Can you move to opportunity? Evidence from the Great Migration

Ellora Derenoncourt*

January 25, 2019

JOB MARKET PAPER

Click here for most recent version.

Abstract

The northern United States long served as a land of opportunity for black Americans, but today the region's racial gap in intergenerational mobility rivals that of the South. I show that racial composition changes during the peak of the Great Migration (1940-1970) reduced upward mobility in northern cities in the long run, with the largest effects on black men. I identify urban black population increases during the Migration at the commuting zone level using a shift-share instrument, interacting pre-1940 black southern migrant location choices with predicted outmigration from southern counties. The Migration's negative effects on children's adult outcomes appear driven by neighborhood factors, not changes in the characteristics of the average child. As early as the 1960s, the Migration led to greater white enrollment in private schools, increased spending on policing, and higher crime and incarceration rates. I estimate that the overall change in childhood environment induced by the Great Migration explains 43% of the upward mobility gap between black and white men in the region today.

*Derenoncourt: Harvard University. Email: elloraderenoncourt@fas.harvard.edu. I thank my advisors Lawrence Katz, Claudia Goldin, Nathaniel Hendren and Edward Glaeser whose support and feedback have proved essential to this project. I am grateful to my mentors Suresh Naidu and Leah Boustan for innumerable conversations and guidance. Heather Sarsons, Maximilian Kasy, Nathan Nunn, Melissa Dell, Thomas Piketty, Emmanuel Saez, Ran Abramitzky, Marianne Wanamaker, James Feigenbaum, Robert Margo, Christopher Muller, Joshua Abel, Elizabeth Mishkin, Sydnee Caldwell, Chenzi Xu, Felix Owusu, Krishna Dasaratha, and Niharika Singh all provided many helpful comments. Price Fishback, Bill Collins, Robert Margo, Augustin Bergeron, Marco Tabellini, Vicky Foukas, and Soumyajit Mazumder generously shared data. Ariel Gomez, Sergio Gonzales, Julian Duggan, and Seung Yong Song provided excellent research assistance. This paper benefited from participants' comments at the First WID World Conference, the Calvo-Armengol Workshop, and seminars at Harvard, Berkeley, and Stanford. This work has been supported by the Harvard Lab for Economic Applications and Policy and by award #83-17-19 from the Russell Sage Foundation. 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 apparent advantage exists for black children growing up outside the South (Davis and Mazumder, 2018). Racial gaps in upward mobility—defined as children's adult outcomes conditional on parent economic status—are similar across the country (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 link between these two phenomena—the migration North and declines in upward mobility—is important for assessing the stability of childhood location effects, which have been shown to be substantial in a variety of experimental and quasi-experimental contexts (Chetty et al., 2016; Chetty and Hendren, 2018a,b).

This paper tests whether the Great Migration (termed "Migration") reduced northern cities' ability to facilitate black intergenerational progress. I do so by comparing commuting zones in the North that exogenously experienced larger or smaller increases in their urban black population. I estimate the impact of these increases on average outcomes for individuals born in the 1980s. I find that the Migration lowered upward mobility in the long run. Black men were the most affected subgroup, implying a widening of the racial gap in former Great Migration destinations. Using an alternative measure of upward mobility–the childhood exposure effects of commuting zones–I show that the Migration's impact was mediated by childhood environment, not unobserved family characteristics. Northern cities responded endogenously to racial composition changes in ways that reduced the gains from growing up in the North for future generations of black children.

I draw on several data sources to conduct the analysis in this paper. I use the complete count US censuses from 1920-1940, commuting zone level upward mobility estimates from Chetty and Hendren (2018b) and Chetty et al. (2018), and a database I assembled on local government expenditures, private schools, murder counts, and other characteristics

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

of commuting zones spanning the period 1920-2015. The core sample is 130 non-southern commuting zones covering roughly 85% of the non-southern US population in 2000 and 58% of the US population overall. These locations contained 97% of the non-southern black population in 2000 and 50% of the black population overall.

The empirical strategy makes use of the fact that black migrants during the Great Migration settled 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 these variables is 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 Great Migration shock to a commuting zone to be the percentile of predicted black population increase.[4]

Using this strategy, I show that the Great Migration led to a reduction in upward mobility in destination commuting zones in the North today. A 30-percentile greater increase in the black population, approximately 1 standard deviation of 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 child; 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

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

whether the causal effect of a commuting zone varies with exogenous historical increases in the black population. The interpretation is as follows: if a random 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 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. I use observational estimates of upward mobility for different racial groups and show that the largest negative effects manifest for black men. I find no impact of the Migration on the household income of black women, but the impact on their individual earnings is weakly positive. 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 digitized and assembled data on local governments, schools, and crime in commuting zones covering the years 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 starting in the 1960s 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.

First, using data from historical reports and Census questions on school attendance, I show that private school enrollment rates increased as a result of the Migration. These increases were driven by white private school enrollment rates; the impact on black private school enrollment is negative and statistically indistinguishable from zero. Consistent with Boustan (2010), I also find that the Migration led to substantial reductions in the urban white population share in northern commuting zones. These areas continue to be more racially and economically segregated today.

Second, using detailed data on local governments operating in each commuting zone, I show that local governments spent a larger share of public expenditures on police,

increasing the number of police officers and police expenditures per capita, in Great Migration cities. The increase may reflect a change in crime rates: beginning in the late 1960s, Great Migration destination CZs have persistently experienced higher urban murder rates. In contrast with public expenditures on police, I find no reductions in expenditures on education, or other categories of spending over which local governments have a large degree of discretion. These include fire-fighting; health and hospitals; sewage and sanitation; and recreation.

Consistent with the findings on policing and crime, the Great Migration also led to higher incarceration rates. As early as 1960, the Migration was associated with higher local county jail rates for non-whites. For the more recent period, I use rich new data on the county-of-commitment to federal and state prison for the incarcerated population starting in 1983. Larger historical black in-migration predicted higher rates of federal and state incarceration, particularly in the early 1990s. The magnitude of these increases for the black population is large. In 1992, at the peak of the Great Migration's impact on incarceration rates, a 30-percentile greater shock resulted in an additional 300 per 100,000 black 15-to-64 year-olds versus 30 per 100,000 white 15-to-64 year olds being sent to federal and state prison.

