices within a commuting zone, giving rise to various forms of segregation (Alesina et al., 2004). These choices may in turn affect the outcomes of children growing up in these locations in the future, independent of their familes' characteristics, $\theta_{ip,CZ}$. In that case, the Migration would affect average outcomes through the channel of $\mu_{p,CZ}$.

One example of θ_i includes the race of the child, which if unobserved, could

[33]Further, positively selected families may move away locations with high levels of in-migration, further potentially affecting $\bar{\theta}_{p,CZ}$.

explain a substantial portion of the Migration's estimated impact on $\bar{y}_{p,CZ}$. Several studies have found persistent differences in intergenerational mobility by race, and these gaps persist even among those observed to be growing up in the same census tract.[34] Areas with a higher black share of the population likely have lower average upward mobility.

Another example of θ_i would be a family's propensity to invest in the human capital of their children. Even after conditioning on parent income, if families tend to value or invest in human capital differently, this may lead to divergent adult outcomes for children from these families, even after conditioning on parent income rank.

5.2 Impact on childhood exposure effects

To address sources of selection θ_i that may be driving the relationship between the Migration and average upward mobility in Figure 6, I turn to an alternative metric of upward mobility in locations that attempts to isolate the causal effect of childhood location.

I take these estimates from Chetty and Hendren (2018b). The authors estimate the causal effect of growing up in every commuting zone in the United States using federal income tax data on families that moved across commuting zones and exploiting variation in children's ages at the time their families moved.[35] Under the assumption that the age of a child at the time a family moved is orthogonal to unobserved family characteristics θ_i, estimating the effect of one additional year of childhood exposure to a location and multiplying this effect by number of years of childhood provides a direct estimate of $\mu_{p,CZ}$ in the model in equation 8.[36]

[34] See, for example, Mazumder (2014); Davis and Mazumder (2018); Chetty et al. (2018).

[35] Parents and children are assigned commuting zones based on the ZIP Code information available on their tax returns.

[36] See Chetty and Hendren (2018a) for several checks of this identifying assumptions including instrumenting for moves with displacement shocks to families and the inclusion of family fixed effects.

The advantage to using these measures is that they provide metrics of upward mobility that isolate the effect of childhood location. Thus, any impact of the Great Migration on this alternative measure of upward mobility can be interpreted as follows: a child randomly assigned to spend an additional year in CZ A that experienced a large shock versus CZ B that experienced a small shock has greater or lower adult income rank. One downside to these measures is that they are not available separately by race. This means I identify impacts of the Migration on childhood exposure up to an average effect across black and white children. Data limitations prevent me from exploring potentially heterogeneous impacts of the Migration on $\mu_{pr,CZ}$, or location effects for black versus white children. In Section 5.3, I explore this heterogeneity using race-specific measures of average upward mobility in commuting zones and discuss the plausibility of the findings being driven by race-specific selection stories ($\bar{\theta}_{pr,CZ}$).

Figure 7 shows a binned scatterplot of the impact of the Great Migration on CZ childhood exposure effects for individuals with parents from the 25th percentile of the parent income distribution. Both the outcome and $\hat{GM}$ have been residualized on the baseline set of controls discussed in Section 4. Each dot represents average outcomes across commuting zones within 5-percentile bins of the shock. The figure shows a strong negative relationship between historical black migrant inflows and the effects of childhood exposure to destination CZs.

Table 3 reports OLS and 2SLS estimates of the relationship. The 2SLS estimates cfan be interpreted as follows: a 1-percentile larger increase in the historical black population lowers household income rank by -0.0087 percentile points (s.e.=.0028).[37] The first stage is 0.27, so the 2SLS coefficients are three times larger in magnitude than the OLS, indicating again that there may be omitted characteristics positively correlated with both childhood exposure effects and black population change that then bias the OLS estimates towards

[37]For individuals from high income families, I find effects of about half the size—consistent with the results on average upward mobility. See Appendix Table C4 for these results.

zero.

5.2.1 Interpretation of results on childhood exposure effects

The results thus far support the hypothesis that one way responses to the Great Migration lowered upward mobility was through a changing environment for families. These estimated impacts on childhood environment can be combined with the first set of results on upward mobility to quantify the impact of the Migration through $\mu_{p,CZ}$ versus $\bar{\theta}_{p,CZ}$. I do this by scaling the 2SLS estimated effect on one year of childhood exposure to represent full childhood exposure to a Great Migration destination and comparing the resulting scaled estimate with the 2SLS estimated impact on observed upward mobility.

Scaling the estimated impact on childhood exposure effects requires making assumptions about the relationship between the average effect of a year of childhood exposure to a location and the age at which the child is exposed to the location. In other words, if the effect of childhood location remains constant over years of childhood, then multiplying the impact of one year by total number of years exposed yields the effect of full childhood exposure.

Chetty and Hendren (2018a) and Chetty and Hendren (2018b) assume constant location effects over each year of childhood and multiply exposure effects by 20 to approximate full childhood exposure. In more recent work, however, Chetty et al. (2018) using data on earlier cohorts of individuals find that the relationship between age at move and predicted income rank in a destination exhibits a kink around age 13, with pre-teen years of childhood exposure having a smaller effect on adult outcomes than teen and post-teen years (see Appendix Figure C3).

The table below decomposes the impact of the Great Migration on upward mobility through the channels of location versus selection using each assumption in turn. Estimates have been scaled to represent the effects of a 30-percentile increase, or approximately 1 s.d., in the historical black popu-

lation. The first row reports results from assuming constant effects over 20 years of childhood exposure leading to a multiplier of 20, and the second row assumes muted effects in the pre-teen years, leading to a multiplier of 15.53. Appendix Section C.3 provides the exact numbers used to calibrate this scaling exercise.

Column 1 reports the impact of the Great Migration on location effects, scaled to represent full childhood exposure to those locations. Column 2 reports the impact of the Great Migration on average upward mobility. The latter estimate combines the Migration's effects through the selection channel as well as the location channel. The ratio of Column 1 estimates to Column 2 estimates gives a sense of what share of the impact of the Migration is driven by location versus selection effects.

The first row suggests that the channel of childhood location explains 140% ($\frac{5}{3} \times 100$) of the impact of the Migration on upward mobility, or that selection effects are in fact positive. The second row makes this comparison using the assumption of more muted impacts of early years of childhood exposure. In this case, I find that the location channels explain 108% of the Migration's effect on upward mobility.

Contribution of location versus selection in Great Migration effects

	CZ Childhood Exposure Effects	Average Upward Mobility
20 years	-5.1	-3.6
15.53 years	-3.9	-3.6

All 2SLS specifications include region fixed effects as well as baseline controls from 1940, including total 1935-1940 black southern migrant share of the population, share of the labor force in manufacturing, and educational upward mobility.

These results suggest that changes in childhood environment are the main mechanism for the impact of the Great Migration on upward mobility. If

the empirical strategy is valid, the estimates reported above reflect the causal effect of black population changes during the Great Migration on childhood environment.

5.3 Heterogeneity by race and gender

In this section, I explore whether different groups of children were affected more or less by the Migration. I do so by estimating the impact of $\hat{GM}$ on race-specific average upward mobility in CZs for black and white men and women from the 1978-1983 birth cohorts. The outcome variable is mean conditional income rank in 2015 by childhood commuting zone. OLS, reduced form, and 2SLS results are reported in Table 4 for black men and women and Table 5 for white men and women.

Figure C2 summarizes these regressions in a plot of the coefficients on percentile of predicted black population change in regressions of each subgroup's average upward mobility on $\hat{GM}$. Here the shock has been scaled to be in 1 standard deviation units. The incidence of the Migration's negative effects on upward mobility fall on black men. A 1-standard-deviation increase in the intensity of a CZ's Great Migration shock lowers the individual income rank of black men by around 2 percentile points, with slightly larger effects on men with higher income parents. By contrast, I find no effects of the Migration on the individual earnings of white men from any parent income group.[38]

The point estimates for the Great Migration's impact on black women's individual earnings are positive and insignificant for black women from low income families and positive and significant at the 10% level for black women from higher income families. These positive effects may represent an income effect. Interracial marriage rates are very low, so black women who marry men likely form households with black men. Given that black men's income is lower in Great Migration destinations, women may increase their labor supply

[38]The term "white" refers to non-Hispanic white population while Hispanic includes those who report their ethnicity as "Hispanic," regardless of race (Chetty et al., 2018).

to compensate for missing men's income. To test this hypothesis, I estimate the effect of the Migration on black women's household income rank as opposed to their individual income rank. The Migration has a negative and insignificant effect on black women's household income rank, consistent with black women increasing their labor supply in locations with a low marriage rate or missing income of black men in shared households. I report these results in Table 4.

The results above do not rule out within-race selection (θ_{ir}) of families into Great Migration locations today as a potential mechanism for the effect of the Migration on average upward mobility. Certain family characteristics, especially family structure or presence of both parents in a household, have been shown to have much stronger effects on boys versus girls (Bertrand and Pan, 2013). Boys' outcomes are also more elastic than girls' to other inputs as well, for example, school quality (Figlio et al., 2016). If black families that invest less in their children's human capital are more likely to live in Great Migration destinations today, then boys from these families may be more affected as adults than girls.

Implications for the racial gap The fact that black households have reduced conditional income as a result of the Migration but white households are unaffected has implications for the racial gap in income upward mobility in the US. In this section I conduct a counterfactual exercise to quantify the contribution of the Great Migration to the gap in upward mobility between black and white individuals with low income, high income, and median income parents.

The counterfactual seeks to address the following question: what would the racial gap in upward mobility in North be without the changes induced by Great Migration? I define the counterfactual as one in which black families grow up in locations that receive the lowest percentile of shock.[39] I then

[39]Alternatively, I can compute counterfactual upward mobility for both black and white families and take the difference. The point estimate for the Migration's effect on white men is negative but close to zero, and this approach ignores the fact that the effect is

compute the average racial gap under this counterfactual to the observed racial gap across the sample commuting zones.

The results are reported in the table below.

Contribution of the Migration to the northern racial upward mobility gap

	Parent Income		
	25th pctile	50th pctile	75th pctile
Observed	12.03	13.45	15.30
CF w/o GM (se)	9.1 (.13)	9.83 (.14)	11.01 (.2)
Pct Change	-24%	-27%	-28%

The first row reports the average observed racial gap, ranging from 12.03 income rank percentiles for individuals with parents at the 25th percentile to 15.30 income rank percentiles for men with parents at the 75th percentile. The second row reports the counterfactual average gap where northern black families experience the lowest percentile of Great Migration shock. Under this counterfactual, the average racial gap across northern commuting zones ranges from 9.1 percentiles (s.e. = .13 percentiles) for individuals with low income parents to 11.01 percentiles (s.e. = .20 percentiles) for individuals with high income parents.

These estimates suggests the Migration increased the racial gap by 24% for low income families, 27% for median income families, and 28% for high income families. These substantial effects on upward mobility and the racial gap warrant an exploration of the local mechanisms through which the Migration affected outcomes. Before assessing these potential mechanisms in Section 6, I first discuss several alternative theories for the findings, namely, deindustrialization and other historical migrations affecting northern commuting zones.

statistically insignificant. Taking this effect on white families as the true effect, the gap in upward mobility for individuals growing up in median income families is 20% rather than 27%.

5.4 Alternative explanations

Deindustrialization Many black southerners were drawn north by manufacturing sector jobs in cities like Gary, Detroit, and Baltimore. These once booming industrial centers subsequently underwent devastating job loss, with the US losing 2 million manufacturing jobs between 1970s and 2000 (Charles et al., 2019). My empirical strategy leverages variation in black population changes stemming from southern factors, rather than economic conditions in the North. Nonetheless, a potential concern is that my measure of black in-migration is correlated with specialization in manufacturing and hence ensuing job loss.

To what extent does deindustrialization account for the reduction of opportunity for black families in Great Migration CZs? It is worth noting that the correlation between black in-migration and the share of the labor force in manufacturing in 1940 is 0.2. Furthermore, in all specifications, I control for the share of the labor force in manufacturing in 1940, which largely accounts for variation in manufacturing shares in subsequent decades.[40] In a final check, I instrument for employment changes in the destination commuting zones using a Bartik demand shock, interacting industry shares with national, leave-one-out changes in manufacturing employment between 1940 and 1970. Including this demand shock as a control in my main specification alters neither the magnitude nor the precision of the estimated effect of the Great Migration on upward mobility for today's cohorts.

Most significantly, I find no effect of the Great Migration on the adult outcomes of white men raised in low income families, a demographic group that would likely be affected if the findings were driven by deindustrialization rather than changes in racial composition. Lastly, black men from higher income families in Great Migration CZs fare worse than those from locations with less historical black in-migration. These empirical patterns suggest that

[40]The correlation between 1950 share of the labor force in manufacturing and the baseline period share is 0.96. By 1970, this drops only slightly, to 0.84.

manufacturing job loss alone cannot explain reduced outcomes in former Great Migration destinations.

Rather, my findings are consistent with stark segregation and urban decline leading to restructuring of economic opportunity in destination CZs. Black families left in the urban core potentially faced disproportionate job loss compared to white families relocating en masse to growing suburban zones. I discuss changes to the quality of the urban environment in detail in Section 6.

European immigrant labor Prior to their reliance on southern black labor, major industrial centers in the North employed European immigrants. Sequeira et al. (2019) demonstrate that counties that received larger influxes of European immigrants subsequently had higher growth and less poverty. It's possible that the effect of the Great Migration confounds the loss of this labor supply during World War I and after the Immigrant Exclusion Act of 1924, which induced these areas to begin hiring black workers from the South. I do not find evidence consistent with historical European immigrant shares driving my findings: controlling for lagged European immigrant shares prior to 1940 does not alter the precision or magnitude of the impact of the Great Migration on upward mobility.

White southern migration A further consideration is the effect of changing the southern born share of the population. Southerners may have different policy preferences than northern incumbents. The increase in the southern born share of the population is therefore a confounding factor in the Great Migration's estimated impact on upward mobility. I explore this alternative explanation by leveraging the fact that white southerners also migrated to northern cities during this period. In a placebo exercise, I show that instrumenting for the change in the white southern population during this period has no effect on black men's upward mobility. White southern in-migration also

has no impact on childhood exposure effects.[41] Appendix Figure C7 shows the relationship between white southern in-migration and black men's outcomes in binned scatterplots. The relationship is insignificant and the coefficient has the opposite sign as the effect of black population increases.

Other fixed characteristics of Great Migration CZs Finally, I examine the extent to which the findings are driven by other potential fixed characteristics of Great Migration CZs. If locations with high black population shares are fundamentally different from those with lower black population levels, for example, this fixed characteristic could be a confounding factor for my findings.[42] Results are reported in tables 6 and 7. In the case of childhood exposure effects, the point estimates are similar in magnitude and precision across these specifications. The coefficient attenuates slightly for the impacts on black men. However, the upward mobility estimates for black men are less precise in places with very small black populations, which may lead to attenuation in the estimated impact of the Migration due to down-weighting locations with well measured outcomes for black men. In the case of childhood exposure effects, which rely on a different source of variation (children's ages at the time their families relocated across commuting zones) results are highly robust to including flexible controls for the black population share in 1940.

Finally, I show consistent results using a first-differenced specification where I take the difference between standardized educational upward mobility for black boys in 1940 and standardized income upward mobility for black men in the 2000s. These results are reported in Appendix Figure C8. This specification estimates the impact of the Great Migration on the change in black men's upward mobility variation. I find a strong negative relationship which

[41] White southern migration appears associated with lower outcomes for white men and women from lower income parents. The lack of an effect on childhood exposure effects suggests that the channel is the composition of the average white child as opposed to changes in local public goods or neighborhood quality in response to historical in-migration of white southerners.

[42] Historical black inflows were higher in locations with a higher initial black population in 1940.

suggests that the Migration's impacts are not driven by fixed characteristics of the CZs in the pre-1940 period.

Together, the evidence presented thus far supports the interpretation that racial composition shocks during the Great Migration lowered upward mobility in destination commuting zones through a deterioration of the northern urban environment for families. Further, these changes appear to have been particularly detrimental for black men growing up in former destination locations.

6 Evidence on local mechanisms

Why did the northern United States cease to be a land of opportunity for black families in the wake of the Great Migration? The historical and sociological literature on urban crisis point to the role of white flight combined with declining economic opportunity in the urban core. Wilson (1990) highlighted the importance of economic factors—reduced prospects for black men in the labor market and subsequently in the marriage market, thus contributing to increased crime and the rise of single households headed by women. Sugrue (1996) also points to the confluence of isolation of poor black households in urban areas and a long trend of manufacturing jobs relocating out of central cities into suburban and rural locations.[43]

Guided by this historical and sociological literature, I focus my analysis on rising segregation, racial tensions, urban decline, and the policy choices of local governments as plausible mechanisms. I use data I assembled on urban northern commuting zones, which spans the period 1920 to 2015. I detail the

[43]Contemporaneous government documents also attest to the extreme inequality in US cities in the 1960s. The 1968 "Report of the National Advisory Commission on Civil Disorders," popularly known as the Kerner Commission Report, analyzed the riots occurring in major cities at the time and concluded that they were the culmination of decades of segregation, discrimination, and racial inequality. Despite the fact that the black population made up a majority of the urban population in several northern cities, black residents largely lived in cities with all-white governments and interacted with all-white police forces.

construction of this database in Appendix D. I estimate the following:

$$M_{CZ}^{t} = \eta + \mu G\hat{M}_{CZ} + \mathbb{X}'_{CZ}\phi + \nu_{CZ} \tag{10}$$

where t refers to the period the mechanism is measured, and M refers to the mechanism of interest. I standardize all mechanism variables and scale the Migration shock $G\hat{M}_{CZ}$ so that the units are one standard deviation (approximately 30 percentiles of predicted black population increase). I estimate the effect of the Great Migration on average pre-period mechanisms (1920-1940) to check for pre-trends and average post-period (1970-2015) mechanisms to assess the long-run impacts of the Migration.

Figure 9 summarizes the results from this analysis. As can be seen in panel (b), Great Migration commuting zones remain more segregated than locations less affected by the Migration, as indicated by opposite effects on white and black private school enrollment (with a highly statistically significant effect on the gap) and residential racial segregation. Analysis of various measures of neighborhood quality suggest that urban decline followed the Great Migration. Destinations exhibit higher murder rates, are more segregated by income, and exhibit greater economic sprawl in the post-1970 period. Local governments increased investment in police and incarcerate at higher rates. By contrast, I see no systematic re-allocation of spending towards or away from other types of spending over which local governments exercise discretion.[44] Appendix Figure E1 documents the lack of a pre-trend across a large number

[44]Although I find no impact of the Migration on educational expenditures per capita or on the share of total spending by local governments devoted to education, these aggregate measures may mask differences across individual school districts within commuting zones.[45] I find suggestive evidence of a higher fraction of white children and a lower fraction of black children enrolled in private schools in Great Migration CZs. Private school enrollment rates tend to be higher in urban areas, so these results are suggestive of lower school quality in urban public school districts. If school spending decreased in urban school districts, which serve more minority student populations, and simultaneously increased in suburban school districts, these two effects could cancel each other out at the commuting zone level. Further analysis utilizing individual school district data is needed to test whether this reallocation within commuting zones explains the null results on education. Appendix Table D1 provides a breakdown of local, state, and federal contributions to different public spending categories.

of these mechanisms, suggesting that the Migration shifted the nature of the urban environment in key ways. One exception is murder rates, which are positively associated with the Migration in the pre-period. All main results are robust to controlling for pre-period murder rates.[46]

Finally, I assess the extent to which sorting within commuting zones contributes to disparate outcomes for black and white individuals growing up in destination locations. I compute the census-tract-level racial gap in income for black and white men from across the parent income distribution and estimate the impact of the Migration on the population-weighted average census-tract-level racial gap. Appendix Figure E12 reports these results. Segregation is not the only mechanism through which the Migration worsened black outcomes in destination CZs. The within-census-tract racial gap in upward mobility is larger in Great Migration destinations. These results suggest that black boys growing up even in predominantly white neighborhoods face a different effective environment than their white counterparts. Policy responses such as investments in the criminal justice system may have a disproportionate impact on black male youth, no matter the neighborhood in which they reside.