A key competing explanation for these long-run declines in Great Migration cities is deindustrialization. Black southerners moved to manufacturing centers that subsequently underwent greater job loss than more economically diversified locations. 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. Further, 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. White southerners also migrated to northern cities over the 20th century. In a placebo exercise, I show that instrumenting for white southern inflows has no effect on black upward mobility or on commuting zone childhood exposure effects. Finally, the impact of black population increases is robust to flexibly controlling for lagged black population shares prior to 1940, suggesting that changes in the racial composition, not simply the levels of the black population, help explain the findings.

This paper provides evidence that responses to the Great Migration in destination cities eroded some of the gains of migrating for future generations of black children. The effects were particularly deleterious for black boys. A counterfactual exercise suggests that without the causal impact of the Great Migration, the racial gap in upward mobility among men in the North would be roughly 43% smaller.[6] An important component of the

[6]Defined as the commuting zone level gap in mean income rank between black and white men with

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.

This paper relates to several literatures. First, 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, 2014; Chetty and Hendren, 2018a,b). Although the correlates of location effects and differences in upward mobility have been well documented, the stability of these effects in response to shocks is much less understood. I provide novel evidence that mid-century racial composition shocks altered the effects locations had on children, turning high opportunity locations into opportunity deserts, particularly for black families.

Second, I contribute to the literature on the effects of the Great Migration on destination cities. Boustan (2010, 2009) showed that black in-migration spurred post-war white flight into suburban neighborhoods and increased labor market competition among black workers in the North. 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 (Foukas et al., 2018). In this paper, I provide evidence of long-run effects of the Great Migration on upward mobility and shed light on a new intermediate impact on cities: higher crime and incarceration rates and greater relative investment of public expenditures in policing.[8]

Finally, my findings relate to theories of local public finance and population heterogeneity (Alesina et al., 2004; Tiebout, 1956). If locally provided public goods can improve children's outcomes, then substitution out of public goods and into private alternatives can lower outcomes for children from lower and middle income families. In areas with

median income parents.

[7]For literature on this topic, see Ananat (2011); Andrews et al. (2017); Cutler and Glaeser (1997); Kasy (2015); Massey and Denton (1993); Graham (2016); Sampson et al. (2002); Wilson (2012).

[8]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.

urban residential racial and income segregation, lower income central cities may need to spend additional resources addressing crime as opposed to spending on education. Such a reallocation of public spending could worsen inequality between urban black and suburban white children in metropolitan areas.

The paper is structured as follows. Section 2 gives an overview of the historical context. Section 3 describes the data sources 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 effect of the Migration on upward mobility. 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.[9] 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.[10] 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, daughter of a migrant and later an activist in the Civil Rights Move-

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

[10]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 during World War I and immigration controls put in place in the 1920s.

ment, 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, 1991). 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 to the South. Furthermore, racial gaps in teen school attendance were much lower in the North. In Appendix A, I document that this regional difference remained fairly constant over the period 1880-1940. 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 declining black upward mobility in the North.

3 Data

In this section I describe the data used to measure upward mobility, the construction of the analysis sample of commuting zones, and my measure of urban black population change during the Great Migration. I conclude the section by discussing key correlates of the Migration and upward mobility for the commuting zones in my sample. Throughout, I define upward mobility in a location as the average outcomes of children conditional on parent income or educational status.

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.[11] Following Card et al. (2018), I define educational upward mobility as the fraction of 14 to 17 year-olds in each commuting zone with 9 or more years of schooling from households where the household head has between 5 and 8 years of schooling, approximately the median for adults in the US at the time.[12] 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.

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.[13]

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.[14]

[11] The Integrated Public Use Microdata Series ("IPUMS") version.

[12] I use the household head's years of education as the measure for parent educational attainment while Card et al. (2018) use the maximum of father and mother educational attainment.

[13] 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.

[14] 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, and half of household self-employment income where relevant.

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.

How comparable are educational upward mobility in 1940 and income upward mobility in the 2000s? The two measures are strongly correlated across US CZs, where both are available, with a correlation coefficient is 0.43. Additionally, income upward mobility is strongly correlated with high school graduation rates in low income families today, with a correlation coefficient of 0.53.

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 percentile of the parent income distribution by childhood commuting zone. In this dataset, outcomes are measured in 2015 when individuals are between the ages of 32 and 37.

3.2 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 census and the City and County Data Books 1944-1977 series ("CCDB")[15], 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 use the CCDB to collect information on the black population in cities in 1970.[16] I restrict the sample to cities that are not missing population data in 1940 in the CCDB. I further restrict the dataset to those cities that had at least one recent black southern migrant, defined as an individual who listed a southern county of residence in 1935, but resided in a northern city in 1940. The total number of cities that meet these criteria is 294.[17] 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.

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

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

[16] 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.

[17] 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.

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—Steubenville, OH and Milwaukee, WI. Both were major manufacturing centers in the 1940s. Steubenville's urban black population share increased by 3.2 percentage points (38th percentile) while Milwaukee's increased by 14.8 percentage points (78th percentile). At the tail end, commuting zones like Detroit, MI, Gary, IN, and Washington, DC, saw very large increases, ranging between 30-50 percentage points.

The descriptive relationship between black population change during the Great Migration and average income upward mobility today can be seen in Figure B1a. The relationship is strikingly negative and linear.[18] A 1-percentile greater black population increase between 1940 and 1970 is associated with a .15 percentiles reduction in adult income rank for individuals with lower income parents. However, this relationship cannot be interpreted as causal as correlates of black population change may drive this relationship.

3.3 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.[19]

[18]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 the changes in the black population between 1940 and 1970. Smaller increases take place in a context of lower levels of the black population share and therefore may also have a large impact. As I discuss in Section 5, my results are robust to including controls for the level of the black population share in 1940.

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

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. Figure B1b shows the correlation between the percentile of black population increase during the Great Migration and several baseline 1940 characteristics: educational upward mobility, the share of the labor force in manufacturing, and the share of the population made up of recent black southern migrants.[20]

Black urban populations increased more in places with higher levels of educational upward mobility, a greater share of the labor force in manufacturing, and in locations that already had a substantial number black southern migrants. 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 such 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 3 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 by building on a standard shift-share approach for the 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 com-

[20]Data on recent black southern migrants come from the 1940 complete count census. I define a recent southern black migrant 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.

munities had settled, thus generating chain migration patterns similar to those observed in the international migration context. The resulting variation in migrant composition is plausibly orthogonal to characteristics of destinations that influence the location choices of future migrants as well as the evolution of upward mobility in destination locations. I evaluate potential violations of this assumption in detail in Section 5.4.