I am limited in my ability to identify the relative importance or contribution of the above mechanisms to the decline in black upward mobility in Great Migration commuting zones. Doing so would require additional natural experiments or instruments to separately estimate each mechanism's causal effect,

[46]Results on upward mobility are nearly identical after controlling for pre-period murder rates. These results are available from the author upon request. Results on post-period mechanisms are very similar after controlling for pre-period murder rates and are reported in Appendix Figure E2. In Appendix E, I include year-by-year results to demonstrate the timing of changes in Great Migration cities. Appendix Figure E4 shows a steady fall in the urban white share, as white families contemporaneously relocated to suburbs, consistent with the evidence from Boustan (2010). In Appendix Figures E5, E11, and E9, I show that the 1960s were a turning point in terms of policing, crime, and incarceration. To understand the underlying context, I also explore racial tensions and attitudes in the destinations during this period. White residents grew more conservative in their racial views in response to the Migration. As I report in Appendix Table E2, voters in major Great Migration destinations were more likely to support segregationist George Wallace during his run for the presidency in 1968. Rising racial tensions in cities across the US erupted in major riots in the late 1960s. As reported in Appendix Table E1, I find that race riots were of greater intensity in Great Migration cities, lasting longer and involving more injuries and arrests.

which is beyond the scope of this paper. However, the economics literature on the effects of segregation and negative spillovers from the criminal justice system suggest these local changes are likely to have played a role in worsening racial inequality in destination commuting zones.[47]

6.1 Discussion: Aggregate effects of the Migration

This paper addresses the following counterfactual regarding the long-run, intergenerational effects of the Great Migration. Do children growing up in major destination locations today fare worse than those growing up in places less affected by the Migration? The results presented thus far indicate that this is indeed the case, and particularly true for black families. Evidence on post-1970 characteristics of the destinations suggests that absent the Great Migration, black children in the North would have grown up in less segregated neighborhoods, with lower exposure to violent crime, fewer encounters with police, and with a lower likelihood of incarceration.

An important question this analysis abstracts from and which is beyond the scope of this paper, is the aggregate effect of the Great Migration on black economic status. In a simple counterfactual exercise conducted in Appendix Section C.4, I explore these aggregate effects by plotting intergenerational mobility curves by race region, including the counterfactual curve for black families in the North had the Migration not taken place. The latter is shifted up based on the estimates in this paper for the effect of the Migration on black families at different points in the parent income distribution. I make

[47]The evidence presented in Johnson (2019), for example, finds positive effects on earnings and lower rates of incarceration for black children exposed to school desegregation, suggesting that policies encouraging integration may mitigate negative responses to the Great Migration. On policing, Legewie and Fagan (2018); Ang (2018) find that police activity can disrupt educational outcomes for black teens. The evidence on incarceration is mixed with Norris et al. (2018) finding some deterrence effects of sibling incarceration and Dobbie et al. (2018) finding that parental incarceration increases teen crime and pregnancy and lowers subsequent employment for youths from disadvantaged families. A large literature in sociology finds that rising incarceration has increased black-white inequality (Western, 2006).

two important assumptions: 1) that average black parent income in the North and South reflect the equilibrium effect of the Migration on parents; and 2) that the Migration had no effect, either positive or negative on the South. The last parameter needed to estimate the aggregate effect of the Migration is the geographic distribution of the black population in the US in the absence of the 1940-1970 Migration, roughly 77% in the South and 23% in the North.

I conclude that while the Migration likely did reduce gains to parent income for black children in the North—shifting down the intergenerational mobility curve for northern black families—only 23% of black children would have experienced those higher gains to parent income in the absence of the Migration. This effect combined with the substantial positive effect of the Migration on the income of earlier generations, moving the average black child *up* the intergenerational mobility curve, is likely to have resulted in a net positive gain. Any positive impacts of black emigration on the South, which experienced improvements in upward mobility during the second half of the 20th century,[48] would only magnify this positive effect.

7 Conclusion

Over the 20th century, black Americans engaged in perhaps the largest natural experiment in "moving to opportunity" in US history. The Great Migration of African Americans out of southern states into Detroit, Chicago, New York, Los Angeles, and hundreds of other cities across the North and West secured concrete gains for migrants that they reasonably might have believed would persist for future generations.

This paper shows that this was not to be. Using exogenous variation in the extent to which northern locations became destinations during the Great Migration, I show that racial composition changes during this period reduced

[48]Based on the author's analysis of changes in standardized measures of black upward mobility for commuting zones between 1940 and the 2000s.

northern cities' ability to promote positive outcomes for today's cohorts, and specifically harming black men growing up in affected locations.

In response to mid-century changes in the racial composition of northern cities, white families withdrew from shared public schools and urban neighborhoods. Starting in the 1960s, the quality of the urban environment sharply deteriorated. Local governments in Great Migration destinations increased public spending on police in both absolute and relative terms, a reallocation possibly driven by increases in crime or in response to race riots in the late 1960s. These locations remained differentially invested in policing over the next several decades, however, potentially crowding out investments in education, which would have benefited an increasingly disadvantaged urban population.

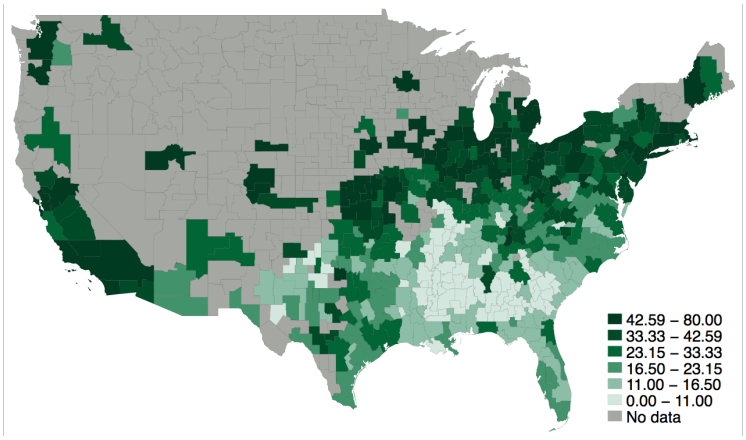
At the height of the rising incarceration in the 1980s and early 1990s, major Great Migration destinations sent substantially more of the black population to federal and state prison than locations less affected by the Migration. Cohorts growing up in the 1960s and 1970s would have been particularly at risk for incarceration. Many studies suggest that incarceration reduces black employment prospects and increases the prevalence of single-parent families, effects that may propagate to future generations. Further research will have to disentangle the long-run impact that increased crime, the race riots of the 1960s, and local governments' responses to each have had on black men's outcomes. A key question is whether alternative strategies for reducing racial inequality in cities can be identified given the sizable gaps under the existing set of policies.

My findings also have implications for policies that incentivize families to move to areas with better opportunities and, in particular, the general equilibrium effects of scaling such programs. In response to millions of black migrants moving North to improve their economic outcomes, receiving northern cities changed in ways that eventually shuttered this pathway to black economic progress. In addition to better understanding the specific policies in locations that contribute to intergenerational mobility, more concerted efforts aimed at

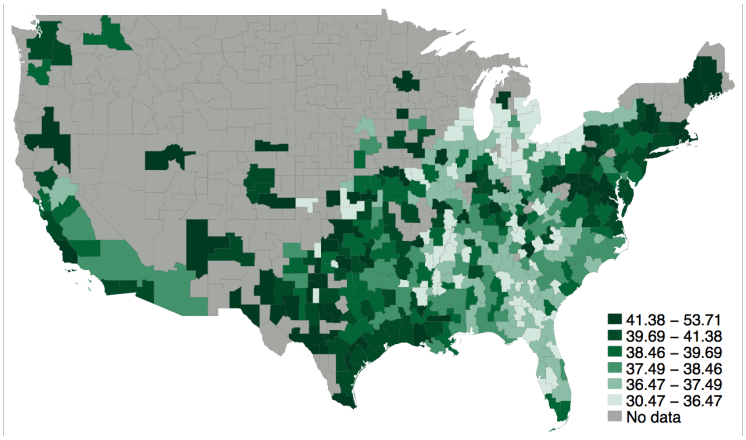
reducing disparities within locations, rather than relocating the disadvantaged, may be warranted.

Figures and tables

Figure 1: Black Upward Mobility in 1940 and 2015



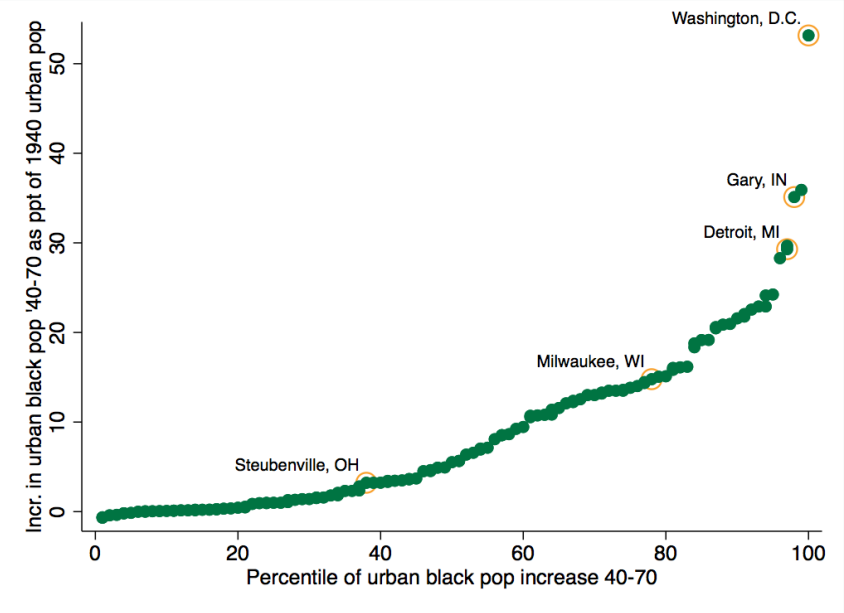
(a) Percentage black teens in median-educ. families with 9-plus years of schooling, 1940



(b) Household inc. rank of black individuals from below-median-income families, 2015

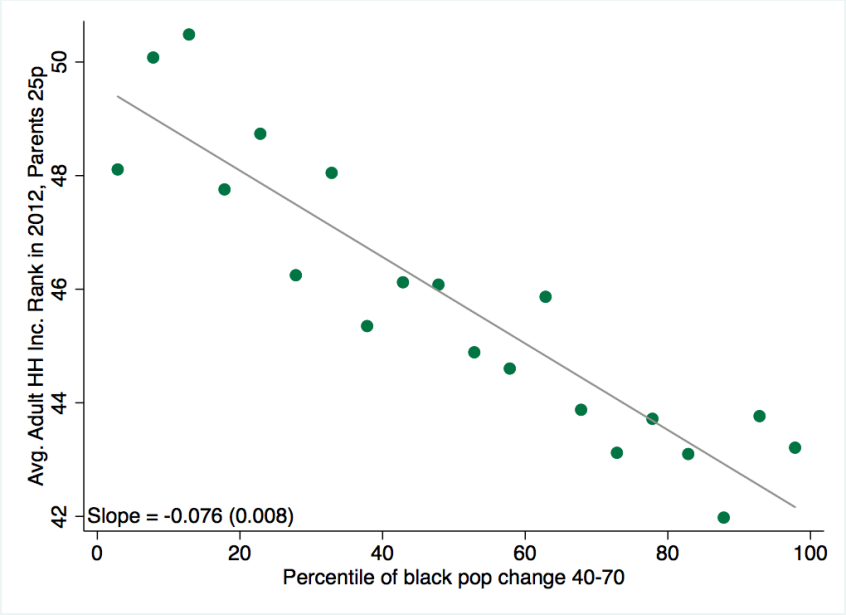
Notes: This figure depicts geographic patterns in black upward mobility in 1940 and 2015. Panel (a) depicts black educational upward mobility in 1940 defined as the percentage of 14-18 year-old boys and 14-16 girls who have at least 9 years of schooling, from households where the household head has between 5 and 8 years of schooling. Panel (b) shows expected mean household income rank in 2015 by childhood commuting zone for the 1978-1983 birth cohorts of black men and women from families at the 25th percentile of the parent income distribution. Darker shades indicate commuting zones with higher levels of upward mobility. *Data sources*: IPUMS 1940 complete count census for panel (a), measure following Card et al. (2018) and Chetty et al. (2018) for panel (b).

Figure 2: Quantiles of urban black share increases, 1940-70



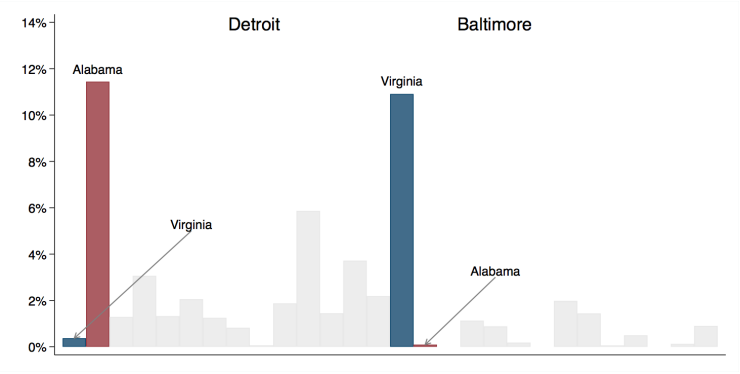
Notes: This figure plots the quantile function of 1940-1970 increases in the urban black population in commuting zones as a share of the total initial 1940 urban population, multiplied by 100 so that the units are percentage points. The CZs in sample are those containing the 294 non-southern mainland cities with information on the black population in both 1940 and 1970 from the *City and County Data Books, 1944-1977* ("CCDB"). Non-southern mainland excludes cities in the following states: Alabama, Alaska, Arkansas, Florida, Georgia, Hawaii, Kentucky, Louisiana, Mississippi, North Carolina, Oklahoma, South Carolina, Tennessee, Texas, Virginia, and West Virginia. Note, Washington, D.C. and cities in Delaware and Maryland were net-receivers of black migrants during the Great Migration and are included in the sample. The city of New Albany, IN is in the Louisville, KY commuting zone, which is included in the sample. Results are robust to excluding this commuting zone. *Data sources*: CCDB.

Figure 3: Relationship between 1940-1970 black population change and upward mobility in 2012

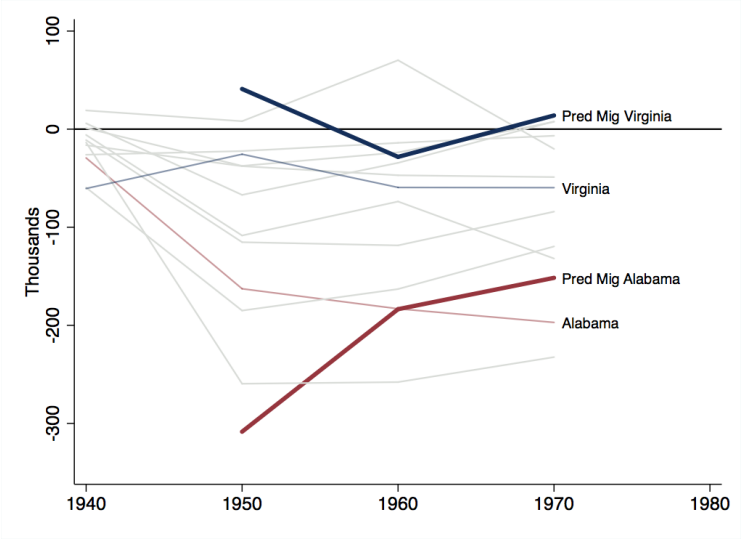


Notes: This binned scatterplot depicts the relationship between average upward mobility in the 2000s for men and women with low income parents and the percentile of actual black population increase during the Great Migration (1940 to 1970) for northern commuting zones. The unit of observation is a commuting zone. The right hand side variable is grouped into 20 bins (5 percentiles each). Upward mobility is defined as expected mean household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. *Data sources*: IPUMS complete count 1940 US census; CCDB.

Figure 4: Great Migration shift-share instrument



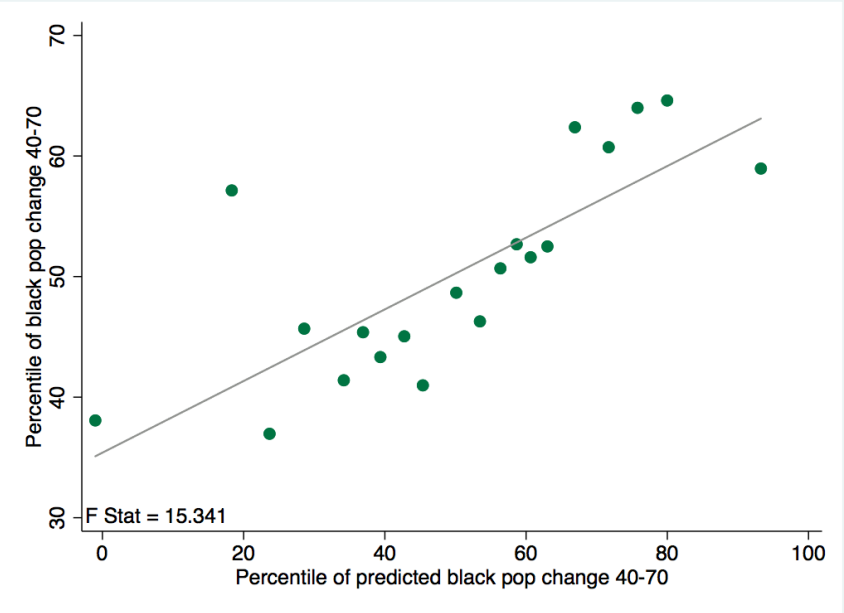
(a) 1935-1940 black southern migrants' origin counties, Detroit vs. Baltimore



(b) Southern state net-migration, 1940-1970

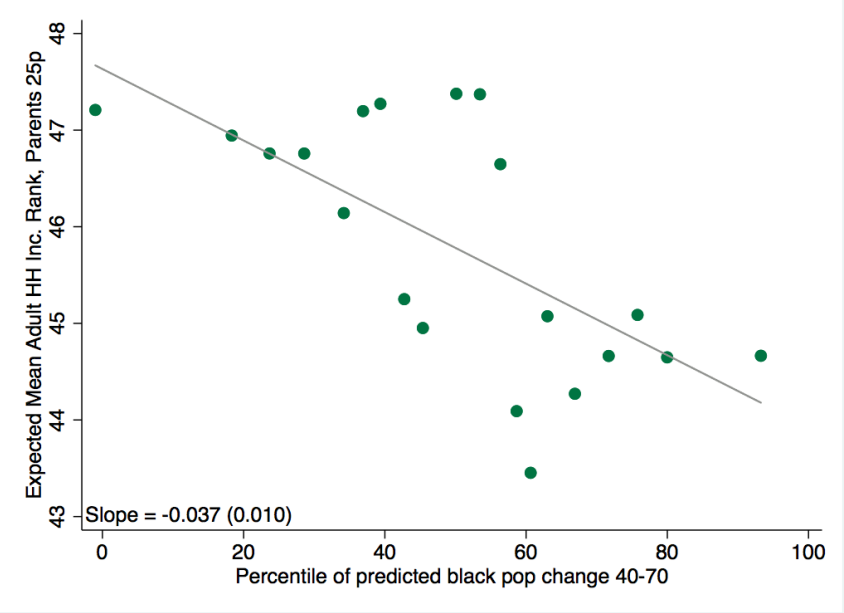
Notes: This figure illustrates the variation underlying the shift-share instrument for urban black population change in northern commuting zones. Panel (a) shows the share of recent black southern migrants (those who migrated between 1935 and 1940) living in Detroit and Baltimore from the largest sending county in each southern state. For Alabama and Virginia, these are Jefferson County (Birmingham) and Richmond City County, respectively. Detroit received the plurality of its migrants from Alabama, Baltimore from Virginia. Panel (b) shows net-migration and predicted net-migration for southern states each decade from 1940-1970, with Alabama and Virginia highlighted. Negative numbers indicate outmigration. Darkened lines indicate net-migration predicted using one-decade lagged southern county agricultural and World War II spending measures. Appendix B describes the construction of the instrument based on this variation. I use LASSO to select predictors each decade, interacting predicted migration with the share of recent black southern migrants from each county, summing up over all southern counties. The procedure yields counterfactual increases in the urban black population from 1940-1970. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); and Fouka et al. (2018).

Figure 5: First stage on black population change



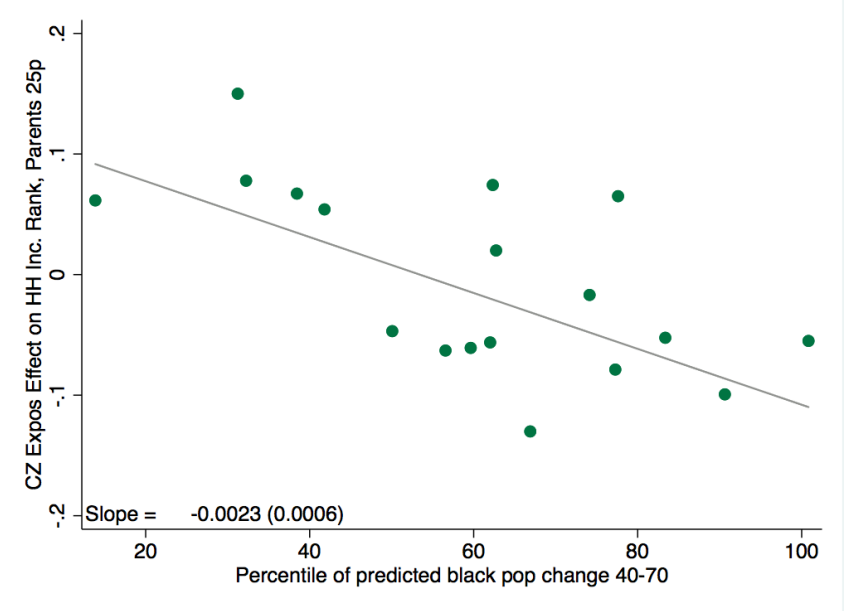
Notes: This binned scatterplot depicts the relationship between the percentile of actual black population increase during the Great Migration (1940 to 1970) for northern commuting zones and the instrument for black population increase over the same period. The instrument is the percentile of predicted 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. The unit of observation is a commuting zone. The right-hand-side variable is grouped into 20 bins (5 percentiles each). Both the left-hand- and right-hand-side variables have been residualized on the set of baseline 1940 controls, including share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016).

Figure 6: Great Migration reduced average upward mobility in northern commuting zones



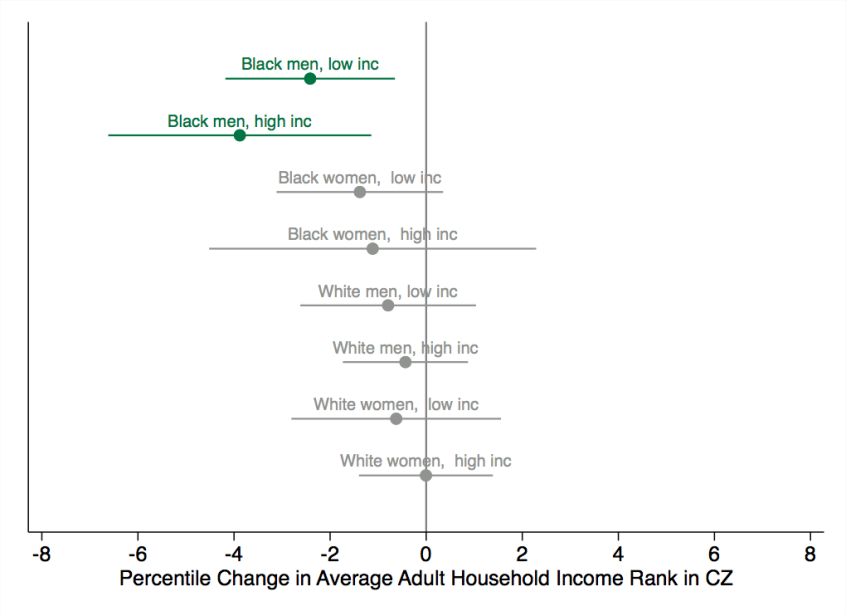
Notes: This binned scatterplot depicts the relationship between average upward mobility in the 2000s for men and women with low income parents and the instrument for black population increases during the Great Migration. The unit of observation is a commuting zone. The right hand side variable is grouped into 20 bins (5 percentiles each). Upward mobility is defined as expected mean household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. The instrument 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. Both the left hand and right hand side variables have been residualized on the set of baseline 1940 controls, including share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Figure 7: Childhood in Great Migration CZs lowers adult income of children from low income families



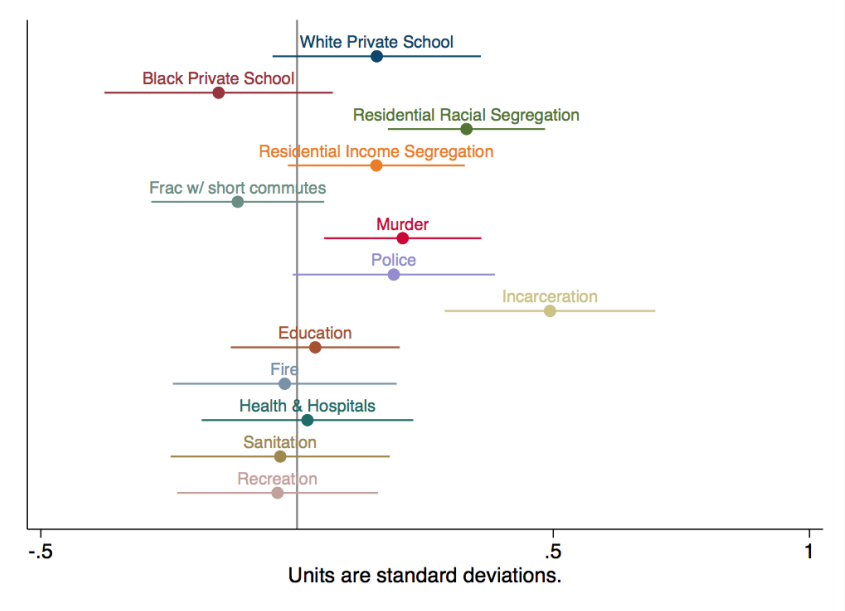
Notes: This binned scatterplot depicts the relationship between commuting zone childhood exposure effects in the 2000s for men and women with low income parents and the instrument for black population increases during the Great Migration. The unit of observation is a commuting zone. The right hand side variable is grouped into 20 bins (5 percentiles each). Childhood exposure effects are the estimated causal impact of one additional year of childhood in the commuting zone on adult household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. The instrument 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. Both the left hand and right hand side variables have been residualized on the set of baseline 1940 controls, including share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Figure 8: Race and gender heterogeneity in impact of Great Migration on upward mobility



Notes: This figure plots coefficients from regressions of average upward mobility in the 2000s for men and women from low and high income parents on the instrument for black population increases during the Great Migration, in approximately one standard deviation units. The unit of observation is a commuting zone. Upward mobility is defined as expected mean household income rank where income is measured from IRS tax returns for cohorts born between 1978 and 1983. Pooled income refers to mean household income rank, pooling across men and women. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); Chetty et al. (2018).

Figure 9: Increased segregation, crime, policing, and incarceration in Great Migration CZs



Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions. The dependent variables are standardized mean 1970-2000 white and black private school enrollment rates; the Theil indices in residential racial and income segregation in 2000; the fraction of families in 2000 with commute times less than 15 minutes; mean 1977-2002 murders per 100,000 of the population; mean 1983-2000 incarcerated per 100,000 of the population; and mean 1972-2002 government expenditure shares by category. The unit of observation is a commuting zone. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); see Appendix D for the full list of data sources on each of the mechanisms.

Table 1: Placebo test of identification strategy using pre-1940 upward mobility and educational attainment

	Fraction of teens with low occ. score fathers attending school			Median adult education
	1920	1930	1940	1940
$\hat{GM}$	0.011	0.023	0.018	-0.013
	(0.024)	(0.029)	(0.015)	(0.009)
Baseline mean	65.477	74.912	80.676	27.355
Std Dev	7.425	8.674	5.710	2.863
Observations	130	130	130	130
Baseline Controls	Y	Y	Y	Y

Notes: This table reports the effect of the Great Migration on pre-1940 educational upward mobility and attainment. In columns 1 through 3, the dependent variable is the school attendance rate of 14-17 year-old boys and girls with below-median occupation score fathers in 1920, 1930, and 1940, respectively. In column 4 the dependent variable is median education attainment of adults aged 25 and older in 1940. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016).

TABLE 2: LOWER AVERAGE UPWARD MOBILITY IN 2000S FOR LOW INCOME FAMILIES IN GREAT MIGRATION CZS

	First Stage on GM					
$\hat{GM}$	0.297	0.297	0.297	0.297	0.297	0.297
	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)
F-Stat	15.34	15.34	15.34	15.34	15.34	15.34
	Household Income Rank			Individual Income Rank		
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.0655	-0.0570	-0.0742	-0.0331	-0.00375	-0.0618
	(0.00995)	(0.0101)	(0.0104)	(0.0108)	(0.0137)	(0.0108)
R-squared	0.571	0.528	0.593	0.345	0.254	0.492
	Reduced Form					
$\hat{GM}$	-0.0370	-0.0308	-0.0432	-0.0282	-0.0128	-0.0439
	(0.00974)	(0.00973)	(0.0103)	(0.00965)	(0.0121)	(0.0101)
R-squared	0.481	0.451	0.495	0.341	0.260	0.443
	Two-stage least squares					
GM	-0.125	-0.104	-0.145	-0.0950	-0.0432	-0.148
	(0.0328)	(0.0318)	(0.0354)	(0.0353)	(0.0410)	(0.0386)
R-squared	0.447	0.446	0.435	0.169	0.203	0.230
N	130	130	130	130	130	130
Mean Rank	45.79	47.04	44.55	45.54	42.74	48.29
SD Rank	3.379	3.283	3.617	2.972	3.527	3.375
SD GM	28.98	28.98	28.98	28.98	28.98	28.98

Notes: This table reports the estimated impact of the Great Migration on average upward mobility in the 2000s for men and women with low income parents. The unit of observation is a commuting zone. Dependent variable is expected mean individual or household income rank for individuals with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. Pooled income refers to household income, pooling across men and women. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Table 3: Childhood exposure to Great Migration CZs lowers upward mobility for low income families

	First Stage on GM					
$\hat{GM}$	0.266	0.263	0.269	0.264	0.263	0.269
	(0.0640)	(0.0639)	(0.0645)	(0.0641)	(0.0642)	(0.0645)
F-Stat	17.27	16.91	17.38	16.99	16.72	17.35
	Household Income Rank			Individual Income Rank		
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.00256	-0.00169	-0.00438	-0.00210	0.000437	-0.00433
	(0.000848)	(0.00125)	(0.00126)	(0.000865)	(0.00125)	(0.00134)
R-squared	0.224	0.115	0.233	0.190	0.0345	0.208
	Reduced Form					
$\hat{GM}$	-0.00232	-0.00209	-0.00318	-0.00189	-0.00111	-0.00276
	(0.000631)	(0.000930)	(0.000967)	(0.000647)	(0.000939)	(0.00103)
R-squared	0.249	0.138	0.226	0.206	0.0445	0.188
	Two-stage least squares					
GM	-0.00871	-0.00794	-0.0118	-0.00716	-0.00424	-0.0103
	(0.00279)	(0.00381)	(0.00393)	(0.00271)	(0.00368)	(0.00397)
R-squared	-0.110	-0.0656	0.0159	-0.0369	-0.0766	0.0793
N	130	130	130	130	130	130
Precision Wt	Y	Y	Y	Y	Y	Y
Mean Expos FX	-0.0160	-0.0151	-0.0303	0.0223	0.0236	-0.0000692
SD Expos FX	0.172	0.235	0.259	0.172	0.226	0.271
SD GM	24.82	24.42	24.84	24.99	24.76	24.95

Notes: This table reports the estimated impact of the Great Migration on commuting zone childhood exposure effects. The unit of observation is a commuting zone. Dependent variable is the estimated causal impact of one additional year of childhood in the commuting zone on adult household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. The unit of observation is a commuting zone. Pooled income refers to household income, pooling across men and women. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Table 4: Great Migration impact on average upward mobility for black families in the 2000s

	First Stage on GM					
$\hat{GM}$	0.297	0.297	0.297	0.297	0.297	0.297
	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)
F-Stat	15.34	15.34	15.34	15.34	15.34	15.34
		Low Income			High Income	
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.000563	-0.000435	-0.000747	-0.000767	-0.000553	-0.00105
	(0.0000956)	(0.000112)	(0.000114)	(0.000147)	(0.000219)	(0.000175)
R-squared	0.428	0.295	0.449	0.358	0.184	0.360
	Reduced Form					
$\hat{GM}$	-0.000183	-0.000148	-0.000258	-0.000269	-0.000119	-0.000415
	(0.0000930)	(0.000103)	(0.000113)	(0.000140)	(0.000196)	(0.000170)
R-squared	0.286	0.220	0.284	0.237	0.143	0.207
	Two-stage least squares					
GM	-0.000591	-0.000476	-0.000833	-0.000869	-0.000384	-0.00134
	(0.000260)	(0.000305)	(0.000311)	(0.000401)	(0.000599)	(0.000482)
R-squared	0.427	0.294	0.447	0.355	0.180	0.346
N	129	129	129	129	129	129
Mean Rank	0.332	0.352	0.312	0.453	0.467	0.442
SD Rank	0.0275	0.0290	0.0333	0.0398	0.0528	0.0476
SD GM	28.80	28.80	28.80	28.80	28.80	28.80

Notes: This table reports the estimated impact of the Great Migration on average upward mobility in the 2000s for black men and women with high income parents. The unit of observation is a commuting zone. Dependent variable is expected mean household income rank for individuals with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1978 and 1983. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty et al. (2018).

Table 5: Great Migration impact on average upward mobility for white families in the 2000s

	First Stage on GM					
$\hat{GM}$	0.297	0.297	0.297	0.297	0.297	0.297
	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)
F-Stat	15.34	15.34	15.34	15.34	15.34	15.34
		Low Income			High Income	
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.000155	-0.000114	-0.000183	-0.000218	-0.000186	-0.000241
	(0.000120)	(0.000132)	(0.000111)	(0.0000793)	(0.0000829)	(0.0000787)
R-squared	0.284	0.278	0.292	0.374	0.342	0.393
	Reduced Form					
$\hat{GM}$	-0.0000757	-0.0000640	-0.0000813	-0.0000238	-0.000000508	-0.0000442
	(0.000108)	(0.000118)	(0.0000993)	(0.0000726)	(0.0000752)	(0.0000725)
R-squared	0.277	0.275	0.280	0.336	0.314	0.348
	Two-stage least squares					
GM	-0.000255	-0.000215	-0.000274	-0.0000802	-0.00000171	-0.000149
	(0.000350)	(0.000384)	(0.000322)	(0.000233)	(0.000245)	(0.000229)
R-squared	0.280	0.274	0.288	0.358	0.315	0.386
N	130	130	130	130	130	130
Mean Rank	0.452	0.463	0.443	0.606	0.618	0.596
SD Rank	0.0316	0.0346	0.0293	0.0223	0.0227	0.0225
SD GM	28.98	28.98	28.98	28.98	28.98	28.98

Notes: This table reports the estimated impact of the Great Migration on average upward mobility in the 2000s for white men and women with high income parents. The unit of observation is a commuting zone. Dependent variable is expected mean household income rank for individuals with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1978 and 1983. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty et al. (2018).

Table 6: Robustness of effects of childhood exposure to Great Migration CZs

$\hat{GM}$	-0.00142	-0.00260	-0.00232	-0.00213	-0.00255	-0.00237	-0.00226	-0.00232
	(0.000602)	(0.000537)	(0.000636)	(0.000801)	(0.000642)	(0.000616)	(0.000707)	(0.000629)
R-squared	0.0694	0.232	0.249	0.289	0.256	0.252	0.254	0.249
N	130	130	130	130	130	130	130	130
Precision Wt	Y	Y	Y	Y	Y	Y	Y	Y
Census Div FE	N	Y	Y	Y	Y	Y	Y	Y
Baseline Controls	N	N	Y	Y	Y	Y	Y	Y
1940 Black Share Quartile FEs	N	N	N	Y	N	N	N	N
Southern Mob	N	N	N	N	Y	N	N	N
White South Mig	N	N	N	N	N	Y	N	N
Eur Mig	N	N	N	N	N	N	Y	N
Emp Bartik	N	N	N	N	N	N	N	Y

Notes: This table reports robustness of the estimated impact of the Great Migration on commuting zone childhood exposure effects to several alternative specifications. The unit of observation is a commuting zone. Dependent variable is commuting zone childhood exposure effects in the 2000s for men and women with low income parents. Childhood exposure effects are the estimated causal impact of one additional year of childhood in the commuting zone on adult household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. Pooled income refers to household income, pooling across men and women. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors clustered by state. Conley standard errors do not change precision of estimates. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Table 7: Robustness of Great Migration's effects on black men's upward mobility

$\widehat{GM}$	-0.000381	-0.000364	-0.000264	-0.000162	-0.000236	-0.000270	-0.000240	-0.000275
	(0.000103)	(0.000116)	(0.000112)	(0.0000958)	(0.000124)	(0.000112)	(0.000102)	(0.000116)
R-squared	0.127	0.138	0.185	0.296	0.188	0.197	0.189	0.188
N	129	129	129	129	129	129	129	129
Precision Wt								
Census Div FE	N	Y	Y	Y	Y	Y	Y	Y
Baseline Controls	N	N	Y	Y	Y	Y	Y	Y
1940 Black Share Quartile FEs	N	N	N	Y	N	N	N	N
Southern Mob	N	N	N	N	Y	N	N	N
White South Mig	N	N	N	N	N	Y	N	N
Eur Mig	N	N	N	N	N	N	Y	N
Emp Bartik	N	N	N	N	N	N	N	Y

Notes: This table reports robustness of the estimated impact of the Great Migration on black men's upward mobility to several alternative specifications. The unit of observation is a commuting zone. Dependent variable is expected mean individual income rank for individuals with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1978 and 1983. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors clustered by state. Conley standard errors do not change precision of estimates. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); Chetty et al. (2018).

References

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

Alesina, A., R. Baqir, and C. Hoxby (2004). Political jurisdictions in heterogeneous communities. *Journal of Political Economy 112*(2), 348–396.

Altonji, J. G. and D. Card (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pp. 201–234. University of Chicago Press.

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.

Andrews, R., M. Casey, B. L. Hardy, and T. D. Logan (2017). Location matters: Historical racial segregation and intergenerational mobility. *Economics Letters 158*, 67 – 72.

Ang, D. (2018). The Effects of Police Violence on Inner-City Students. Working Paper.

Belloni, A., V. Chernozhukov, and C. Hansen (2011). Lasso methods for gaussian instrumental variables models. Working Paper.

Belloni, A., V. Chernozhukov, and C. Hansen (2014). High-dimensional methods and inference on structural and treatment effects. *Journal of Economic Perspectives 28*(2), 29–50.

Bertrand, M. and J. Pan (2013). The trouble with boys: Social influences and the gender gap in disruptive behavior. *American Economic Journal: Applied Economics 5*(1), 32–64.

Black, D. A., S. G. Sanders, E. J. Taylor, and L. J. Taylor (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., P. Hull, and X. Jaravel (2019, September). Quasi-experimental shift-share research designs. WP 24997, NBER.

Boustan, L. and M. Tabellini (2018). Black Out-Migration and Southern Political Realignment. Working Paper.

Boustan, L. P. (2009). Competition in the promised land: Black migration and racial wage convergence in the north, 1940–1970. *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*. Princeton University Press.

Calderon, A., V. Fouka, and M. Tabellini (2019). Racial Diversity, Electoral Preferences, and the Supply of Policy: the Great Migration and Civil Rights. Working Paper.

Card, D., C. Domnisoru, and L. Taylor (2018). The intergenerational transmission of human capital: Evidence from the golden age of upward mobility. Working Paper.

Card, D., A. Mas, and J. Rothstein (2008). Tipping and the dynamics of segregation. *Quarterly Journal of Economics 123*(1), 177–218.

Carter, G. L. (1986). The 1960s black riots revisited: city level explanations of their severity. *Sociological Inquiry 56*(2), 210–228.

Charles, K. K., E. Hurst, and M. Schwartz (2019). The transformation of manufacturing and the decline in us employment. *NBER Macroeconomics Annual 33*(1), 307–372.

Chernozhukov, V., D. Chetverikov, M. Demirer, E. Duflo, C. Hansen, W. Newey, and J. Robins (2018). Double/debiased machine learning for treatment and structural parameters. Working Paper.

Chetty, R. and N. Hendren (2018a). The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects. *Quarterly Journal of Economics 133*(3), 1107–1162.

Chetty, R. and N. Hendren (2018b). The impacts of neighborhoods on intergenerational mobility II: County-level estimates. *Quarterly Journal of Economics 133*(3), 1163–1228.

Chetty, R., N. Hendren, M. R. Jones, and S. R. Porter (2018). Race and economic opportunity in the united states: An intergenerational perspective. WP 24441, NBER.

Chetty, R., N. Hendren, and L. F. Katz (2016). The effects of exposure to better neighborhoods on children: New evidence from the Moving to Opportunity experiment. *American Economic Review 106*(4), 855–902.

Collins, W. J. (1997). When the tide turned: Immigration and the delay of the Great Black Migration. *Journal of Economic History 57*(3), 607–632.

Collins, W. J. and R. A. Margo (2007). The economic aftermath of the 1960s riots in American cities: Evidence from property values. *Journal of Economic History 67*(4), 849–883.

Collins, W. J. and M. H. Wanamaker (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.

Cutler, D. M. and E. L. Glaeser (1997). Are ghettos good or bad? *Quarterly Journal of Economics 112*(3), 827–872.

Davis, J. and B. Mazumder (2018). Racial and ethnic differences in the geography of intergenerational mobility. Working Paper.

Dobbie, W., H. Grönqvist, S. Niknami, M. Palme, and M. Priks (2018). The intergenerational effects of parental incarceration. WP 24186, NBER.

Eriksson, K. (2018). Moving North and into jail? The Great Migration and black incarceration. *Journal of Economic Behavior & Organization*.

Eriksson, K. and G. T. Niemesh (2016). Death in the Promised Land: the Great Migration and Black Infant Mortality. Working Paper.

Figlio, D., K. Karbownik, J. Roth, M. Wasserman, et al. (2016). School quality and the gender gap in educational achievement. *American Economic Review 106*(5), 289–95.

Fouka, V., S. Mazumder, and M. Tabellini (2018). From Immigrants to Americans: Race, Status, and Assimilation During the Great Migration. Working paper.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018). Bartik instruments: What, when, why, and how. WP 24408, NBER.

Graham, B. (2016). Identifying and estimating neighborhood effects. WP 22575, NBER.

Hilger, N. G. (2015). The Great Escape: Intergenerational Mobility in the United States since 1940. WP 21217, NBER.