Variation in pre-1940 migrant composition can then be interacted with variation in outmigration from origin locations driven by origin factors alone ("push factors"). Push factors include war spending and shocks to cotton as well as other economic sectors in the South, for example, tobacco and mining.

Building off of the approach described above, I define the following instrument to analyze the long-run impact of the Migration on upward mobility: the percentile of black population change if the black population were to increase solely through the interaction of the two sources of variation described above. 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.[21] The term $\hat{m}_j^{1940-1970}$ consists of the sum of fitted values of decadal predictions of southern county net migration rates (from 1940-1950, 1950-1960, and 1960-1970) using lagged southern economic predictors of migration:

$$\hat{m}_j^{1940-1970} = \sum_{t=1950}^{1970} \hat{m}_{jt}$$

where fitted values, $\hat{m}_{jt} = m_{jt} - \varepsilon_{jt}$, come from the following prediction of net-migration:

$$m_{jt} = \beta_0 + \mathbb{Z}'_{jt-10}\beta_1 + \varepsilon_{jt}.$$

[21] Recently migrated black southern migrants in 1940 are black individuals who reported a southern county of residence in 1935 and were enumerated in a different county of residence in 1940. The 1940 census was the first census to record internal migration in this way. Enumerators asked individuals about their prior residence (city, county, and state) in 1935.

Appendix C describes the construction of ω_{jc} and $\hat{m}_j^{1940-1970}$, and the procedure for choosing predictors $\mathbb{Z}'_{jt-10}$ 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 as 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) to estimate impacts of the Great Migration on destination cities. There are two differences worth noting between my approach and the one developed by Boustan (2010). Boustan (2010) uses southern state of residence in 1935 to assign future waves of migrants from southern states to northern cities. 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.[22] 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.

Because there are over 1200 southern counties from which migrants moved North, this approach yields highly specific linkages between southern and northern locations. Using the microdata to construct the migration matrix also gives me the flexibility to explore alternative specifications of the instrument using age, gender, and demographic characteristics of the individual migrants. As I discuss further in Section 5.4, I construct several different instruments for black population change and conduct overidentification tests, which fail to reject the null that the estimates of the impact of the Great Migration are statistically indistinguishable from one another.[23] As I discuss in the same section, the results are also robust to defining migration shares using southern state of birth in 1940, 1930, or 1920.

The second difference is that I use machine learning to improve the prediction of

[22]Boustan (2010) uses tabulated census reports that report 1935 state of residence to construct southern-state-to-northern-city weights to assign migrants. The complete count 1940 census was declassified in 2012.

[23]Hornbeck and Moretti (2018) performs overidentification tests using several instruments for productivity changes across cities.

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 the relationship between these variables and county net migration rates using OLS.[24]

Although the instrument alleviates some concerns regarding the endogeneity of black population increases in northern urban areas, the shares themselves reflect migration patterns during the first wave of the Great Migration and may themselves be endogenous to characteristics of destinations that affect the course of upward mobility in subsequent decades. I therefore include controls for the baseline 1940 characteristics discussed in Section 3, including the share of the labor force in manufacturing and the share of the 1940 city population made up of recent black southern migrants.

I also include educational upward mobility in 1940 as a control in all baseline regressions. As I describe in greater detail in Section 4, 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. 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.

My preferred specification includes baseline 1940 characteristics for robustness; however, their inclusion does not significantly alter the point estimates. I report key results with and without this baseline set of controls. Finally, all specifications include census division fixed effects to control for systematic regional differences in migration patterns and upward mobility.[25]

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

[25]Including census division 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 2. 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

Estimating equation

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

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

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

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, then 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.

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

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.

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.[26]

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.

First-stage results

Figure 4 shows a binned scatterplot of the relationship between GM, the percentile of actual black population increase, and $\hat{GM}$, 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 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 23.[27]

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.[28] From the van-

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

[27]In Appendix Figure C4a, I also show the first stage with respect to the level of black population increase, normalized by initial 1940 CZ urban population. Figure C4 shows this increase for a windsorized sample. The levels of predicted black population increases are shifted down relative to the actual. This is due to using a relatively low period of migration (1935-1940) to form the migration matrix generating predicted inflows into cities in the sample.

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

tage 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 results on upward mobility in detail, before investigating plausible local mechanisms in Section 6.

To focus on the more plausibly causal estimates of the impact of the Great 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 5 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 striking negative relationship between historical black migrant inflows and average outcomes for individuals from low income families in the destination CZs today.

Table 4 reports 2SLS estimates of the relationship. A 1-percentile increase in the historical black population lowered household income rank by -0.0981 percentile points

(s.e. = 0.0301). OLS estimates are reported in Table 4 as well. The 2SLS coefficients are larger in magnitude than the OLS. One potential explanation for this is that omitted characteristics that are correlated with both black population change and upward mobility are positively correlated with both. For example, the black population grew more in places with higher levels of median educational 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. Table 5 reports the results for individuals with high income parents. I do not find strong effects of $G\hat{M}$ on the outcomes of individuals with high income parents (at the 75th percentile of the parent income distribution). I find evidence of weak negative effects on the individual income rank of men with high income parents.

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 income 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 $\bar{\theta}_{p,CZ}$.[29] 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. 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 explain a substantial portion of the Migration's estimated impact on $\bar{y}_{p,CZ}$. Several studies have

[29]Further, incumbent families may leave an with high levels of in-migration, further potentially affecting $\bar{\theta}_{p,CZ}$.

found persistent differences in intergenerational mobility by race (Davis and Mazumder, 2018; Mazumder, 2014; Bhattacharya and Mazumder, 2011). Chetty et al. (2018) find that black men have lower income rank than white at every parent income rank, and these gaps persist even among those observed to be growing up in the same census tract. 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 5, 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.[30] 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.[31]

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 upto an average effect

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

[31] These estimates are valid estimates of $\mu_{p,CZ}$ if the identifying assumption that the age of the child at the time of the family's move is orthogonal to omitted determinants of outcomes for children. See Chetty and Hendren (2018a) for several checks of these identifying assumptions including instrumenting for moves with displacement shocks to families and the inclusion of family fixed effects.

across black and white children. Data limitations prevent me from exploring potentially heterogeneous impacts of the Migration on $\mu_{pr,CZ}$, 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 6 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 8 reports OLS and 2SLS estimates of the relationship. The 2SLS estimates an be interpreted as follows: a 1-percentile larger increase in the historical black population lowers household income rank by -0.0083 percentile points (s.e.=.0025). The first stage is around 0.30, 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 zero.