Hornbeck, R. and E. Moretti (2018). Who Benefits From Productivity Growth? Direct and Indirect Effects of Local TFP Growth on Wages, Rents, and Inequality. WP 24661, NBER.

Hornbeck, R. and S. Naidu (2014). When the levee breaks: black migration and economic development in the American South. *American Economic Review 104*(3), 963–90.

Johnson, R. C. (2019). *Children of the dream: Why school integration works*. Hachette UK.

Kling, J. R., J. B. Liebman, and L. F. Katz (2007). Experimental analysis of neighborhood effects. *Econometrica 75*(1), 83–119.

Legewie, J. and J. Fagan (2018). Aggressive policing and the educational performance of minority youth. Working Paper.

Leibbrand, C., C. Massey, J. T. Alexander, and S. Tolnay (2019). Neighborhood attainment outcomes for children of the great migration. *American Journal of Sociology 125*(1), 141–183.

Ludwig, J., G. J. Duncan, L. A. Gennetian, L. F. Katz, R. C. Kessler, J. R. Kling, and L. Sanbonmatsu (2012). Neighborhood effects on the long-term well-being of low-income adults. *Science 337*(6101), 1505–1510.

Margo, R. A. (1990). *Race and schooling in the South, 1880-1950: An economic history*. University of Chicago Press.

Margo, R. A. (1991a). Segregated Schools and the Mobility Hypothesis: A Model of Local Government Discrimination. *Quarterly Journal of Economics 106*(1), 61–73.

Margo, R. A. (1991b). Segregated schools and the mobility hypothesis: A model of local government discrimination. *The Quarterly Journal of Economics 106*(1), 61–73.

Massey, D. S. and N. A. Denton (1993). *American apartheid: Segregation and the making of the underclass*. Harvard University Press.

Mazumder, B. (2014). Black-white differences in intergenerational economic mobility in the United States. Working paper.

Muller, C. (2012). Northward migration and the rise of racial disparity in American incarceration, 1880–1950. *American Journal of Sociology 118*(2), 281–326.

Naidu, S. (2010). Recruitment restrictions and labor markets: Evidence from the postbellum US South. *Journal of Labor Economics 28*(2), 413–445.

Norris, S., M. Pecenco, and J. Weaver (2018). The Effects of Parental and Sibling Incarceration: Evidence from Ohio. Technical report, Working Paper.

Rothstein, J. (2019). Inequality of educational opportunity? schools as mediators of the intergenerational transmission of income. *Journal of Labor Economics 37*(S1), S85–S123.

Sampson, R. J., J. D. Morenoff, and T. Gannon-Rowley (2002). Assessing "Neighborhood Effects": Social processes and new directions in research. *Annual Review of Sociology 28*(1), 443–478.

Sequeira, S., N. Nunn, and N. Qian (2019, 2019). Immigrants and the Making of America. *Review of Economic Studies*.

Shertzer, A. and R. P. Walsh (2016). Racial sorting and the emergence of segregation in American cities. WP 22077, NBER.

Stuart, B. A. and E. J. Taylor (2017). The Effect of Social Connectedness on Crime: Evidence from the Great Migration. Working Paper.

Stuart, B. A., E. J. Taylor, et al. (2018). Migration Networks and Location Decisions: Evidence from US Mass Migration. Working Paper.

Sugrue, T. J. (1996). The origins of the urban crisis. *Princeton, NJ: Princeton University Press*.

Tabellini, M. (2018). Racial Heterogeneity and Local Government Finances: Evidence from the Great Migration. Working Paper.

United States National Advisory Commission on Civil Disorders, Kerner, O. and Wicker, T. (1968). *Report of the National Advisory Commission on Civil Disorders: Special Introduction by Tom Wicker of The New York Times; [Chairman, Otto Kerner].* Bantam Books.

Western, B. (2006). *Punishment and inequality in America*. Russell Sage Foundation.

Whatley, W. C. (1985). A history of mechanization in the cotton South: The institutional hypothesis. *Quarterly Journal of Economics 100*(4), 1191–1215.

Wilkerson, I. (2011). *The warmth of other suns: The epic story of America's Great Migration*. Vintage.

Wilson, W. (1990). *The Truly Disadvantaged: The Inner City, the Underclass, and Public Policy*. Sociology, urban studies, black studies. University of Chicago Press.

For online publication: Appendices

Appendix A Geographic crosswalks

A.1 Historical county to 1990 CZ crosswalks

To construct the geographic crosswalks used in the analysis, polygon shapefiles for US geographic areas were downloaded from *National Historical Geographic Information Systems* ("NHGIS") and merged based on spatial location using ArcGIS software. Listed below are the raw files and the website where they can be downloaded.

The following procedure was used to crosswalk between historical county boundaries (1920-1940) and 1990 commuting zones. Using ArcGIS, polygon shapefiles were converted to points representing the centroid of the polygon and then merged to the commuting zone polygon containing the centroid.[49]

Raw data files from NHGIS (`https://www.nhgis.org/`)

1. US_county_1940.shp
2. US_county_1990.shp
3. US_msacmsa_1990.shp
4. US_necma_1990.shp
5. US_smsa_1970.shp

Because CZs are aggregations of 1990 counties, historical counties are matched to the CZ in which the geographic centroid of their 1940 borders falls. This procedure allows me to rapidly assign many historical county-level datasets to 1990 commuting zones. However, this procedure may result in assignment errors if county borders change substantially over time.

[49]The commuting zone polygon was created by dissolving borders between counties in the commuting zone using the crosswalk between 1990 counties and commuting zones provided by David Dorn at `http://www.ddorn.net/data/cw_cty_czone.zip`.

The vast majority of the counties in the paper's sample did not experience boundary changes over the timeframe of analysis. The *Inter-university Consortium for Political and Social Research* ("ICPSR") code assigned to each county has a numeric flag for counties that were dissolved and/or merged before 1970 (a final digit of 5); no counties in the sample fall into this category.

Data on changes in the county boundaries can be obtained from the Atlas of Historical County Boundaries at the Newberry Library.[50] Of the 776 counties in the sample, only 32 had mapped boundary changes, representing 4.12 percent. The majority of these boundary changes are referred to by the Newberry Library researchers as "small," many too small to map. While most counties in the sample changed boundaries rarely if at all, one notable outlier is the Denver, Colorado metro area, where Denver, Jefferson, Adams, and Arapahoe Counties swapped patches of land at a sustained pace between 1940 and 1970. However, these counties fall within the Denver-Boulder-Longmont, CO commuting zone, therefore their border changes do not affect the accuracy of the matching procedure.

A.2 City name standardizing

Names of cities in all city-level data digitized or collected for this paper were standardized first to be consistent with those in the 2010 U.S. place point shapefile from National Historical Geographic Information Systems. Places in the 2010 US place point file were matched to the county or CZ they fell within, allowing for the matching of city-level datasets to counties and commuting zones. In instances where a city did not appear in the 2010 US place point file, the city was assigned to the same commuting zone as places geographically close to the city in the 2010 US place point file.

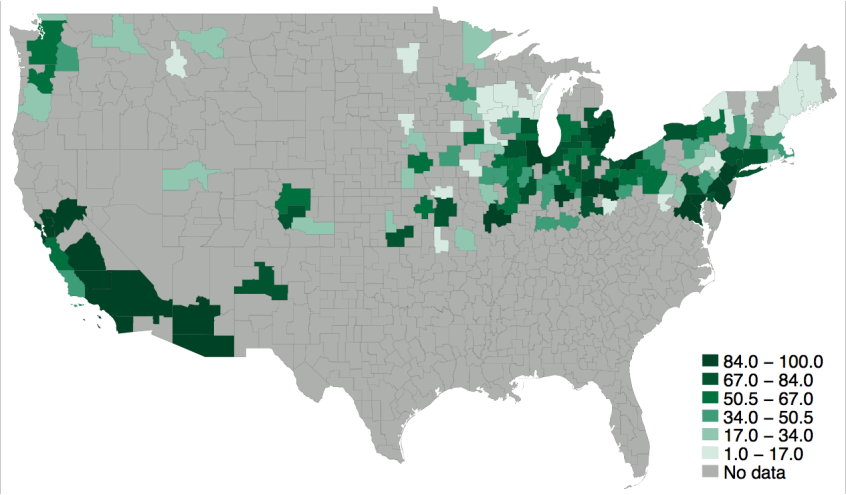
[50] See the information available at the following webpage: `https://publications.newberry.org/ahcbp/`.

Table A1: 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.

Figure A1: Map of 1940-70 change in the black population



Notes: This map depicts Great Migration commuting zones and each CZ's percentile change in the black population between 1940 and 1970. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016).

Appendix B Great Migration shift-share instrument

This appendix details the construction of the instrument from these two sources of variation, beginning with the construction of the shares from pre-1940 migrant location choices and following with the prediction of migration from southern counties using a machine learning approach.

B.1 Pre-1940 black southern migrant shares

I measure black southern migrant shares using the complete count 1940 census. The 1940 census was the first census in which enumerators asked individuals to report their place of residence in 1935. There are several advantages to this approach of measuring pre-1940 black migration patterns. The first is that I am able to observe the universe of enumerated recent black southern migrants, generating a nearly complete picture of recent migration flows into northern cities. The second is that the census microdata allow me to observe fine geographies for individuals' 1935 place of residence, including city and county. I define a recent black southern migrant as a black individual who reported a southern county of residence in 1935, but was enumerated in a different county (whether southern or not) in 1940. There are over 340,000 such individuals.

Using this population of recent black southern migrants, I construct the share of migrants from each 1935 southern county j who settled in a northern city c by 1940:

$$\omega_{jc}^{1935-1940} = \frac{b_{cj}}{b_j} \tag{11}$$

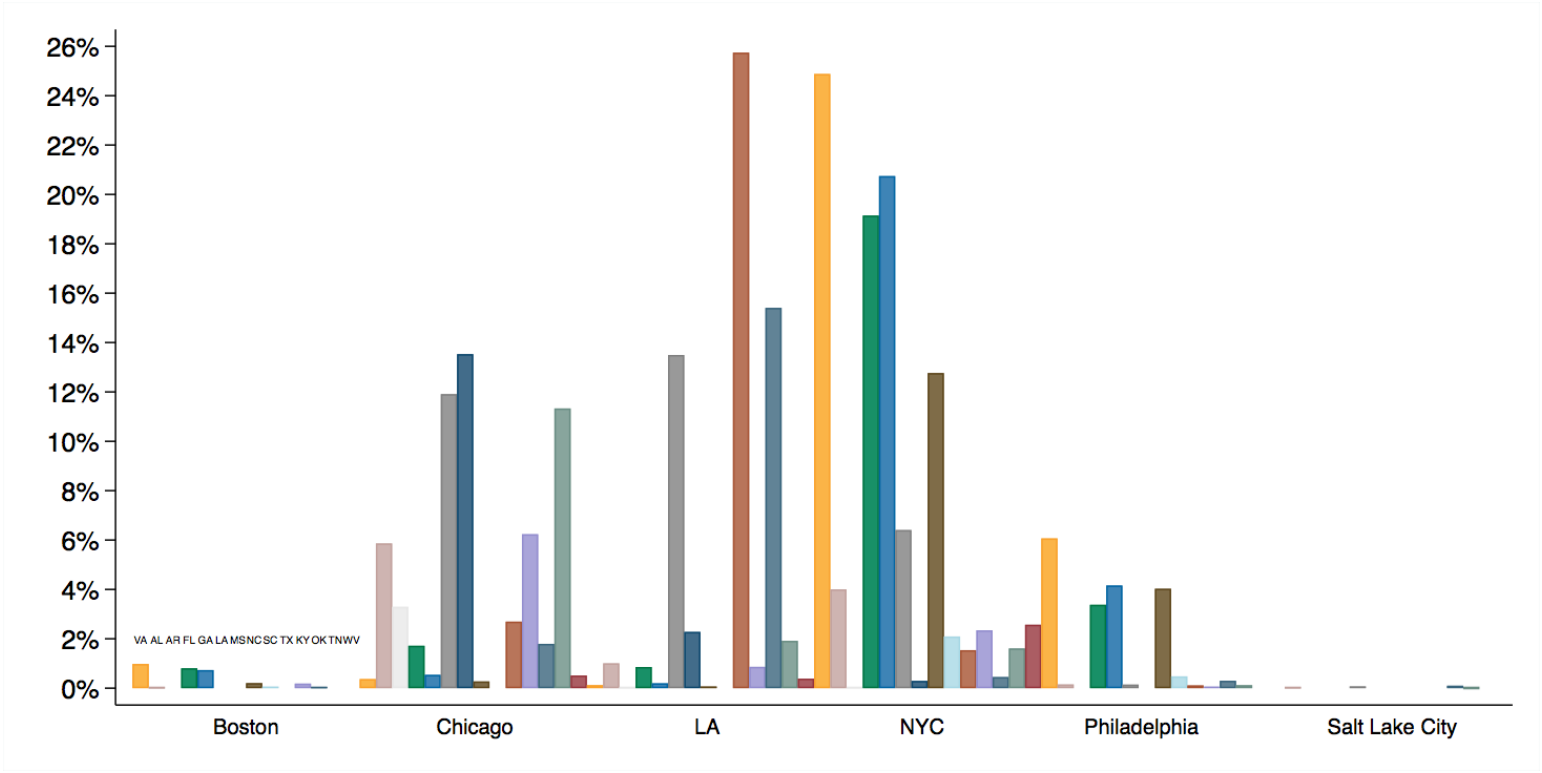
where b_j is the number of black individuals who listed j as their county of residence in 1935, and b_{cj} is the number of black individuals who were enumerated

in city c.

Figure B1 depicts $\omega_{jc}^{1935-1940}$ for a select group of cities and southern counties. Depicted is the share of 1935-1940 black migrants from the largest sending county for each southern state who settled in the following cities: Boston, Chicago, Los Angeles, New York, Philadelphia, and Salt Lake City. The figure captures the immense heterogeneity in settlement patterns across and volume of migration into the cities in question.

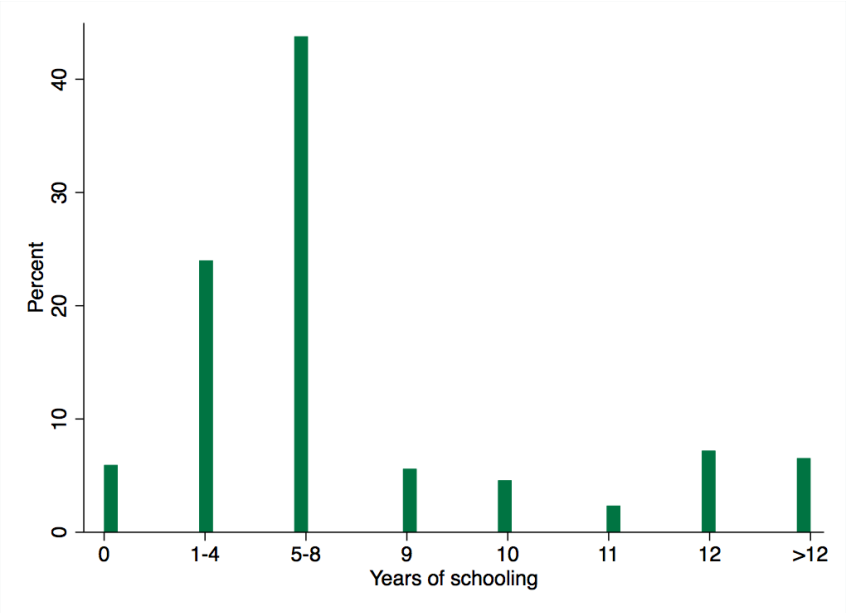
Descriptive evidence on migrant characteristics Figure B2 shows the educational distribution for 1935-1940 black southern migrants aged 25 plus. Finally, Figure B3 explores selection of migrants relative to northern incumbent black families in the North. If anything, black children from low socioeconomic status families whose parents were southern born had better educational outcomes than those whose parents were northern born.

Figure B1: 1935-39 black southern migrant composition in select northern cities



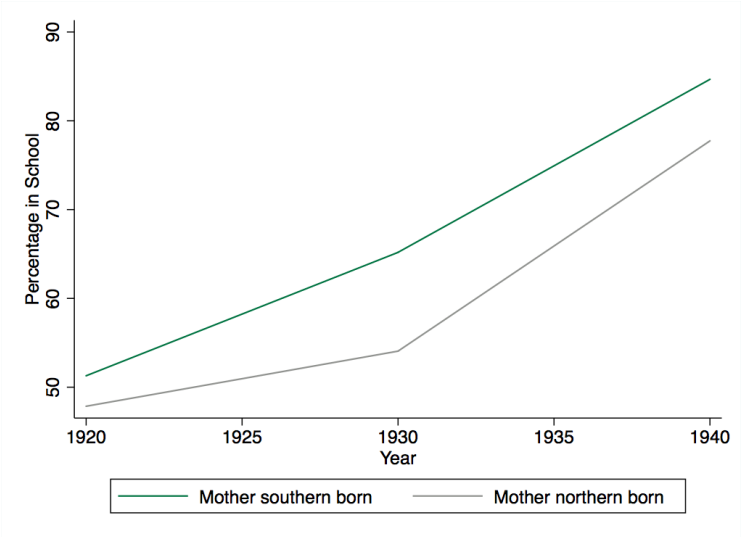
Notes: This figure shows the share of 1935-1939 black migrants from the largest sending county for each southern state who settled in the following cities: Boston, Chicago, Los Angeles ("LA"), New York City ("NYC"), Philadelphia, and Salt Lake City. *Data source*: IPUMS Complete Count 1940 US Census.

Figure B2: 1935-1940 black southern migrant educational attainment

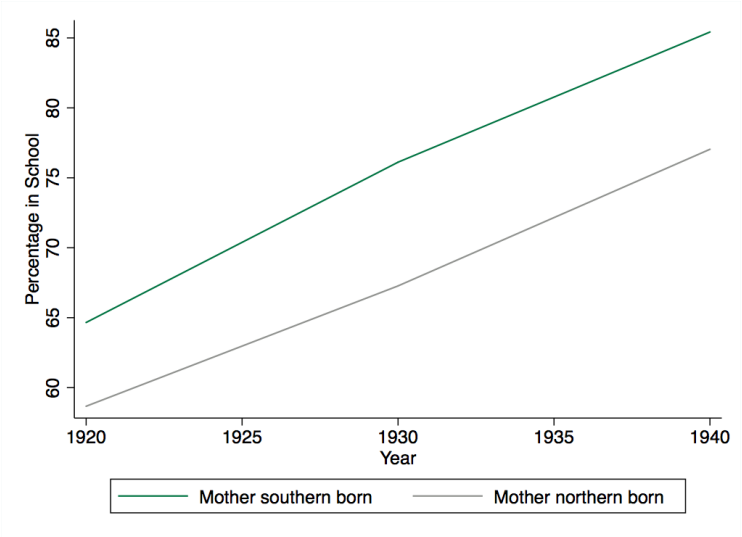


Notes: Histogram of years of schooling for 1935-1940 black southern migrants aged 25 plus. *Data source*: IPUMS Complete Count 1940 US Census.

Figure B3: School attendance for black teens in North with southern- vs. northern-born mothers



(a) Black teens with illiterate mothers



(b) Black teens with low-occ-score fathers

Notes: 1920-1940 school attendance rates (in percentage points) for black 14-17 year-old boys and girls by mother birth region. *Data sources*: IPUMS Complete Count 1920-1940 US Censuses.

B.2 Post-LASSO prediction of southern county net migration

Under the assumption that county-level variation in southern economic indicators from 1940-1970 is uncorrelated with northern destination city characteristics for migrants from those counties, I view estimating southern county net migration rates as a pure prediction problem. Belloni et al. (2011) propose a machine learning based estimation of the first stage in an instrumental variables context where the number of instruments is large relative to the number of observations. In my case, I use this approach to select predictors in the "zero" stage prediction of migration out of southern counties.[51]

In a "zero stage," I predict net migration from southern counties using southern push factors:

$$\text{mig rate}_{jt} \quad = \quad \beta_0 + Z'_{jt-10}\boldsymbol{\beta} + \varepsilon_{jt}, \tag{12}$$

for $t \in \{1950, 1960, 1970\}$ where m_{jt} is net migration for southern county j from the decade of $t-10$ to t and $\mathbf{Z}'_{jt-10}$ is the set of predictors measured in $t-10$. Using LASSO, I shrink the set of predictors to an optimal subset. Excluding a predictor from the subset corresponds to setting the respective element of $\boldsymbol{\beta}$ to zero. More explicitly, LASSO solves the following problem:

$$\min_{\beta_0,\boldsymbol{\beta}} \left\{ \frac{1}{N} \sum_{j=1,..,1223} \left(\text{mig rate}_{jt} - \beta_0 - Z'_{jt-10}\beta + \varepsilon_{jt}\right) \right\} \quad \text{subject to} \quad \sum_{k=1}^{9} |\beta_k| \leq p, \tag{13}$$

where p is the tuning parameter and β_k are the coefficients on each of the nine predictors in Z'_{jt-10} as suggested by Boustan (2010): the percent acreage in cotton; percent tenant farms; share of the labor force in agriculture; indi-

[51] See Sequeira et al. (2019) where the authors first predict European outmigration using local weather shocks and then interact predicted migration flows with railway expansion across US counties.

cator for being in tobacco growing state and the interaction between tobacco growing state and share in agriculture; WWII spending per capita; share of the labor force in mining, an indicator for being in a mining state (OK and TX), and the interaction between the two.

For each decade, I use five-fold cross-validation to choose the tuning parameter p that minimizes the expected prediction error.

In my case, LASSO selects the following for each year:

Variables selected in 1940:

- Percent tenant farms
- Share of the labor force in agriculture
- WWII spending per capita
- Percent acreage in cotton
- Share of the labor force in agriculture $\times$ Tobacco growing state
- Indicator for mining state
- Indicator for mining state $\times$ Share of the labor force in mining

Variables selected in 1950:

- Percent tenant farms
- Share of the labor force in agriculture
- WWII spending per capita
- Percent acreage in cotton
- Percent acreage in tobacco
- Indicator for mining state
- Indicator for mining state $\times$ Share of the labor force in mining
- Share of the labor force in mining

Variables selected in 1960:

- Percent tenant farms
- Share of the labor force in agriculture
- Indicator for tobacco growing state
- Share of the labor force in agriculture × Tobacco growing state
- Percent acreage in cotton
- Indicator for mining state
- Indicator for mining state × Share of the labor force in mining
- Share of the labor force in mining

Using LASSO-selected variables improves the F-statistic for county outmigration prediction from 1940-1950 from 11.56 to 14.78. The F-statistics in the models for county outmigration prediction from 1950-1960 and 1960-1970 are identical using the original set of variables in Boustan (2010) and the LASSO-selected set.[52]

Given this choice of included predictors, I estimate Equation (12) using OLS to predict net migration from county j, $\hat{m}_{jt}$, for each decade $t \in \{1950, 1960, 1970\}$.[53] Next, I generate predicted migration into northern city c, $\hat{m}_{ct}$, by multiplying the share of pre-1940 migrants from each county by the

[52]Chernozhukov et al. (2018) discuss inference adjustment in empirical settings where machine learning is used; they show that in a variety of empirical examples, qualitative conclusions of results remain unchanged after inference adjustment.

[53]Direct measures of county-level in-migration and out-migration is not available for this time period, so I use net migration estimates produced by Boustan (2010) and made available in Boustan (2016).

predicted number of migrants leaving that county between 1940 and 1970:[54]

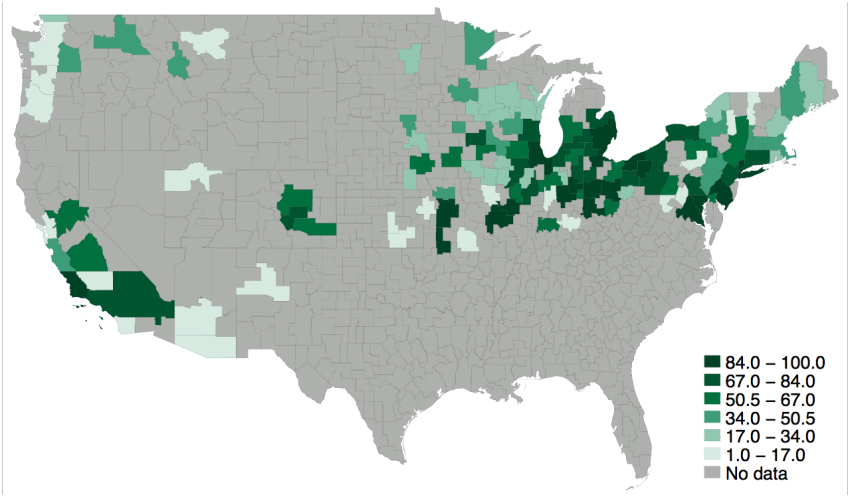
$$\hat{m}_{ct} = \sum_{j=1,..,1223} (\omega_{cj}^{1935-40} \cdot \hat{m}_{jt}) \tag{14}$$

where $\omega_{cj}^{1935-40}$ is the share of black migrants from southern county j living in city c. The estimated total black in-migration is calculated as $\hat{m}_c = \sum_{t \in \{1950,1960,1970\}} \hat{m}_{ct}$. Finally, I update the estimated share of black residents in city c, $\hat{b}_{c,t}$, as

$$\hat{b}_{c,t} = \hat{b}_{c,t-10} + \hat{m}_{c,t} \tag{15}$$

where $\hat{b}_{c,1940} = b_{c,1940}$ as observed in the data.

Figure B4: Map of Great Migration instrument



Notes: This map depicts Great Migration commuting zones and each CZ's predicted percentile change in the black population between 1940 and 1970, predicted using the methods described in Appendix B. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016).

[54]Because the available figures are net migration figures, and some southern counties experienced positive net migration (in-migration) as opposed to negative (out-migration), this procedure may result in predicted *decreases* in the black population. This is the case for a small share of the commuting zones in the sample, particularly those in western states that are more likely to be connected to counties in Oklahoma or Texas, for example, some of which experienced net in-migration between 1940 and 1970.

Appendix C Additional upward mobility results and robustness

This appendix provides additional results on upward mobility as well as further robustness checks on the main findings. I begin with descriptive analysis of the change in the geography of upward mobility between 1940 and 2015. I then provide additional results and supporting evidence on the impact of the Great Migration on upward mobility for recent cohorts.

C.1 Long run change, 1940-2015

In this section, I provide correlations between measures of educational upward mobility in 1940 with income upward mobility in 2015. Following a method similar to Card et al. (2018), I measure the fraction of teenagers from households in which the household head has 5-8 years of schooling[55] who obtain at least 9 years of education. The measure of income upward mobility in 2015 consists of estimated average adult income rank at the commuting zone level, for children from different parent income percentiles, where adult income is measured between the ages of 32 and 37.[56] Section 3.1 describes these data in much greater detail.

In Table C1, I report the correlation coefficients between historical and contemporary upward mobility measures separately by race and gender. For white men and women, historical educational upward mobility is positively correlated with income upward mobility across commuting zones today. However, for black men and women, these measures are virtually uncorrelated. This racial difference is particularly pronounced among men. Figures C1 shows the correlation between the historical measure and the contemporary measure for black men in the top panel and for white men in the bottom panel.

[55] Approximately the median of adult education in 1940.

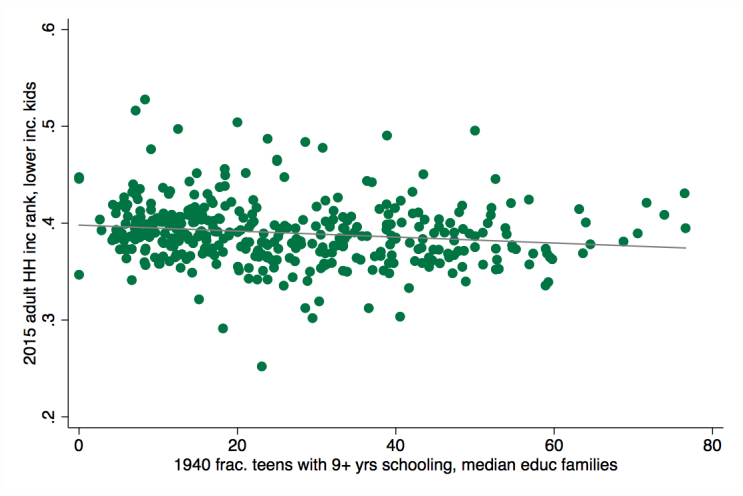
[56] The children come from 1980s birth cohorts (1978-1983).

Table C1: Correlation between historical and contemporary upward mobility measures, by race and gender

	Men	Women
Black	-.09	.11
White	.46	.43

Correlation coefficients between 1940 and 2015 measures of upward mobility, by race and gender. The sample in each column is the set of CZs within each gender for which both black and white upward mobility measures can be computed.

Figure C1: Correlation 1940 & 2015 Upward Mobility



(a) Black men

(b) White men

Notes: This figure depicts scatter plots of the relationship between historical upward mobility and contemporary upward mobility for black and white men. In panel (a), the right hand side ("RHS") is 1940 educational upward mobility defined as fraction of 14-18 year old black boys who have at least 9 years of schooling, from families where the household head has 5-8 years of education. The left hand side ("LHS") is expected average individual adult income rank based on 2014-2015 IRS tax returns of black men from 1978-1983 birth cohorts who come from families at the 25th percentile of the parent income distribution. Panel (b) shows the same relationship as in panel (a) for white men. In order to compare the same set of commuting zones and to minimize the influence of CZs with small numbers of black children, I restrict the sample of CZs in both panels to those with at least 10 14-17 year old black boys in 1940 and at least 10 black men in the IRS sample. The correlations between historical and contemporary upward mobility are reported for black and white women in Appendix Table C1. *Data sources*: IPUMS for 1940 measure; and Chetty, Hendren, Jones, and Porter (2018) for 2015 measures.

C.2 Supporting evidence and additional results

Table C2: Upward mobility results with coefficients on baseline controls

	Average	Expos. Effects	Black, p25	Black, p75
$\hat{GM}$	-0.0370	-0.00232	-0.0264	-0.0386
	(0.00974)	(0.000631)	(0.0114)	(0.0169)
Edu. Upward Mobility 1940	0.0163	-0.000572	0.00554	-0.0348
	(0.0391)	(0.00212)	(0.0457)	(0.0679)
Share of LF employed in manufacturing, 1940	-0.152	-0.00323	-0.0835	-0.00524
	(0.0271)	(0.00192)	(0.0317)	(0.0471)
Black Southern Mig 1935-1940	-4.312	-0.0820	-0.383	-2.014
	(1.446)	(0.0671)	(1.698)	(2.523)
Midwest	-0.536	0.0981	-1.449	-0.870
	(0.603)	(0.0365)	(0.705)	(1.048)
South	-2.004	0.167	-0.294	1.430
	(1.306)	(0.0758)	(1.527)	(2.270)
West	-2.682	-0.100	-1.575	-1.691
	(0.872)	(0.0459)	(1.028)	(1.528)
R-squared	0.481	0.249	0.185	0.110

Notes: Dependent variable is mean individual income rank, where income is measured from IRS tax returns for cohorts born between 1978 and 1983. Pooled income refers to household income, as opposed to individual income. Independent variable is predicted change in black population share between 1940 and 1970. Baseline controls include share of CZ population made up of 1935-1939 black southern migrants from any southern county, median education levels in 1940, and share of employment in manufacturing in 1940. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

TABLE C3: GREAT MIGRATION IMPACT ON AVERAGE UPWARD MOBILITY OF HIGH INCOME FAMILIES IN 2000S

	First Stage on GM					
$\hat{GM}$	0.297	0.297	0.297	0.297	0.297	0.297
	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)	(0.0759)
F-Stat	15.34	15.34	15.34	15.34	15.34	15.34
	Household Income Rank			Individual Income Rank		
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.0413	-0.0373	-0.0453	-0.0169	-0.00149	-0.0316
	(0.00793)	(0.00789)	(0.00828)	(0.00796)	(0.0102)	(0.00809)
R-squared	0.529	0.521	0.530	0.503	0.467	0.470
	Reduced Form					
$\hat{GM}$	-0.0161	-0.0137	-0.0184	-0.00839	-0.000228	-0.0165
	(0.00766)	(0.00752)	(0.00804)	(0.00717)	(0.00911)	(0.00748)
R-squared	0.445	0.448	0.438	0.490	0.467	0.426
	Two-stage least squares					
GM	-0.0541	-0.0462	-0.0618	-0.0283	-0.000768	-0.0556
	(0.0232)	(0.0230)	(0.0244)	(0.0233)	(0.0297)	(0.0243)
R-squared	0.519	0.516	0.515	0.494	0.467	0.432
N	130	130	130	130	130	130
Mean Rank	58.82	60.40	57.28	57.95	55.39	60.44
SD Rank	2.570	2.533	2.684	2.510	3.118	2.470
SD GM	28.98	28.98	28.98	28.98	28.98	28.98

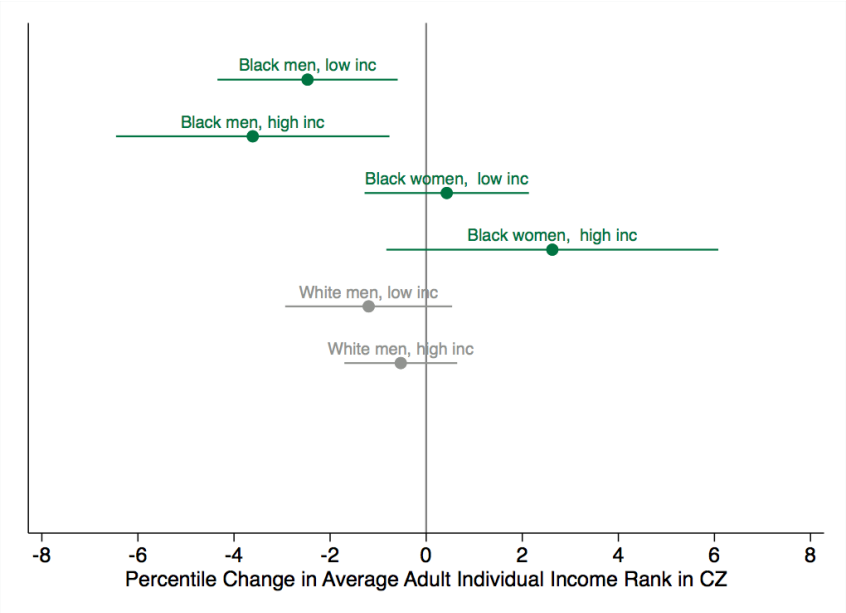
Notes: This table reports the estimated impact of the Great Migration on average upward mobility in the 2000s for men and women with high income parents. The unit of observation is a commuting zone. Dependent variable is expected mean individual or household income rank for individuals with parents at the 75th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. Pooled income refers to household income, pooling across men and women. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Table C4: Great Migration impact on childhood exposure effects in the 2000s for high income families

	First Stage on GM					
$\hat{GM}$	0.274	0.273	0.274	0.273	0.274	0.274
	(0.0618)	(0.0617)	(0.0619)	(0.0618)	(0.0619)	(0.0618)
F-Stat	19.68	19.53	19.65	19.56	19.63	19.71
	Household Income Rank			Individual Income Rank		
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.00119	-0.0000847	-0.00267	-0.000736	0.00134	-0.00278
	(0.000829)	(0.00115)	(0.00111)	(0.000804)	(0.00121)	(0.00113)
R-squared	0.305	0.233	0.154	0.472	0.388	0.204
	Reduced Form					
$\hat{GM}$	-0.00131	-0.00104	-0.00153	-0.00203	-0.00171	-0.00241
	(0.000604)	(0.000839)	(0.000827)	(0.000564)	(0.000880)	(0.000823)
R-squared	0.320	0.243	0.138	0.520	0.400	0.219
	Two-stage least squares					
GM	-0.00479	-0.00382	-0.00557	-0.00742	-0.00622	-0.00877
	(0.00232)	(0.00313)	(0.00297)	(0.00262)	(0.00361)	(0.00325)
R-squared	0.197	0.167	0.107	0.173	0.190	0.0211
N	130	130	130	130	130	130
Precision Wt	Y	Y	Y	Y	Y	Y
Mean Expos FX	-0.00323	-0.0253	-0.0162	0.0305	0.0182	-0.00525
SD Expos FX	0.175	0.228	0.212	0.195	0.270	0.222
SD GM	24.40	24.08	24.29	24.52	24.33	24.38

Notes: This table reports the estimated impact of the Great Migration on commuting zone childhood exposure effects. The unit of observation is a commuting zone. Dependent variable is the estimated causal impact of one additional year of childhood in the commuting zone on adult household income rank for men and women with parents at the 75th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. The unit of observation is a commuting zone. Pooled income refers to household income, pooling across men and women. 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 share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Figure C2: Race and gender heterogeneity in impact of Great Migration on upward mobility, individual earnings



Notes: This figure plots coefficients from regressions of average upward mobility in the 2000s for men and women from low and high income parents on the instrument for black population increases during the Great Migration, in approximately one standard deviation units. The unit of observation is a commuting zone. Upward mobility is defined as expected mean individual income rank where income is measured from IRS tax returns for cohorts born between 1978 and 1983. Pooled income refers to mean household income rank, pooling across men and women. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); Chetty et al. (2018).

C.3 Supplementary info. on childhood exposure effects

Chetty and Hendren (2018b) use variation in age of child at time of family's move to purge place effect estimates of bias from sorting on family unobservables, θ_i:

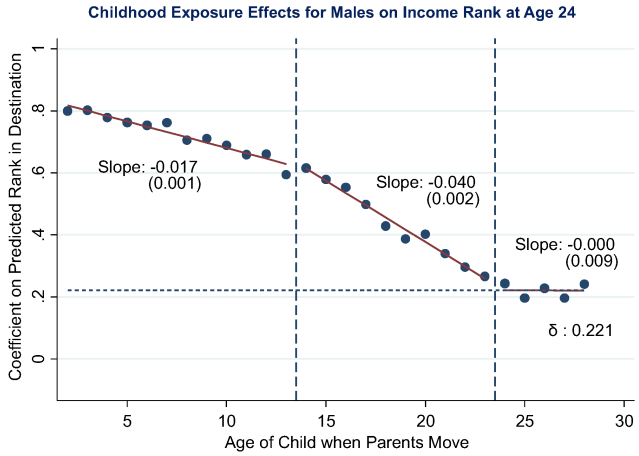
$$\begin{aligned} y_i &= \delta_c + \theta_i \\ &\downarrow \\ \Delta y_i &= \alpha_c \Delta t_i \end{aligned}$$

α_c is an unbiased estimate of the effect of additional year of childhood exposure to location c on adult outcome y_i.

Scaling childhood exposure effects Assuming muted effects for early years according to Figure C3, the effect of full childhood exposure for 23 years should be adjusted in the following manner:

$$\text{Years} = (23 - 13) + (17/40) * 13 = 15.525$$

FIGURE C3: HETEROGENEITY IN CHILDHOOD EXPOSURE EFFECTS BY AGE OF CHILD (CHETTY ET AL., 2018)



Notes: This image from Chetty et al. (2018) depicts heterogeneity in childhood exposure effects by age of exposure. Early years of childhood exposure have more muted impacts compared to teen years of exposure.

C.4 Net effect of the Great Migration

This appendix discusses the overall impact of the Great Migration on black economic status over the 20th century, through the lens of intergenerational mobility. The main analysis in the paper poses the counterfactual of upward mobility for children in the northern US had they grown up in locations less affected by the Great Migration. This counterfactual does not consider the impact of the Migration on earlier generations, which affects the adult income of black children today through their parents and grandparents, or on southern stayers, who may have been affected by black *emigration* from South.

The Great Migration moved black grandparents dramatically up in the national income distribution: estimates suggest that migrants could approximately double their earnings by moving North (Collins and Wanamaker, 2014; Boustan, 2016). At the same time, racially segmented labor markets in the North led to increased competition between black incumbents and new arrivals

such that racial earnings convergence in the destinations slowed (Boustan, 2009). Evidence on the timing of changes in conditions in northern cities, presented in Appendix E, suggests that the cohorts growing up in the 1970s would have been exposed to negative environmental factors including extreme segregation, high crime rates, and spillovers from greater greater police presence. Nonetheless, the sharp increase in average grandparent income through migration likely outweighed the competition effect in the North, and potentially even the harsher environment faced by the second generation.

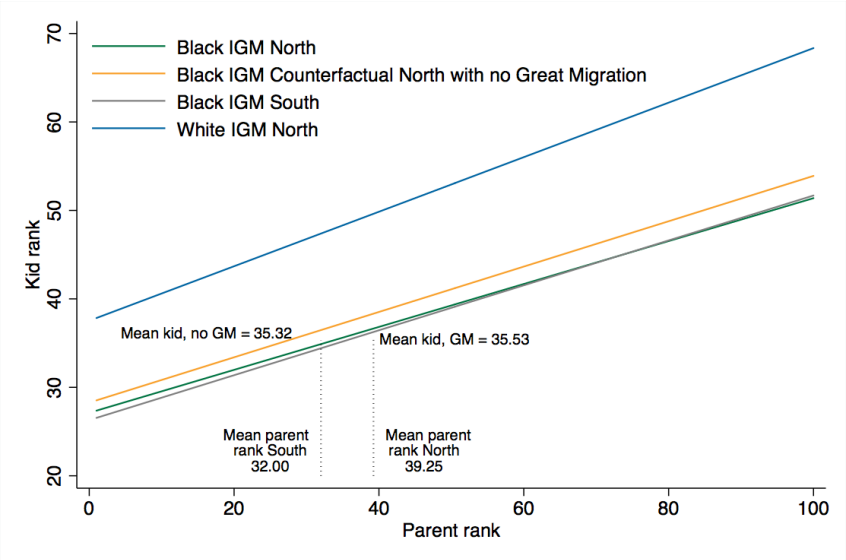
A final relevant factor for understanding the effect of the Migration is the impact black *emigration* on southern locations. For the Great Migration to have had a net *negative* impact on black economic status, it would be necessary to assume that in the absence of black emigration, southern locations would have been better off. There are two key reasons why this is unlikely to be the case. First, emigration put direct pressure on southern jurisdictions to offer better amenities for black workers. Boustan and Tabellini (2018) find that votes for segregationist policies decreased in places where black migrants left in greater numbers. This echoes the "voting with one's feet" hypothesis explored by Margo (1991b). Second, Calderon et al. (2019) find that the Great Migration may have played a role in bringing Civil Rights issues to the national stage and helped civil rights legislation get passed. The effects of Civil Rights legislation were felt more strongly in the South than in the North, so this suggests another mechanism through which the Migration may have improved southern conditions.