I find strong negative effects of predicted black population increases on exposure effects for individuals with high income parents (see Table 9), in contrast with the impact of the Migration on average upward mobility for high income families. The impact on exposure effects on men's individual income rank is particularly strong, compared to effects on household income. Overall, the impact of predicted black population increases on childhood exposure effects is similar for individuals with low and high income parents.

5.2.1 Interpretation of results on childhood exposure effects

The results thus far support the hypothesis that one way the Great Migration lowered upward mobility was by negatively impacting childhood environment. 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 D2).

The table below decomposes the impact of the Great Migration on upward mobility through the channel 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 population. 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 15.52. Appendix Section D.1.1 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 Migrations effect 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 167% ($\frac{5}{3} \times 100$) 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 130% 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	-3
15.52 years	-3.9	-3

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 D2 for black men and women and Table **??** for white men and women.

Figure 7 summarizes these regressions in a plot of the coefficients on percentile of predicted black population change in regressions of each sub-group's average upward mobility on $\hat{GM}$. Here the shock has been scaled to be in 30-percentile units, approximately 1 standard deviation. The negative effects of the Migration load on black men. A 30-percentile increase in the intensity of a CZ's Great Migration shock lowers the individual income rank of black men by around 1 percentile point, 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.[32]

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.

[32]The term "white" refers to non-Hispanic white population (Chetty et al., 2018).

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 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 D1.

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 men have reduced conditional income as a result of the Migration but white men are unaffected has implications for the racial gap in income upward mobility in the US. Chetty et al. (2018) find that conditional on parent income, black and white women have identical income rank as adults. This implies that the substantial racial gap in upward mobility is driven primarily by differences in the outcomes of black and white men with similar parent income during childhood. 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 men with low income, high income, and median income parents.

The counterfactual seeks to address the following question: what would the racial gap in men's upward mobility in North be without the changes induced by Great Migration? I first calculate the racial gap in each commuting zone at each of the listed parent income rank by taking the difference in mean white and black men's conditional income rank. Figure D1 illustrates the positive relationship between $\hat{GM}$ and the racial gap. I then predict what the racial gap would be if each location instead received the lowest percentile

of shock.

A comparison of the actual average racial gap across northern urban commuting zones and the counterfactual gap in absence of the effect of the Great Migration is reported in the table below.

Contribution of the Migration to the racial upward mobility gap among men

	Low Parent Income	High Parent Income
Average gap	10.46	11.03
Counterfactual gap (s.e.)	6.9 (.16)	5.0 (.24)
Pct Change	-34%	-55%

The first row reports the average gap in mean black and white men's upward mobility across commuting zones in the sample: 10.46 income rank percentiles for men with parents at the 25th percentile and 11.03 income rank percentiles for men with parents at the 75th percentile. The second row reports the counterfactual average gap if each location received the lowest percentile of Great Migration shock. Under this counterfactual, the average racial gap across northern commuting zones would be 6.9 percentiles (s.e. = .16 percentiles) for men with low income parents, and 5 percentiles (s.e. = .24 percentiles) for men with high income parents.

These estimates suggests the Migration increased the racial gap by 36% for low income families and 55% for high income families. Finally, looking at men with median income parents, I estimate the Migration increased the gap between black and white men with median income parents by 43%. 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 explanations for the findings, namely, the role of potential omitted variables.

5.4 Alternative explanations

Deindustrialization A key competing explanation for reductions in children's outcomes in Great Migration destinations is deindustrialization: black southerners moved to booming industrial centers, and these areas subsequently underwent greater job loss than locations less specialized in manufacturing. 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.[33] I find no effect of the Migration on the share of the labor force in manufacturing from 1950 to 1970. Further, I find no effect of the Great Migration on the adult outcomes of white men with low income parents, a demographic group that would likely be affected if the findings were driven by deindustrialization rather than changes in racial composition.

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.[34] Appendix Figure D5 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.

[33]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.

[34]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.

Black population shares Finally, I examine the extent to which the findings are driven by variation in historical black population levels as opposed to increases. If locations with high black population shares are fundamentally different from those with lower black population levels, this fixed characteristic could be a confounding factor for my findings. The impact of exogenous black population increases ($G\hat{M}$) is robust to controlling for lagged black population shares prior to 1940, suggesting that changes in the racial composition, not simply the levels of the black population, contributed to the changes I document. I include controls for the 1920, 1930, and 1940 black population shares.

Results are reported in tables 2 and 3. 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's upward mobility. 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 lagged black population shares and flexible controls for the black population share in 1940 (e.g., separate controls for quartiles of the black population share—see Appendix Figure D6).

Bartik: potential endogeneity of shares and shifters Recent analyses of Bartik or shift-share instruments commonly used to estimate the causal impact of immigration or of demand shocks to labor markets have generated new intuitions on the exogeneity of shift share instruments. Two views have emerged on the source of exogeneity in the instrument. Goldsmith-Pinkham et al. (2018) argue for exogeneity in terms of shares: the identification condition in a shift-share or Bartik style instrument is satisfied if, for example in the context of manufacturing labor demand shocks, industry shares are exogenous to location characteristics. The authors recommend investigation of whether specific shares drive the findings as endogeneity of these shares are a particular threat to the identification strategy.

In my setting, over 1200 shares link northern destinations to southern county migration. The large number of shares that can provide identification is reassuring: both Adão et al. (2018) and Goldsmith-Pinkham et al. (2018) find that identifying assumptions are more likely to be met the greater the number of shares. To test whether there are particular southern counties driving the results on upward mobility, I directly regress outcomes

on city-population weighted average shares at the commuting zone level.[35] Because the number of shares far exceeds the number of destinations, I use LASSO to select shares that best predict outcomes, focusing on the childhood exposure effects outcome and including baseline controls. Four shares are ultimately selected from the following southern counties: Tucker County, WV; Armstrong County, TX; Mitchell County, NC; and Hamilton County, FL. Results are robust to controlling for these shares in the baseline regression.