In a simplified counterfactual exercise, I explore the aggregate effect of the Great Migration while making several conservative assumptions. First, I assume a zero effect of the Great Migration on the South and that the net effect of the Great Migration on parent income (inclusive of the effects on grandparents) is reflected in the difference in average black parent income rank in the North and the South today. In the absence of the Great Migration, 23% of black grandchildren would experience the counterfactual northern intergenerational mobility curve had the Migration not occurred while 77% would remain

on the southern curve. This exercise suggests a positive net effect of the Great Migration on black income of 0.2 income percentiles.

I conclude that while the Migration eventually reduced the gains to parent income for black children in the North, the large positive effect the Migration had on the income of earlier generations (moving black children *up* the IGM curve) makes up for these losses. Any additional positive impacts on the South would only magnify a positive net effect of the Great Migration.

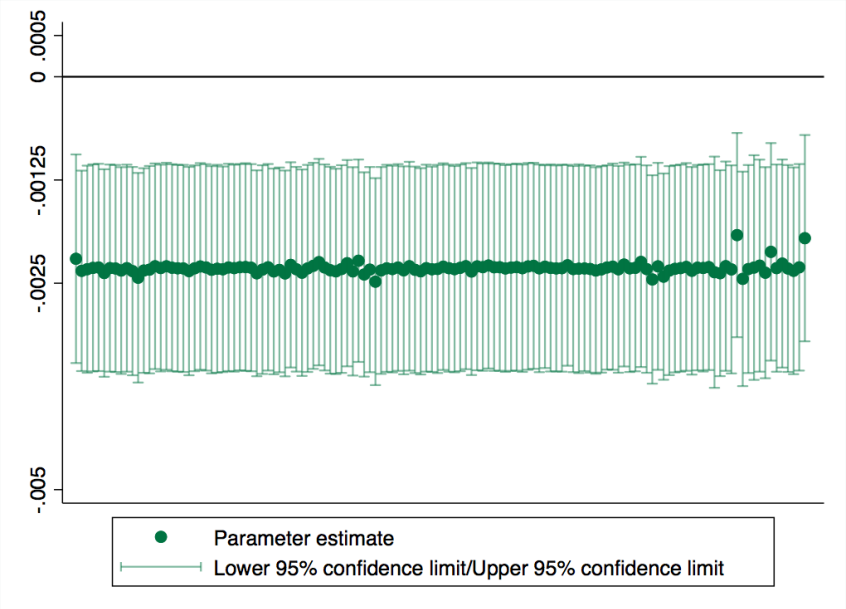
Figure C4: Intergen. mobility by race and region



Notes: This figure plots intergenerational mobility curves by race and region. The y-axis plots the income rank of individuals from the 1978-1983 birth cohorts and the x-axis plots the income rank of their parents. Income is measured from IRS tax returns. The green line plots the intergenerational mobility curve for black families in the North; the gray line plots the intergenerational mobility curve for black families in the South; and the gold line plots a counterfactual intergenerational mobility curve for black families in the North in the absence of the Great Migration. Average parent income rank in the North and South are indicated on the plot. The counterfactual line is plotted using estimates of the Migration's impact on black men from the 25th, 50th, and 75th percentiles of the parent income distribution from regressions described in Section 4. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty et al. (2018).

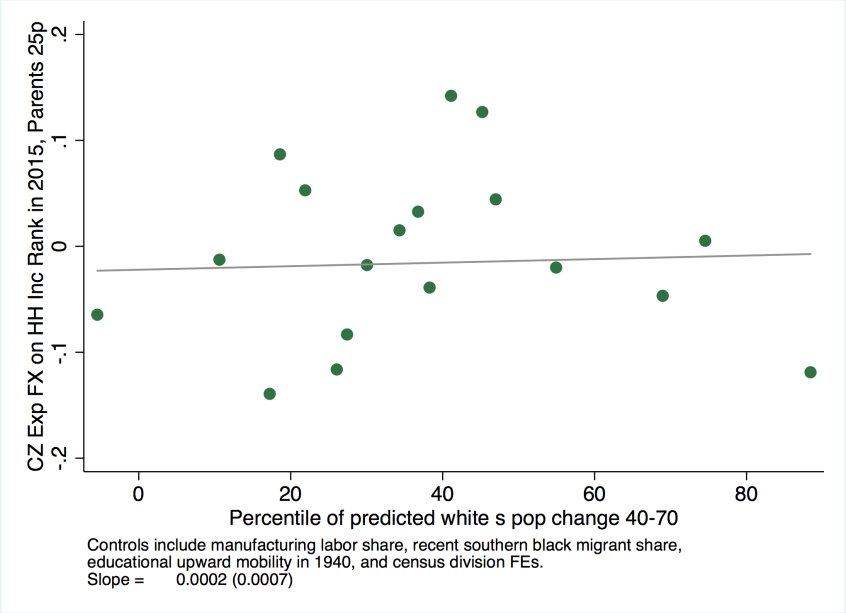
C.5 Additional robustness checks

Figure C5: Great Migration effect robust to leaving out each CZ once from sample



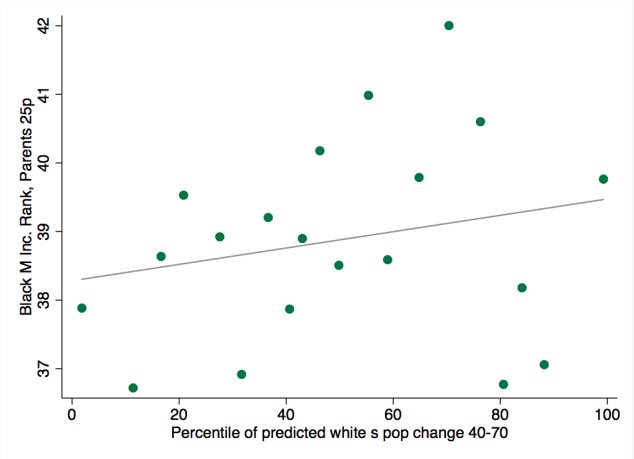
Notes: This figure plots the coefficient on percentile of predicted black population change in 130 separate regressions where each CZ in the sample has been left out of the regression once. 95% confidence intervals indicated. The unit of observation is a commuting zone. Dependent variable is the estimated causal impact of one additional year of childhood in the commuting zone on adult household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: CCDB; IPUMS complete count 1940 US census; Boustan (2016); Chetty and Hendren (2018b).

Figure C6: White southern migration impact on childhood exposure effects

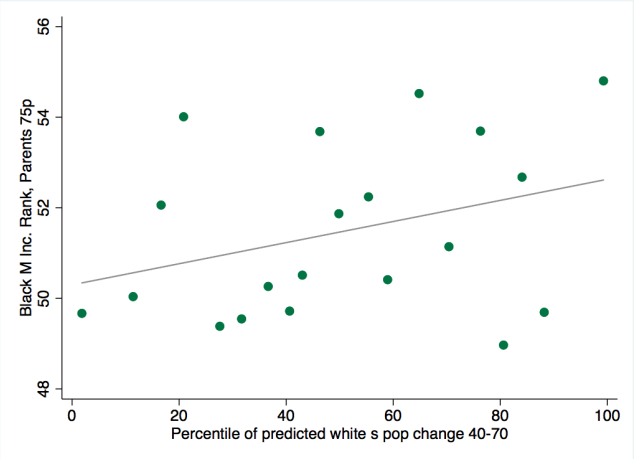


Notes: This figure depicts a binned scatterplot of the relationship between the percentile of predicted white southern in-migration and CZ childhood exposure effects for individuals from low income families. The unit of observation is a commuting zone. The right hand side variable is grouped into 20 bins (5 percentiles each). Both right hand side and left hand side variables have been residualized on the following controls from 1940: share of urban population made up of 1935-1940 black southern migrants, the share of labor force in manufacturing, census division fixed effects, and quartiles of the black population share in 1940. *Data sources*: Chetty, Hendren, Jones, and Porter (2018); IPUMS Complete Count 1940 US Census; Boustan (2016).

Figure C7: White southern migration impact on black men's upward mobility



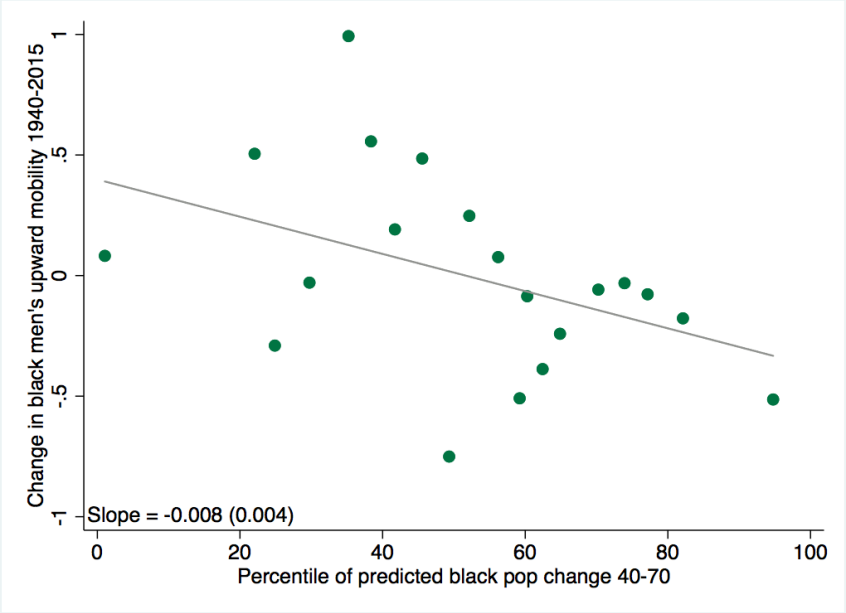
(a) Parents 25th percentile



(b) Parents 75th percentile

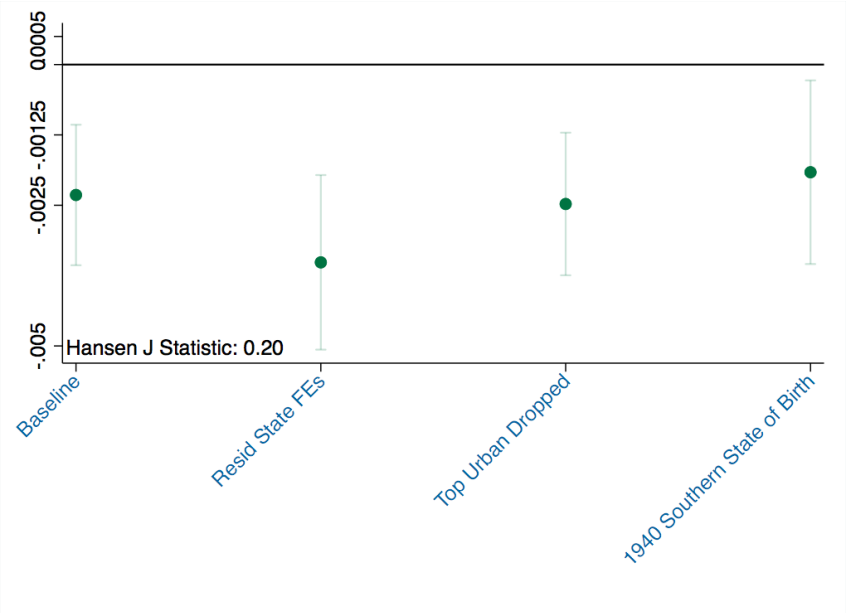
Notes: Panel (a) depicts a binned scatterplot of the relationship between the percentile of predicted white southern in-migration and black men's upward mobility (25th percentile of parent income distribution). Panel (b) depicts the same for black men from the 75th percentile of parent income distribution. The right hand side variable is grouped into 20 bins (5 percentiles each). The unit of observation is a commuting zone. Units of outcome variables are standard deviations. Upward mobility is defined as mean individual or household income rank by childhood commuting zone where income is measured from IRS tax returns for cohorts born between 1978 and 1983. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Chetty, Hendren, Jones, and Porter (2018); IPUMS Complete Count 1940 US Census; Boustan (2016).

Figure C8: Great Migration impact on change in black men's upward mobility, 1940-2015



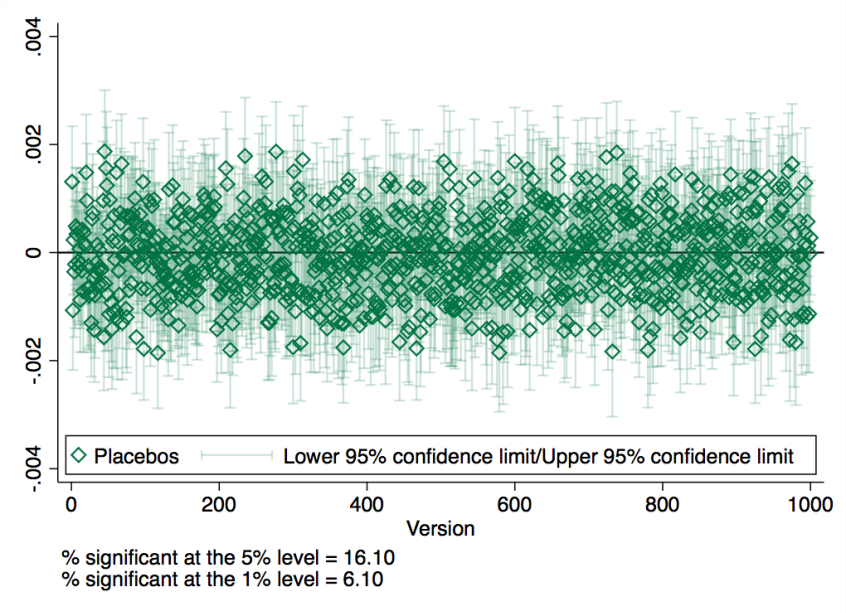
Notes: This figure depicts a binned scatterplot of the relationship between the percentile of predicted black population increase and the change in black men's upward mobility between 1940 and 2015. The unit of observation is a commuting zone. To construct the outcome variable, I take the difference in the Z-score of black male income upward mobility in 2015 (for men from parents at the median of the national parent income distribution) and the Z-score of black male educational upward mobility in 1940 (for boys whose parents had 5-8 years of schooling, the national median for adults). I then standardize this difference, so that the units of outcome variables are standard deviations. The right hand side variable is grouped into 20 bins (5 percentiles each). Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, the share of labor force in manufacturing, and census division fixed effects. *Data sources*: Chetty, Hendren, Jones, and Porter (2018); IPUMS Complete Count 1940 US Census; Boustan (2016).

Figure C9: Alternative instruments for the Great Migration



Notes: This figure plots the coefficient on alternative instruments for black population increases during the Great Migration, where the dependent variable is commuting zone childhood exposure effects in the 2000s for men and women with low income parents. The unit of observation is a commuting zone. Childhood exposure effects are the estimated causal impact of one additional year of childhood in the commuting zone on adult household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. The instrument is the percentile of predicted black population increase, defined in versions 1-3 as the interaction between pre-1940 black southern migration patterns and 1) post-1940 southern county net-migration as predicted by local economic factors alone; 2) southern county net-migration residualized on state fixed effects; and 3) southern county net-migration from less urban counties (dropping the counties coded as central city with populations of 1 million or more in 1990 – 42 total). In version 4, predicted southern county outflows between 1940 and 1970 are aggregated to the state level and assigned to northern cities according to the share of the black population born in that southern state and living in the destination city in 1940. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Chetty, Hendren, Jones, and Porter (2018); IPUMS Complete Count 1940 US Census; Boustan (2016).

Figure C10: Placebo migration shocks



Notes: This figure plots the coefficient on placebo shocks in 1,000 separate regressions, where the dependent variable is commuting zone childhood exposure effects in the 2000s for men and women with low income parents. The unit of observation is a commuting zone. Childhood exposure effects are the estimated causal impact of one additional year of childhood in the commuting zone on adult household income rank for men and women with parents at the 25th percentile of the parent income distribution. Income is measured from IRS tax returns for cohorts and parents of cohorts born between 1980 and 1986. The placebo shock is the percentile placebo increase in the black population, defined as the interaction between pre-1940 black southern migration patterns and a normally distributed random variable with mean 0 and variance 5. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Chetty, Hendren, Jones, and Porter (2018); IPUMS Complete Count 1940 US Census; Boustan (2016).

Appendix D Public Finance and Neighborhoods Database, 1920-2015

D.1 Data sources and key measures

Private school enrollment rates

Data on private school enrollments come from two different sources depending on the time period. For pre-1940 statistics on private school enrollment, I use tabulations on city school systems from the 1922 Biennial Survey of Education report. This report contains the total number of elementary and high school students enrolled in private schools in that city as well as total enrollment in the city.

For 1970 onwards, I use county-level counts of private school enrollments from IPUMS National Historical Geographic Information System ("NHGIS"), which I aggregate up to the CZ level. Starting in 1970 through 2010, enrollment is also reported separately for elementary and high school students and separately by race from 1970 to 2000.

Incarceration rates

For 1920 and 1930, I use the complete count censuses to construct the percent of the population in a county that is incarcerated in jails or local correctional institutions. I do not include the federal or state prison in these estimates as it is not possible to allocate state and federal prisoners back to localities they came from. For 1940, I digitized data from a census report on the incarcerated population. For 1960, I digitized data from the 1960 US Census publication, which includes a table on the incarcerated population and reports the non-white and white incarcerated population by county separately.

For the post 1970 period, I use a rich new dataset from the Vera Institute

of Justice In Our Backyards Symposium ("IOB"), which provides counts of federal and state prisoners by their county-of-commitment to federal and state prison. These data begin in the year 1983. These figures are available separately by race. Due to reliability issues for the local jail population in these data, I focus on total jail rates rather than jail population breakdowns by race.

Crime rates

For crime rates, I focus on murder rates as these are less subject to reporting bias than other crime categories, such as property crime or non-fatal violent crimes. I digitize murder rates for cities with a population of 25,000 or more from the Uniform Crime Repors publications ("UCR") of the FBI in 1931, 1943, and 1950.[57] For the years 1958 to 1969, I use city-level tabulations of murder rates from UCR available from ICPSR. Finally, for the post 1970 period, I use county-level tabulations of UCR murder rates available from the IOB database.

In addition to looking at crime rates as a measure of neighborhood quality, I also use data on the intensity and duration of race riots in major cities in the 1960s.[58]

Local government expenditures

Data on local government expenditures come from surveys of state and substate level governments conducted by the US Census Bureau. For each expenditure category, I focus on aggregate spending by various local governments in a county area. The advantage to this approach is that changes in which levels

[57]Some large cities did not report to the FBI UCR series in these years. A notable case is New York City in 1931 and in 1950. For these cities in 1931, I supplement using data generously shared by Price Fishback. I drop 1950 from the analysis due to missing data from New York City.

[58]These data were generously shared by William Collins and Robert Margo. See Collins and Margo (2007).

or types of government are responsible for providing a certain public good will not affect this measure of spending. The first full set of such data are available in the 1932 publication of *Financial statistics of states and local governments*. I digitize county aggregate and individual local government expenditures from this report.

For post-migration years, I rely on the US Census Bureau Annual Survey of Local Governments, which provides individual government expenditure data in digital format for roughly 15,000 local governments across the United States from 1967 to 2012. I also include data on city government expenditures available for intermittent years from 1948 to 1975 from the City Data Books available from ICPSR, and for the year 1962, I include data available on different expenditure categories from the County Data Book also available from ICPSR. In the case of police expenditures, I supplement these two measures with counts of police officers per capita using the complete count censuses available from IPUMS for the years 1920, 1930, and 1940 and US Census Bureau data surveying public sector employment in cities from 1951-2007.

For each data set, I construct commuting zone area aggregate expenditures for the expenditure categories of interest. I focus on expenditures per capita (or per student), and the share of total expenditures devoted to that expenditure category.

For example, for police spending, CZ-area local government expenditure share is defined as

$$\text{Pol. Exp. Share}_{CZ} = \frac{\text{\$Spent on Police by All Local Governments}_{CZ}}{\text{\$Spent by All Local Governments}_{CZ}}$$

and per capita expenditures at the CZ-area level are defined as

$$\text{Per Cap Pol. Exp.}_{CZ} = \frac{\text{\$Spent on Police by All Local Governments}}{\text{Population}_{CZ}}$$

Finally, I focus on categories of expenditures over which local governments

have a large degree of discretion: police expenditures, education expenditures, and fire expenditures. Table D1 shows the the contribution of different levels of government (e.g., federal, state, county, etc.) to direct expenditures for each category of government spending.

Table D1: Expenditure by government type by spending category

Govt Type	Revenue	Rev (Own)	Dir Exp	Elem + HS	Fire Prot	Health + Hosp	Highway	Parks & Rec	Police	Pub. Welf	Sanitation
Fed	18.93%	24.90%	22.95%	0.00%	0.00%	6.12%	0.36%	3.24%	3.73%	6.27%	0.82%
State	42.92%	42.76%	34.17%	1.14%	0.00%	42.77%	60.37%	14.44%	13.31%	79.42%	5.13%
County	8.73%	7.40%	9.51%	7.84%	13.83%	26.71%	14.76%	16.76%	23.77%	10.33%	16.03%
Muni	12.45%	12.72%	13.90%	8.31%	67.90%	10.64%	19.41%	51.29%	54.80%	3.78%	56.82%
Town	1.19%	1.22%	1.31%	2.24%	6.29%	0.25%	3.79%	3.75%	4.36%	0.09%	5.40%
Spec. Dist.	4.15%	4.13%	4.79%	0.24%	11.99%	13.51%	1.32%	10.52%	0.03%	0.11%	15.81%
School Dist.	11.64%	6.87%	13.38%	80.23%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%	0.00%
Total	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%	100.00%

Notes: This table shows the breakdown in spending by government type for different spending categories, averaged across all reporting years from 1967-2012. Column 1 lists the government types: federal, state, county, municipality, town, special district, and school district. Column 2 shows the total breakdown of government revenue by government type. Column 3 shows this number net of intergovernmental transfers. Column 4 shows total direct government expenditures by government type. Starting with Column 5, the categories of spending from left to right are education for elementary and high school districts; fire protection services; health and hospitals; highways; parks and recreation; police protection; public welfare; and sanitation. Sanitation spending includes sewage and waste management. *Source*: US Census Bureau Annual Survey of Local Governments (1967-2012).

Appendix E Additional results on local mechanisms

E.1 Impact on private schooling and residential segregation

In this section, I report additional results on private schooling and residential segregation. Figure E3 plots the coefficients on predicted black population increases on standardized measures of private school enrollment rates separately for each year that data are available. The outcome variables is the share of elementary and high school students enrolled in private school. Beginning in 1970, these measures are available separately by race. I find no impact of the 1940-1970 Migration shock on private school enrollment rates in 1920. In 1970, the next year that data are available,[59] a 1-standard-deviation increase (approximately 30 percentiles) in the Great Migration shock is associated with a 0.2 standard deviation increase in white private school enrollment rates and a 0.2 standard deviation decrease in black private school enrollment rates. Individually, these results are not statistically significantly different from zero. However, the black-white gap in public school enrollment is significantly larger in Great Migration CZs.

Consistent with Boustan (2010) and Tabellini (2018), I find that black population increases also predicts large declines in the urban white share at the commuting zone level. These results are shown in Appendix Figure E4.

[59]Starting in 1960, the Census began asking about the type of school households enrolled their children in; however aggregate statistics for children attending high school as well as breakdowns by race are only available through the *National Historical Geographic Information System* ("NHGIS") data consortium until 1970. See Appendix D for more details.

E.2 Impact on local government expenditures

Next, I examine the impact of the Migration on the public spending patterns of local governments. I focus on categories of public expenditures over which sub-state governments have a large degree of discretion. Appendix Table D1 shows the contribution of different levels of government to each of several main categories of public expenditures. I focus on two categories in particular, police and school expenditures. Spending on police indicates levels of neighborhood safety and crime, but also may have direct effects on the outcomes of black male youth in particular, which I discuss further in the main text (see Section 6). School spending has natural implications for the average outcomes of children in a given location.

Appendix Figure E5 plots the coefficients on predicted black population increases on standardized measures of police investments separately for each year that the data are available. The outcome variables are police expenditures per capita, the share of local government expenditures on police, and police officers per capita. As can be seen in the Figure, the Migration had no statistically significant or large effects on pre-period police investments from 1920-1940. Starting after 1940, the association between the Migration and police spending increases, peaking in the late 1970s and persisting for several decades after. At the peak of the association between the Migration and police investments, a 1 standard deviation increase in the Migration shock increased the police expenditure share and police expenditure per capita by just over 0.2 standard deviations.

I then look at the impact of the Great Migration on educational investments in affected commuting zones. These investments include direct educational expenditures by school districts, both as a share of all local government expenditures in commuting zones and per pupil. Figures E6 and E7 report these results. I estimate a noisy negative association between the Migration on pre-period (1932) aggregate educational expenditure shares. In E7, I control for 1932 educational expenditure shares and estimate the Migration's impact

on post-1970 educational investments. I find no impact of the Migration on aggregate education expenditures at the commuting zone level in the post-Migration period. I discuss the implications of these findings in the main text (see Section 6).

To check whether the effect of the Migration on police expenditures is simply driven increases in municipal spending in Great Migration destinations, I estimate the impact of the shock on fire fighting expenditures. Figure E8 reports these results. I find no impact of the Migration on fire-fighting expenditures. Higher police expenditures may be associated with higher crime and incarceration rates. I investigate these below.

E.3 Impact on incarceration rates

Appendix Figure E9 plots the coefficients on predicted black population increases on standardized measures of incarceration separately for each year. The outcome variables are the local correctional institution population per 100,000, the non-white local correctional institution population per 100,000 of the non-white population, and the state and federal imprisoned population by commuting-zone-of-commitment per 100,000, for all individuals aged 15-64 and then separately for this group by race. As can be seen in the Figure, the Migration had no statistically significant effects on pre-period incarceration. The Migration is most strongly associated with incarceration in the 1980s and 1990s, during the rise of incarceration rates nationally.

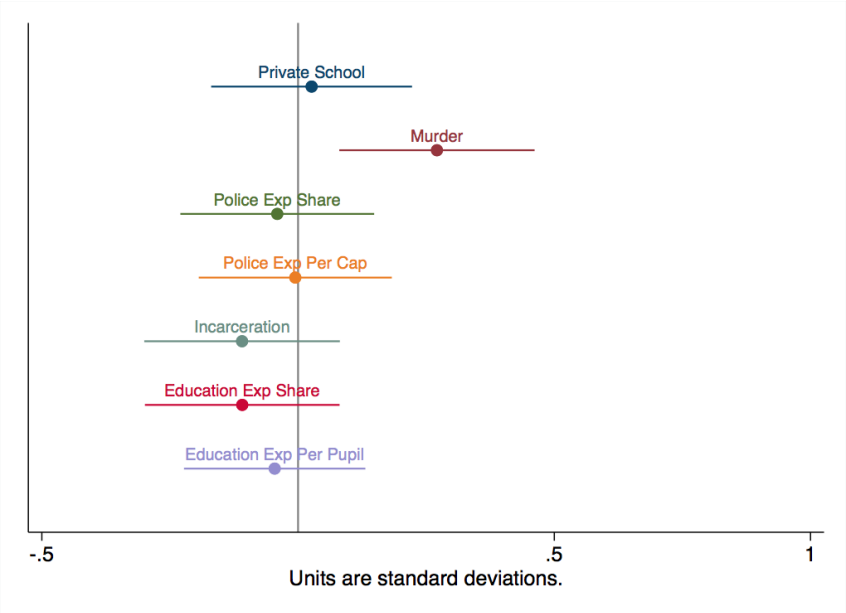
In Figure E10, I report the impact of the Migration on the incarceration rate in levels. At the peak of the association between the Great Migration and black incarceration rates, in 1992, a 30-percentile increase in predicted black population increases was associated with 300 more black people per 100,000 being committed to federal and state prison. The impact for whites was an increase of approximately 30 per 100,000.

E.4 Impact on murder rates

Appendix Figure E11 shows the impact of the migration on standardized measures of murder rates between 1931 and 2015. A 30-percentile increase in the Great Migration shock is associated with 0.3 standard deviations higher murder rates in 1931, before the period of black population change predicted by the shock, but is not associated with higher murder rates in 1936 or 1943. Murder rates are not significantly associated with the Migration again until the late 1960s. In the post-1970 period, a 30-percentile increase in the migration shock is associated with a .5 standard deviation increase in murder rates. Controlling for the 1931 murder rate attenuates some of the impact of the Migration on post-1970 murder rates, but the effect on late 1960s murder rates remains positive and statistically significant.

The late 1960s coincided with increases in the murder rate in cities across the US. At the same time, race riots erupted in urban areas as well. I explore whether the Migration affected the intensity of these riots. Table E1 reports these results. I find that Great Migration destination cities experienced longer riots and that riots in these areas involved more deaths, injuries, and arrests than places with fewer black migrant inflows. The magnitude of the effect of the Migration on arrests is large: a 30-percentile increase in the Migration shock is associated with over 30 more arrests per 100,000 during the 1960s riots. Both of these events may have contributed to rising police investments during this period. Both the impact on police expenditures and incarceration rates appear to have persisted for several decades afterwards.

Figure E1: Great Migration correlation with pre-period measures of local mechanisms



Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions. The dependent variables are standardized 1920 private school enrollment rates; mean 1931-1943 urban murders per 100,000 of the urban population; mean 1920-1940 local jail rate per 100,000; and mean government expenditure shares and per capita or per pupil spending. The unit of observation is a commuting zone. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); see Appendix D for the full list of data sources on each of the mechanisms.

Figure E4: Great Migration impact on urban white share

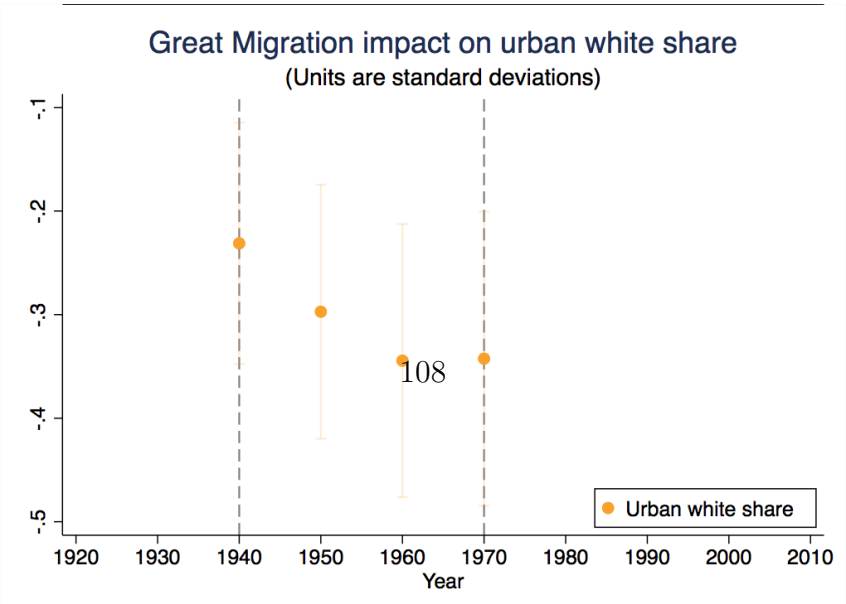
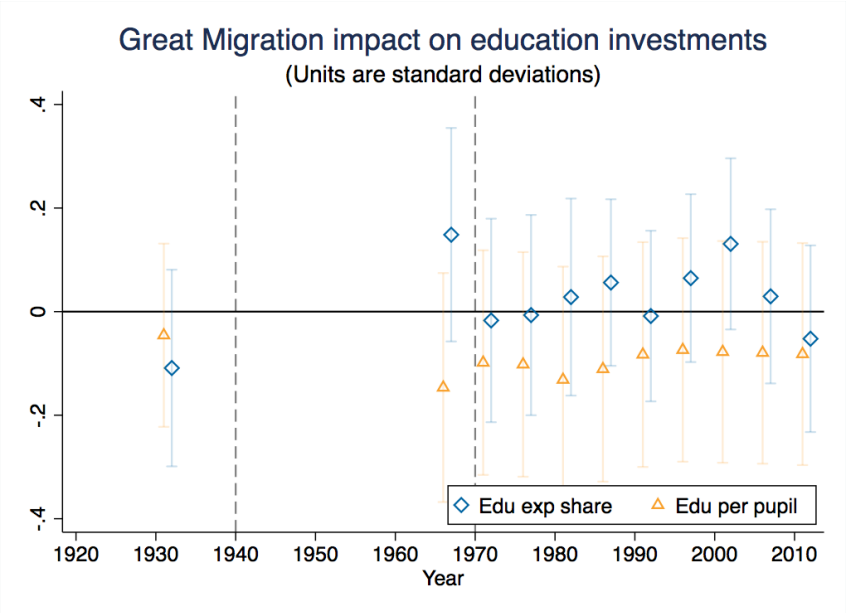
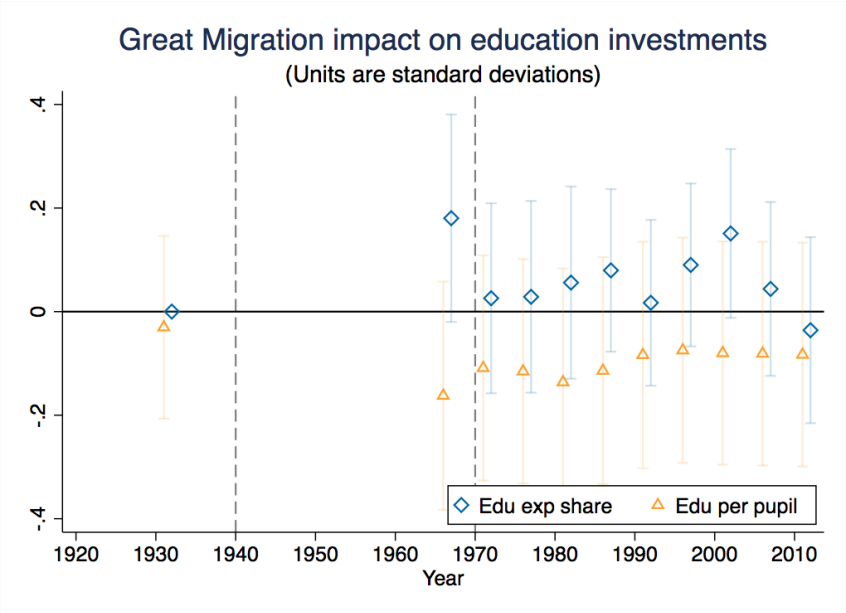


Figure E6: Great Migration impact on schooling investments



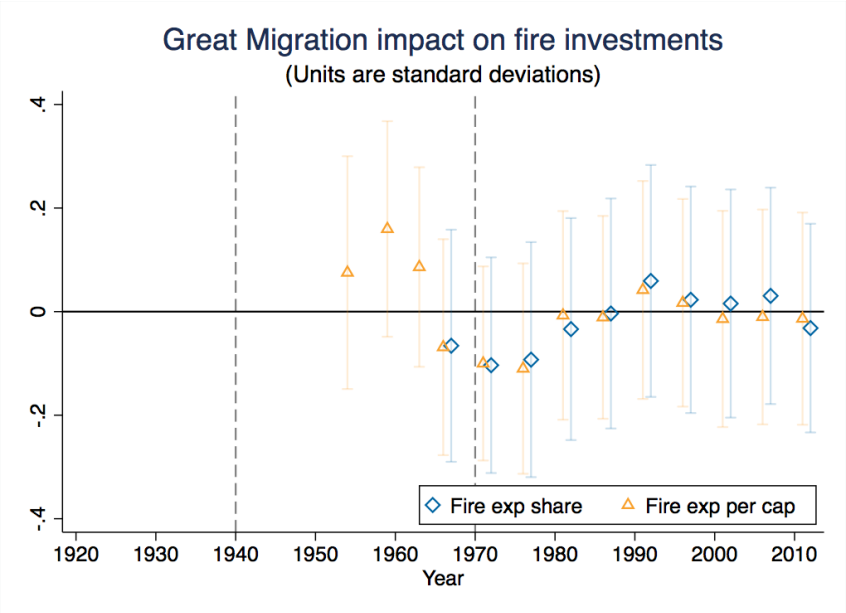
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is either the share of local government expenditures on education or education expenditures per student. Education expenditure data are for elementary and high school districts. The unit of observation is a commuting zone. Units of outcome variables are standard deviations. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Data on education expenditure shares and per student spending come from Financial statistics of states and local governments, 1932; US Census Bureau Annual Survey of Local Governments (1967-2012).

Figure E7: Great Migration impact on schooling investments, with pre-period control



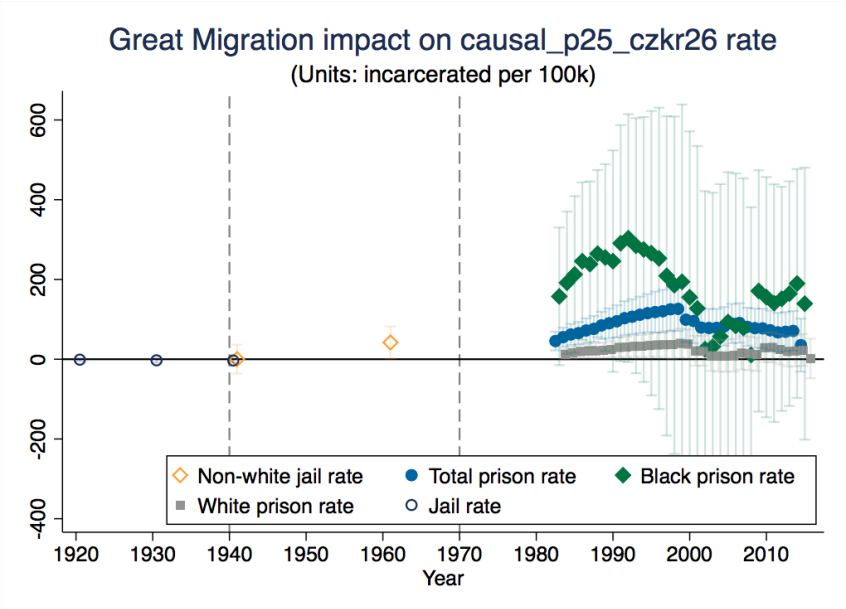
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is either the share of local government expenditures on education or education expenditures per student. Education expenditure data are for elementary and high school districts. All regressions include controls for the 1932 share of local government expenditures on education. The unit of observation is a commuting zone. Units of outcome variables are standard deviations. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Data on education expenditure shares and per student spending come from Financial statistics of states and local governments, 1932; US Census Bureau Annual Survey of Local Governments (1967-2012).

Figure E8: Great Migration impact on fire-fighting investments



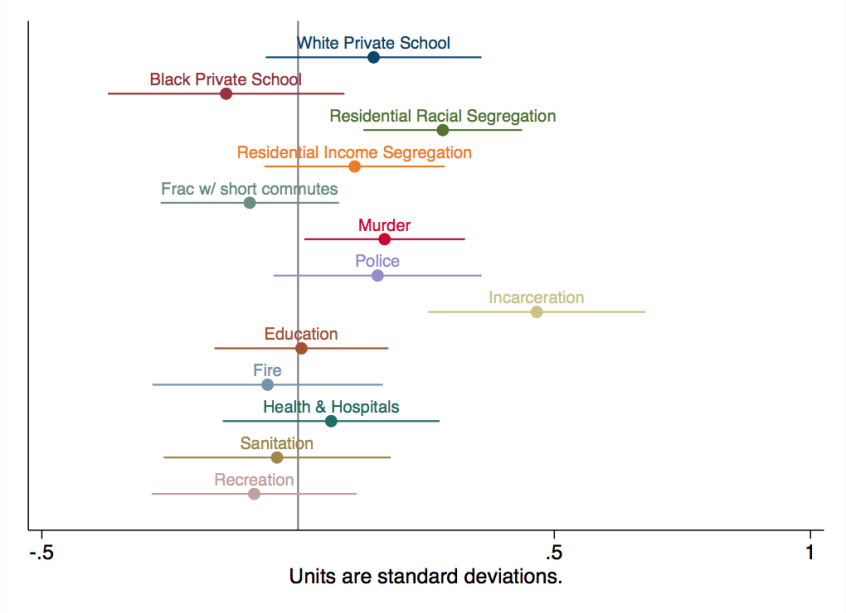
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is either the share of local government expenditures on fire-fighting or fire-fighting expenditures per student. The unit of observation is a commuting zone. Units of outcome variables are standard deviations. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: US Census Bureau Annual Survey of Local Governments (1967-2012).