Borusyak et al. (2018) 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 northern destinations. Relying on predicted rather than actual outmigration from southern counties alleviates some of these concerns. To address the possibility that shocks to particular states are correlated with shocks to, for example, Detroit or Baltimore, I show that my results are robust to first residualizing county net-migration on state fixed effects and then predicting migration on a new set of optimal county-level predictors chosen through my Post-LASSO estimation procedure. In Appendix Figure D4, I report these results.

Finally, I conduct an additional check on the validity of my identification strategy using over-identification tests using three slightly different sources of variation to construct the instrument. In addition to the baseline instrument, a second instrument is constructed using the southern county outmigration rates that are first residualized on state fixed effects. A third version of the instrument uses variation in state of birth across the southern born black population in northern cities in 1940 to address the fact that 1935-1940 migration rates were low as a result of the Great Depression. Results using each of these versions of the instrument are extremely similar. Further, an over-identification test fails to reject the null that the estimated effects on upward mobility are statistically indistinguishable from each other.[36]

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 childhood environment. Further, the change in childhood environment appears to have been particularly detrimental for black men.

[35] If southern counties A and B sent migrants to city C and D, then average A and B shares at the commuting zone level are the sum of the shares of each county in C and D weighted by the population in C and D, divided by the total urban population $C + D$ in the commuting zone.

[36] Hansen J statistic p-value of .15.

6 Evidence on location-based mechanisms

Why did upward mobility decline in northern destination commuting zones? To explore potential mechanisms, I assembled a new database on commuting zones spanning the years 1920-2015. The database covers statistics on schooling, demographics, local government expenditures, incarceration, and crime, among other characteristics of commuting zones. 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.

Additionally I used the complete count censuses between 1920-1940 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 F.

The database allows me to evaluate what aspects of urban neighborhoods causally responded to the Migration by estimating the impact of predicted black population increases on outcomes before, during, and after the period of the Migration. 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 approximately one standard deviation (30 percentiles of predicted black population increase).

The results from this analysis reveal shifts in three areas dating back to the 1960s and 1970s that persist over the next several decades. First, I find increased white enrollment in private schools and declines in the number of whites living in urban neighborhoods in commuting zones that experienced greater in-migration. Second, I find that public spending in Great Migration destinations was reallocated towards policing. Third, some of the increase in police spending may be related to elevated crime rates in Great Migra-

tion cities, particularly in the late 1960s. During this period, a national crime wave and race riots struck cities throughout the US. I find that race riots were of greater intensity in Great Migration cities, lasting longer and involving more deaths, injuries, and arrests.

Section 6.1 briefly reports the specific findings for each mechanism. In section 6.2, I situate the results in the broader literature on the impact of education, segregation, crime and criminal justice policies on children.

6.1 Results on mechanisms

Impact on private schooling and residential segregation

I begin by reporting results on private schooling and residential segregation. Figure 8 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,[37] a 30-percentile increase in the Great Migration shock is associated with a 0.2 standard deviation increase in the private school enrollment rate. This association appears driven entirely by white private school enrollment rates. The point estimate for black children's private school enrollment rates is negative and statistically insignificant.

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 E1. In Appendix Table E2, I show additional results on residential sorting: the long-run impact of the Migration shock on racial and income residential segregation across commuting zones. The results above are consistent with white flight from public schools and urban neighborhoods. I discuss the implications these findings may have for upward mobility of black and white children from low income families in Section 6.2.1.

[37]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 F for more details.

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 F1 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 Section 6.2.2. School spending has natural implications for the average outcomes of children in a given location.[38]

Figure 9 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. Appendix Figures E5 and E6 report these results. I estimate a noisy negative association between the Migration on pre-period (1932) aggregate educational expenditure shares. In E6, 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 detail in Section 6.2.1.

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. Appendix Figure E4 reports these results. I find no impact of the Migration on fire-fighting expenditures. Higher police expenditures

[38]See Section 6.2.1.

may be associated with higher crime and incarceration rates. I investigate these below.

Impact on incarceration rates

Figure 10 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 Appendix Figure E2, 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.

Impact on murder rates

Figure 11 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 10 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. I discuss the implications of this and the other findings reported above in the sections below, first beginning with the results on private schooling and residential segregation.

6.2 Discussion

6.2.1 Education spending, school quality, and residential segregation

Two prominent explanations for local differences in children's outcomes and racial inequality in outcomes are school quality and segregation. Recent work by Jackson et al. (2015) has shown that school spending can improve adult earnings, and Card et al. (2018) show that in the early 20th century, improved school quality improved educational attainment of children and educational upward mobility.

Aggregate differences in school spending across commuting zones do not appear to be a mechanism for the impact of the Great Migration on upward mobility in destination cities. 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. Further, if there were aggregate reductions in school spending, one would expect to see worse outcomes for white men and women and black women from low income families growing up in these locations. However, I find no negative impacts of the Migration on these subgroups. The results are consistent with Rothstein (2018), who finds that differences in school quality can explain very little of the variation in average upward mobility across commuting zones (roughly eleven percent).

Still, aggregate measures of school spending may mask differences across individual school districts within commuting zones. I find that a higher fraction of white children are enrolled in private schools in Great Migration destination cities. Private school enrollment rates tend to be higher in urban areas, so these results are suggestive of lower school quality in urban 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.

Residential racial segregation has been shown to negatively impact black children's outcomes, as well as black-white differences (Ananat, 2011; Cutler and Glaeser, 1997). Further, segregation is a major correlate of average upward mobility in counties and commuting zones (Chetty et al., 2014; Chetty and Hendren, 2018b; Andrews et al., 2017). I find evidence that the Migration increased increased segregation and that destination commuting zones are more segregated today. A large literature in sociology examines the confluence of several factors negatively correlated with black children's outcomes in racially segregated neighborhoods, including the prevalence of single parent families and high crime rates (Wilson, 2012; Massey and Denton, 1993).

My findings are strongly consistent with the existing evidence linking residential racial segregation with worse outcomes for black children. Lower marriage rates and higher single parent families in segregated areas also help explain why the Migration is positively associated with the individual earnings of black women. In areas with lower marriage rates, more black women may be in the labor force and working more. Still, the particularly negative effects on black men's upward mobility warrant a look into facets of neighborhoods and local policy that may disproportionately affect black men.