Figure E10: Great Migration impact on incarceration rates, levels



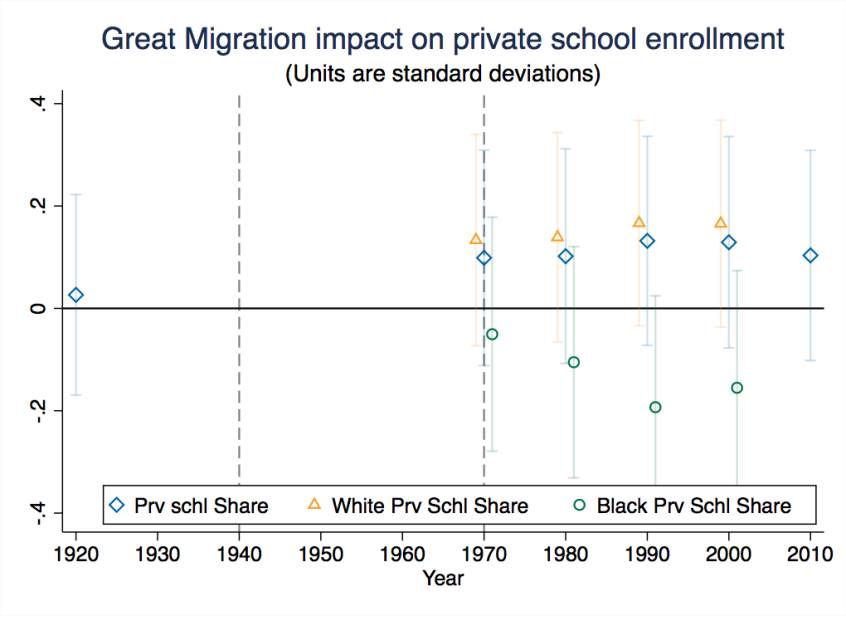
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is county jail population per 100,000 (1940 and 1960) or federal and state prison population by 100,000 by county-of-commitment from 1983-2015. Each jail or prison population group is normalized by the population for that group. Federal and state prison rates are for black and white men aged 15-64. The unit of observation is a commuting zone. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: 1940 county jail rates come from US IPUMS complete count US census; 1960 county jail rates come from 1960 Census report on county correctional institution population; data on 1983-2015 federal and state prison population by county-of-commitment come from Vera Institute of Justice In Our Backyards Database.

Figure E2: Impact of Great Migration on standardized post-1970 mechanisms, controlling for pre-period murder rates



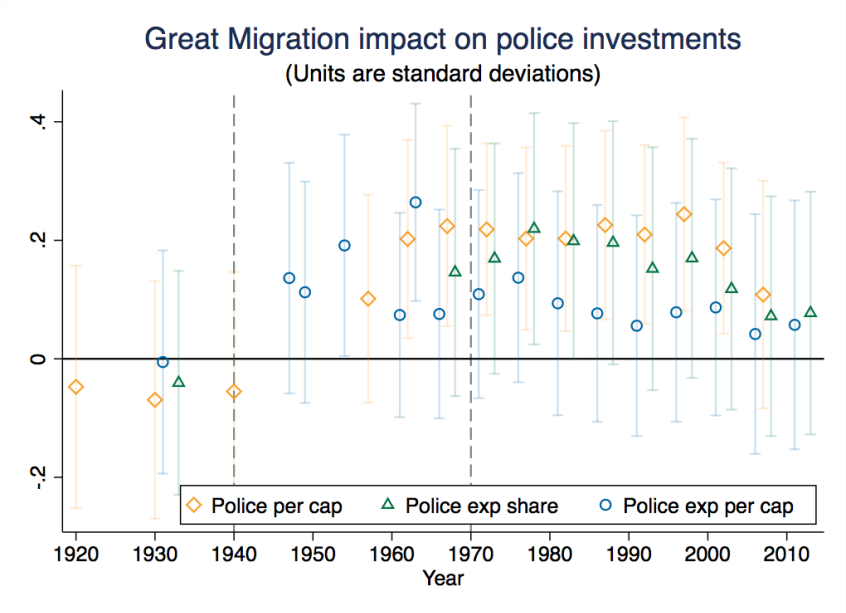
Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions. The dependent variables are standardized mean 1970-2000 white and black private school enrollment rates; the Theil indices in residential racial and income segregation in 2000; the fraction of families in 2000 with commute times less than 15 minutes; mean 1977-2002 murders per 100,000 of the population; mean 1983-2000 incarcerated per 100,000 of the population; and mean 1972-2002 government expenditure shares by category. The unit of observation is a commuting zone. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Controls include standardized mean 1931-1943 murder rates and baseline 1940 controls: share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: IPUMS complete count 1940 US census; Boustan (2016); see Appendix D for the full list of data sources on each of the mechanisms.

Figure E3: Great Migration impact on private schooling



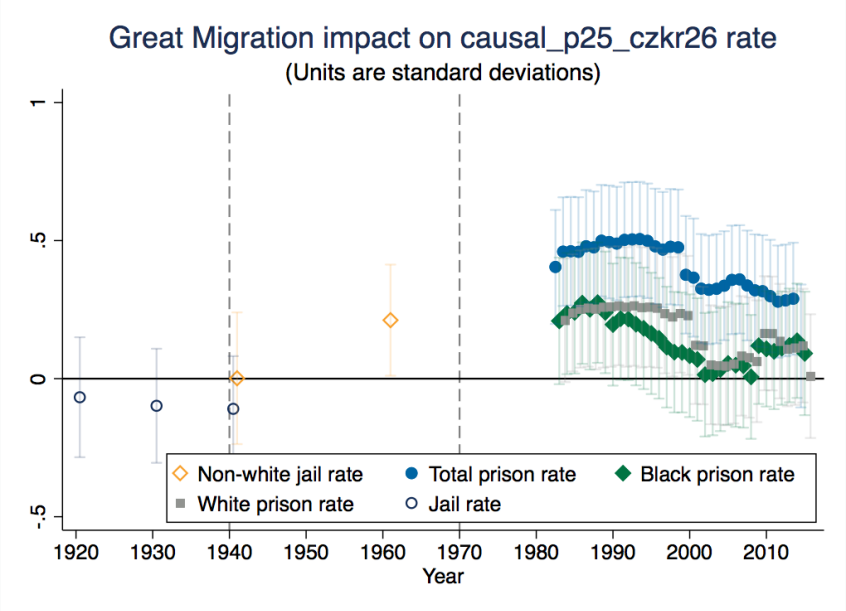
Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions for each year where the dependent variable is private school enrollment rates. The unit of observation is a commuting zone. Units of outcome variables are standard deviations. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Biennial Statistics of Education, 1920-1922; NHGIS county-level aggregates of elementary and high school enrollment by school type (public or private), 1970-2010. Instrument data sources: IPUMS complete count 1940 US census; Boustan (2016).

Figure E5: Great Migration impact on policing investments



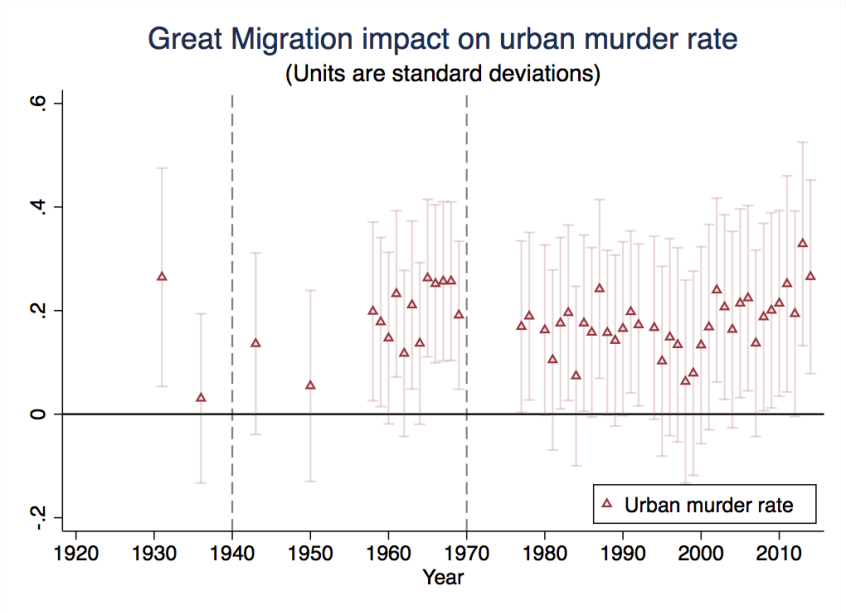
Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions for each year where the dependent variable is either the share of local government expenditures on policing, police expenditures per capita, or city police employees per 100k urban population. The unit of observation is a commuting zone. Units of outcome variables are standard deviations. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Data on police expenditure shares and per capita spending come from Financial statistics of states and local governments, 1932; US Census Bureau Annual Survey of Local Governments (1967-2012); police employees from City Government Employment and IPUMS complete count US censuses (1920-1940). Instrument data sources: IPUMS complete count 1940 US census; Boustan (2016).

Figure E9: Great Migration impact on incarceration rates



Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions for each year where the dependent variable is county jail population per 100,000 (1940 and 1960) or federal and state prison population by 100,000 by county-of-commitment from 1983-2015. Each jail or prison population group is normalized by the population for that group. Federal and state prison rates are for black and white men aged 15-64. The unit of observation is a commuting zone. Units of outcome variables are standard deviations. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: 1940 county jail rates come from US IPUMS complete count US census; 1960 county jail rates come from 1960 Census report on county correctional institution population; data on 1983-2015 federal and state prison population by county-of-commitment come from Vera Institute of Justice In Our Backyards Database. Instrument data sources: IPUMS complete count 1940 US census; Boustan (2016).

Figure E11: Great Migration impact on murder rates



Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions for each year where the dependent variable is urban murder rates per 100,000 in commuting zones. The unit of observation is a commuting zone. Units of outcome variables are standard deviations. The instrument 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. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. *Data sources*: Uniform Crime Reports. Instrument data sources: IPUMS complete count 1940 US census; Boustan (2016).

Table E1: Great Migration CZs experienced more severe 1960s riots

	Ordinary Least Squares					
	Killed Per 100k	Arson Per 100k	Arrests Per 100k	Days of Riots Per 100k	Injured Per 100k	Riots Per 100k
GM	0.00139	0.0755	0.464	0.0265	0.0967	0.0109
	(0.000605)	(0.0335)	(0.166)	(0.00544)	(0.0351)	(0.00214)
R-squared	0.308	0.440	0.605	0.292	0.461	0.311
	Reduced Form					
$\hat{GM}$	0.000815	0.0483	0.420	0.00987	0.0769	0.00406
	(0.000544)	(0.0301)	(0.148)	(0.00521)	(0.0314)	(0.00206)
R-squared	0.291	0.429	0.606	0.179	0.455	0.191
	Two-stage least squares					
GM	0.00274	0.163	1.415	0.0332	0.259	0.0137
	(0.00179)	(0.0997)	(0.543)	(0.0159)	(0.110)	(0.00624)
R-squared	0.279	0.409	0.499	0.283	0.368	0.301
N	130	130	130	130	130	130
Mean Dep Var	0.0589	4.697	24.91	0.950	4.474	0.413
SD Dep Var	0.162	9.952	58.87	1.437	10.65	0.572
SD GM	28.98	28.98	28.98	28.98	28.98	28.98

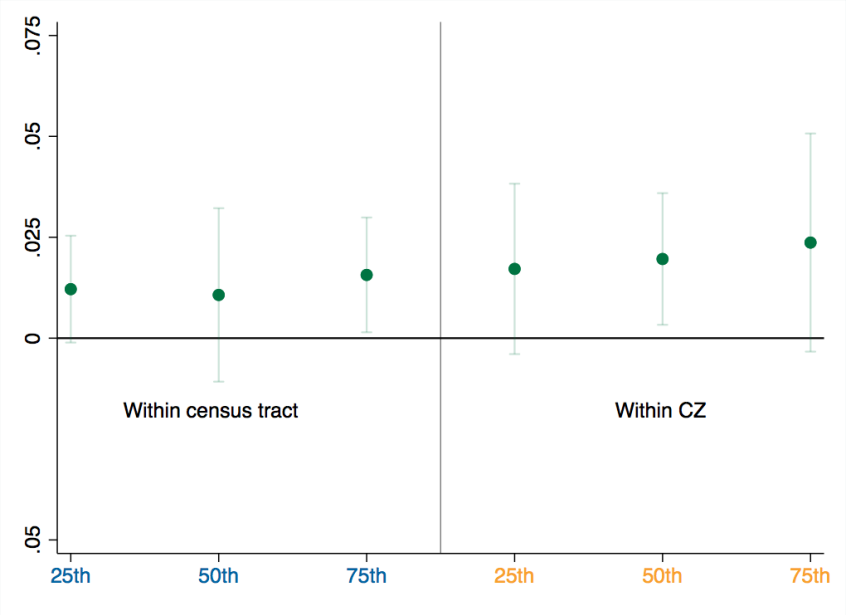
Notes: This table reports the estimated impact of the Great Migration on 1960s race riots and riot severity. Dependent variables in columns 1-5 are individual measures of the severity of riots, including number of individuals killed, number of arson incidents, number of arrests, the duration of the riot in days, and the number of injuries; the final column is total number of riots. All outcomes are normalized by the total CZ population in 1960 and multiplied by 100,000, so they are in per 100,000 of the population units. Independent variable is black population increase between 1940 and 1970. The instrument for black population increase is the predicted black population increase through variation in black southern migration alone. OLS, Reduced Form, and 2SLS estimates are reported. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: Collins and Margo (2007); Carter (1986); CCDB; IPUMS complete count 1940 US census; Boustan (2016).

Table E2: Great Migration and votes for George Wallace, 1968

	Ordinary Least Squares	
	Wallace Vote Per 1k Voters	Wallace Votes Per 1k White Pop
GM	12.41	4.812
	(4.058)	(1.713)
R-squared	0.518	0.514
	Reduced Form	
$\hat{GM}$	12.20	4.840
	(3.642)	(1.536)
R-squared	0.525	0.521
	Two-stage least squares	
GM	40.39	16.03
	(13.87)	(5.771)
R-squared	0.331	0.343
N	130	130
Mean Dep Var	58.49	24.17
SD Dep Var	44.87	18.85
SD GM	1.000	1.000

Notes: This table reports the estimated impact of a one-standard deviation increase in Great Migration inflows on votes for George Wallace, pro-segregation former governor of Alabama and third-party presidential candidate in 1968. Dependent variables in columns 1 is votes for Wallace per 1000 voters, and column 2 is votes for Wallace per 1000 white population in the commuting zone. The unit of observation is a commuting zone. $\hat{GM}$ is the predicted black population increase through variation in black southern migration alone. A one standard deviation increase is approximately 30 percentiles. Baseline 1940 controls include share of urban population made up of 1935-1940 black southern migrants, educational upward mobility, share of labor force in manufacturing, and census division fixed effects. Standard errors are in parentheses. *Data sources*: Collins and Margo (2007). Great Migration data sources: CCDB; IPUMS complete count 1940 US census; Boustan (2016).

Figure E12: Impact of Great Migration on within-census-tract racial gap



Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration in separate regressions where the dependent variable is the racial gap in upward mobility for different geographies and parent income levels. The dependent variable is the difference in mean household income rank between black and white individuals, pooling men and women. Income is measured from IRS tax returns for cohorts born between 1978 and 1983. The first three coefficients reflect the Great Migration's impact on the CZ-average within-census-tract racial gap for individuals with parents at the 25th, 50th, and 75th percentile, weighted by total black plus white population per census tract whose tax returns were used to construct the estimates. The last three coefficients plot the CZ-level gap. Independent variable is the percentile of black population increase during the Great Migration. Independent variable is predicted change in black population share between 1940 and 1970. Baseline controls included. *Data sources*: Chetty, Hendren, Jones, and Porter (2018); IPUMS Complete Count 1940 US Census; Boustan (2016).

Number	Title	file
Main Figures and Tables - produced by 5_main_figures_tables.do		
Figure 1a	Black Upward Mobility in 1940 and 2015, Percentage black teens in median-educ. families with 9-plus years of schooling, 1940	black_edu_mobility_1940_map.png
Figure 1b	Black Upward Mobility in 1940 and 2015, Household inc. rank of black individuals from below-median-income families, 2015	black_p50_mobility_2015_map.png
Figure 2	Quantiles of urban black share increases, 1940-70	bpopchange_percentiles.png
Figure 3	Relationship between 1940-1970 black population change and upward mobility in 2012	permres_GM_nocontrols.png
Figure 4a	Great Migration shift-share instrument, 1935-1940 black southern migrants' origin counties, Detroit vs. Baltimore	detroit_baltimore_migrant_composition.png
Figure 4b	Great Migration shift-share instrument, Southern state net-migration, 1940-1970	southern_netmig_1940_1970_2.png
Figure 5	First stage on black population change	first_stage_GM_GM_hat2.png
Figure 6	Great Migration reduced average upward mobility in northern commuting zones	permres_GM_hat2.png
Figure 7	Childhood in Great Migration CZs lowers adult income of children from low income families	causal_GM_hat2.png
Figure 8	Race and gender heterogeneity in impact of Great Migration on upward mobility	GM_race_kfr_mobility_iv_coefplot.png
Figure 9	Increased segregation, crime, policing, and incarceration in Great Migration CZs	GM_locgov_coefplot.png
Table 1	Placebo test of identification strategy using pre-1940 upward mobility and educational attainment	table1.tex
Table 2	Lower average upward mobility in 2000s for low income families in Great Migration CZs	permres_table_p25.tex
Table 3	Childhood exposure to Great Migration CZs lowers upward mobility for low income families	main_table_p25.tex
Table 4	Great Migration impact on average upward mobility for black families in the 2000s	black_hh_table.tex
Table 5	Great Migration impact on average upward mobility for white families in the 2000s	white_hh_table.tex
Table 6	Robustness of effects of childhood exposure to Great Migration CZs	main_robust_table_p25.tex
Table 7	Robustness of Great Migration's effects on black men's upward mobility	main_robust_table_bmp25.tex
Appendix tables and figures - produced by 6_appendix_figures_tables.do		
Figure A1	Map of 1940-70 change in the black population	GM_map.png
Figure B1	1935-39 black southern migrant composition in select northern cities	northern_cities_migrant_composition.png
Figure B2	1935-1940 black southern migrant educational attainment	blackmig_25plus_educ_hist.png
Figure B3a	School attendance for black teens in North with southern- vs. northern-born mothers, Black teens with illiterate mothers	blackmob_mom_illit_parent_bpl_1920_1940.png
Figure B3b	School attendance for black teens in North with southern- vs. northern-born mothers, Black teens with low-occ-score fathers	blackmob_lowocc_parent_bpl_1920_1940.png
Figure B4	Map of Great Migration instrument	GM_hat2_map.png
Figure C1a	Correlation 1940 \& 2015 Upward Mobility, Black men	black_m_p25_1940_2015_correlation.png
Figure C1b	Correlation 1940 \& 2015 Upward Mobility, White men	white_m_p25_1940_2015_correlation.png
Figure C2	Race and gender heterogeneity in impact of Great Migration on upward mobility, individual earnings	GM_race_kir_mobility_iv_coefplot.png
Figure C3	Heterogeneity in childhood exposure effects by age of child \citep{chetty2018race}	race_hockey_stick.png
Figure C4	Intergen. mobility by race and region	igm_curve_counterfactual_race_region.png
Figure C5	Great Migration effect robust to leaving out each CZ once from sample	leave_one_out_robustness.png
Figure C6	White southern migration impact on childhood exposure effects	causal_GM_hat2_whitemig.png
Figure C7a	White southern migration impact on black men's upward mobility, Parents 25th percentile	bmp25_GM_hat2_whitemig.png
Figure C7b	White southern migration impact on black men's upward mobility, Parents 75th percentile	bmp75_GM_hat2_whitemig.png
Figure C8	Great Migration impact on change in black men's upward mobility, 1940-2015	change_bm_upm_GM_hat2.png
Figure C9	Alternative instruments for the Great Migration	expos_fe_overid_coefplot.png
Figure C4b	Placebo migration shocks	causal_p25_czkr26_versions_95.png
Figure E1	Great Migration correlation with pre-period measures of local mechanisms	GM_locgov_coefplot_pretrends.png
Figure E2	Impact of Great Migration on standardized post-1970 mechanisms, controlling for pre-period murder rates	GM_locgov_coefplot_murder_control.png
Figure E3	Great Migration impact on private schooling	private_school_st_graph.png
Figure E4	Great Migration impact on urban white share	c_whtpop_share_st_graph.png
Figure E5	Great Migration impact on policing investments	police_st_graph.png
Figure E6	Great Migration impact on schooling investments	education_st_graph.png
Figure E7	Great Migration impact on schooling investments, with pre-period control	education_st_graph_preperiod_control.png
Figure E8	Great Migration impact on fire-fighting investments	fire_st_graph.png

Figure E9	Great Migration impact on incarceration rates	incarceration_st_graph.png
Figure E10	Great Migration impact on incarceration rates, levels	incarceration_graph.png
Figure D4a	Great Migration impact on murder rates	murder_rate_st_graph.png
Figure D4b	Impact of Great Migration on within-census-tract racial gap	racegap_geo_coefplot.png

Table A1	Commuting zones in sample	Embedded in tex file appendix.tex
Table C1	Correlation between historical and contemporary upward mobility measures, by race and gender	Embedded in tex file appendix.tex
Table C2	Upward mobility results with coefficients on baseline controls	mob_table_controls.tex
Table C3	Great Migration impact on average upward mobility of high income families in 2000s	permres_table_p75
Table C4	Great Migration impact on childhood exposure effects in the 2000s for high income families	main_table_p75.tex
Table D2	Expenditure by government type by spending category	Embedded in tex file appendix.tex
Table E1	Great Migration CZs experienced more severe 1960s riots	riots_table.tex
Table E2	Great Migration and votes for George Wallace, 1968	wallace_table.tex