6.2.2 Policing, crime, and incarceration

How might higher policing, crime, and incarceration affect upward mobility for black men? Several studies have shown that exposure to crime increases individual criminal behavior, which has consequences for one's probability of incarceration and traditional employment (Case and Katz, 1991; Damm and Dustmann, 2014; Heller et al., 2017; Sviatschi, 2018). Crime and residential racial segregation are highly correlated across urban areas, which suggests that black children are disproportionately exposed to crime and violence compared to white children growing up in the same commuting zones. Childhood exposure to higher crime rates may thus directly reduce black men's income upward mobility relative to white in Great Migration cities.

A growing literature, however, suggests that polices used to curb crime may also have direct exacerbating effects on racial inequality. A recent paper by Liu (2018) examines the impact of incarceration of black men on black women's marriage outcomes and family structure for black children using variation in federal and state sentencing policy from 1986 to 2009. Incarceration lowers the marriage rate for black women and increases black children's likelihood of being born out of wedlock and living in single parent households. The author further finds that incarceration increases the gap in upward mobility between black and white men. The results are consistent with incarceration being a mechanism

for the Great Migration's impact on black men's upward mobility through incarceration of the father generation. Further Liu (2018) finds that higher incarceration rates increase black women's probability of employment. These results are consistent with shocks to black men's incomes having an income effect on black women's labor supply.

Legewie and Fagan (2018) find that a policy increasing police activity in New York City had negative effects on test scores and school attendance of black male teenagers. The mechanism appears to be increased police stops and arrests of black male teens, which disrupted their education. Reductions in crime were small in magnitude, suggesting that police activity may have had a net negative impact on black boys' outcomes in this case. Ang (2018) finds that police shootings of civilians lowered educational and behavioral outcomes of students in Los Angeles, leading to reduced high school graduation and college attendance rates. Dobbie et al. (2018) find that parental incarceration increases teen crime and pregnancy and lowers subsequent employment for youths from disadvantaged families.

Another way that reallocation of public spending to policing may affect black men's upward mobility is through diverting resources from other uses. Jackson et al. (2015) find that school spending has positive effects on the adult income of children, and low income children in particular. Despite a likely increase in the share of children who would benefit from additional educational expenditures, local governments did not increase educational expenditures in response to the Great Migration. Rising crime rates or the perception of rising crime rates in the late 1960s may have led governments to specialize instead in police protection, and this allocation of public spending persisted for several decades since. The empirical strategy in this paper cannot distinguish between the potential direct effects of higher crime rates in Great Migration destinations from the allocation of public spending towards policing as opposed to other goods. Further research will have to disentangle the impact of these two effects of the Migration.

7 Conclusion

The divergent trajectories of black and white boys with observably identical parental income is one of the most striking examples of inequality in the US. One of the policies aimed at reducing disparities in children's outcomes is to incentivize families to "move to opportunity" (MTO), or to relocate to neighborhoods with better opportunities, as measured by poverty rates, average outcomes, or, more recently, childhood exposure effects. Yet the persistent gap in upward mobility among black and white boys growing up in the same neighborhood raises the question of whether these types of policies can

be effective at reducing racial gaps in children's outcomes, and if not, why not.

Over the 20th century, black Americans engaged in perhaps the largest natural experiment in MTO in US history. The Great Migration of African Americans out of the South 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 believed would be enjoyed by future generations. Using a shift share instrument to isolate exogenous increases in the black population in northern urban commuting zones during the Great Migration, I show that racial composition changes during this period reduced the ability of northern cities to promote upward mobility in the long run and harmed black upward mobility in the North specifically.

In response to mid-century changes in the racial composition of northern cities, white families withdrew from shared public schools and urban segregation increased. 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, potentially crowding out investments in, for example, education for an increasingly disadvantaged urban population.

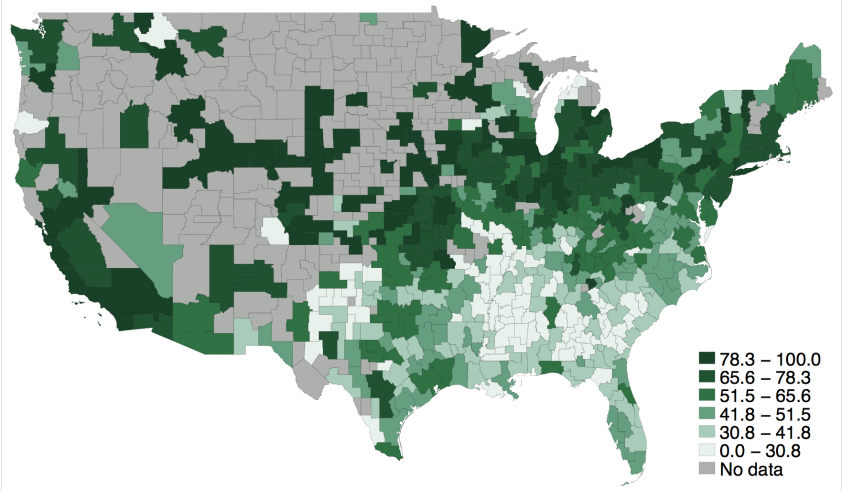
At the time that prison populations began to rise dramatically in the US in the 1980s and early 1990s, places with larger increases in their black population during the Great Migration sent substantially more of the black population to federal and state prison. The timing of the effect of the Migration on incarceration rates suggests that parents of children born in the 1980s would be most affected. 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 separately assess the long-run impact that increased crime, the race riots of the 1960s, and city policy responses to each have had on black men's outcomes. A key question is whether alternative policies can reduce racial inequality in upward mobility given the sizable gaps under the existing set of policies.

My findings have implications for policies that incentivize families to move to areas with better opportunities and, in particular, the general equilibrium effects of scaling such a policy. In response to millions of black migrants moving North to improve 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,

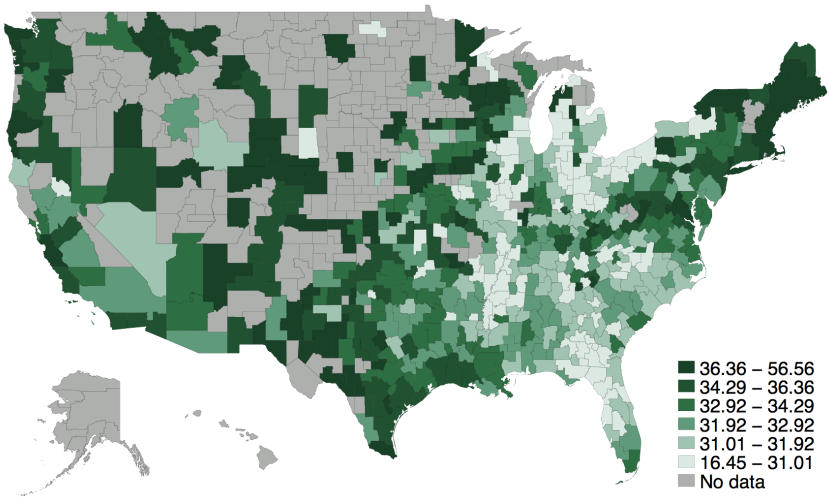
may be warranted.

Figures and tables

Figure 1: Black Upward Mobility in 1940 and 2015



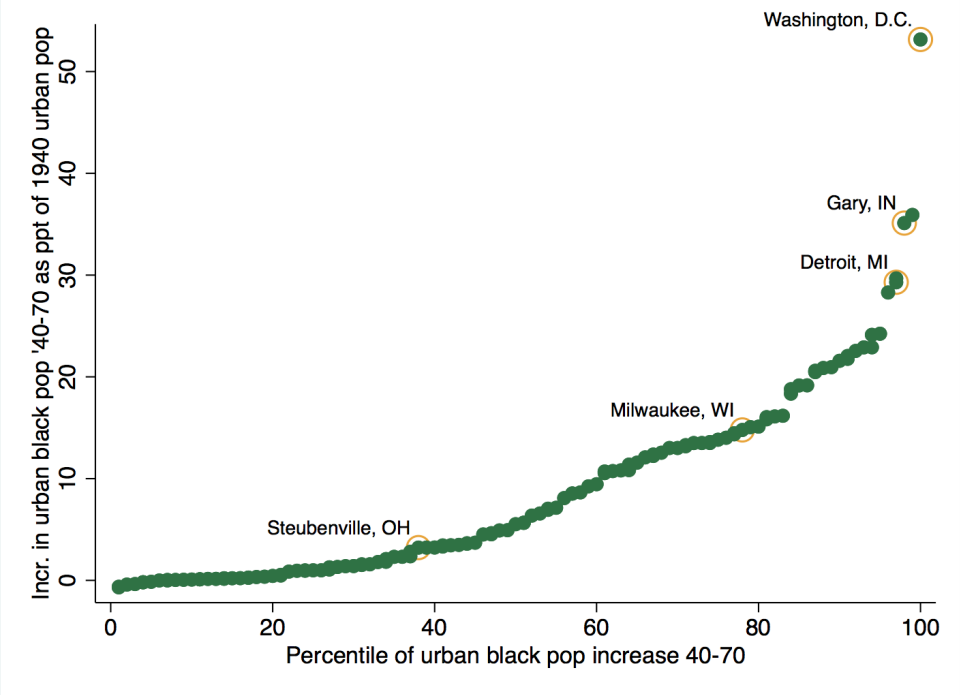
(a) Fraction of black teens from median educated families with 9-plus years of schooling, 1940



(b) Household income rank of black men and women from below median income families, 2015

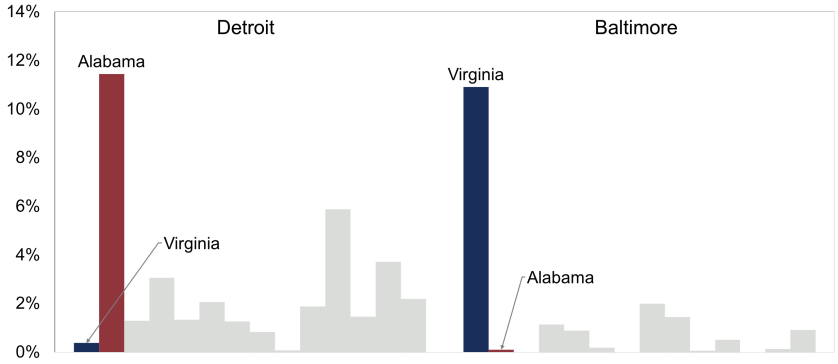
Notes: This figure depicts the geographic patterns in black upward mobility in 1940 and 2015. Panel (a) depicts black educational upward mobility in 1940 defined as fraction of 14-17 year-old boys and girls who have at least 9 years of schooling, from households where the household head has between 5 and 8 years of schooling. Panel (b) shows expected mean household income rank in 2015 by childhood commuting zone for 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) and Chetty, Hendren, Jones, and Porter (2018) for panel (b).

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

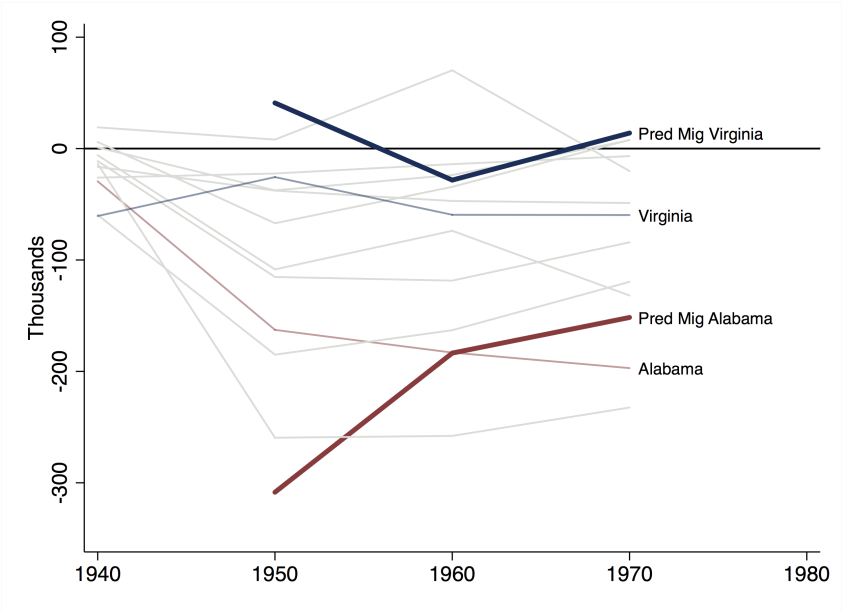


Notes: This figure plots the quantile function of 1940-1970 increases in the urban black population in commuting zones as a share of the total initial 1940 urban population and multiplied by 100, so that the units are in percentage points. The CZs in sample are those containing the 294 non-southern mainland cities with information on the black population in both 1940 and 1970 from the *City and County Data Books, 1944-1977*. ''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: Shift-share Instrument for Great Migration



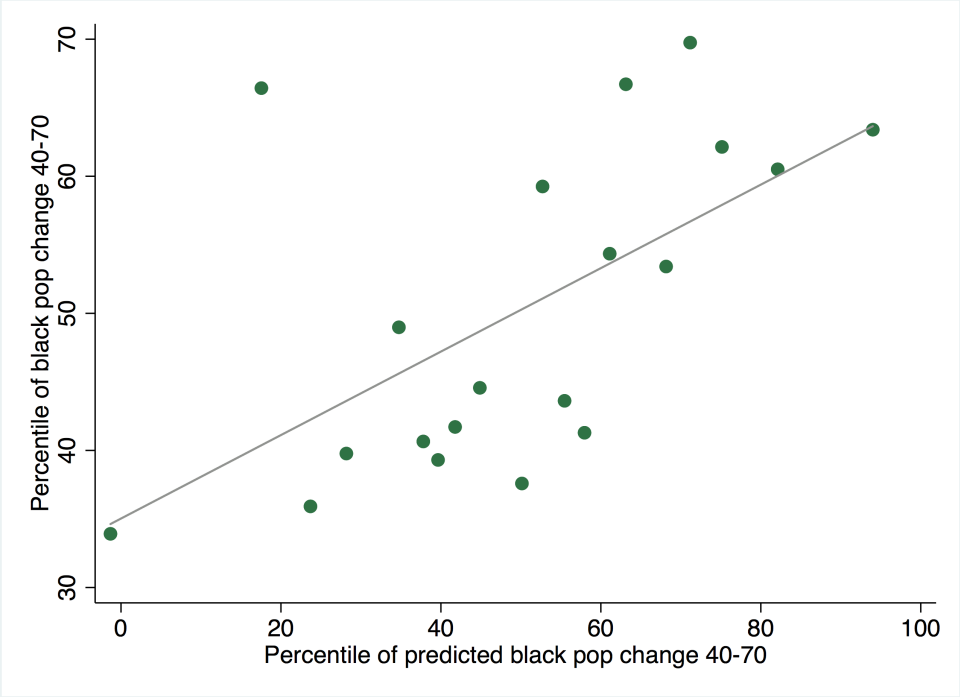
(a) Composition of recent 1935-1940 black southern migrants in Detroit vs. Baltimore



(b) Southern state net-migration, 1910-1970

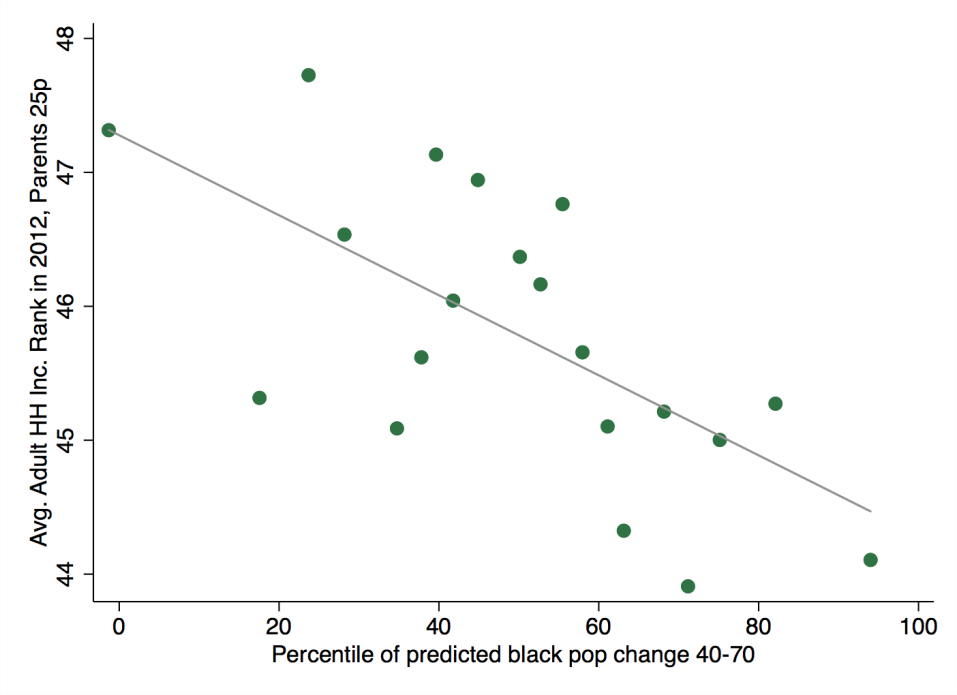
Notes: This figure depicts the variation underlying the shift-share instrument for urban black population change in northern commuting zones in the case of Detroit and Baltimore. Panel (a) shows the share of recent black southern migrants (those who migrated between 1935 and 1940) living in Detroit and Baltimore in 1940 by 1935 county, for the largest sending county in each southern state. For Alabama and Virginia, these are Jefferson County (Birmingham) and Richmond City County, respectively. Detroit receives the plurality of its migrants from Alabama while Baltimore receives the plurality from Virginia. Panel (b) shows net-migration and predicted net-migration for southern states each decade from 1910-1970, with net-migration figures for Alabama and Virginia highlighted. Negative numbers indicate outmigration. In darker lines are net-migration figures predicted using one-decade lagged southern county agricultural and World War II spending measures. Appendix C describes the construction of the instrument based on this variation. I use LASSO to select predictors each decade, interacting predicted migration with share of recent black southern migrants from each county, and summing up over all southern counties to construct counterfactual increases in the urban black population from 1940-1970 using variation in black southern migration alone. *Data sources*: IPUMS complete count US census for 1940; Boustan (2016) and Foukas et al. (2018).

Figure 4: First Stage on Black Population Change



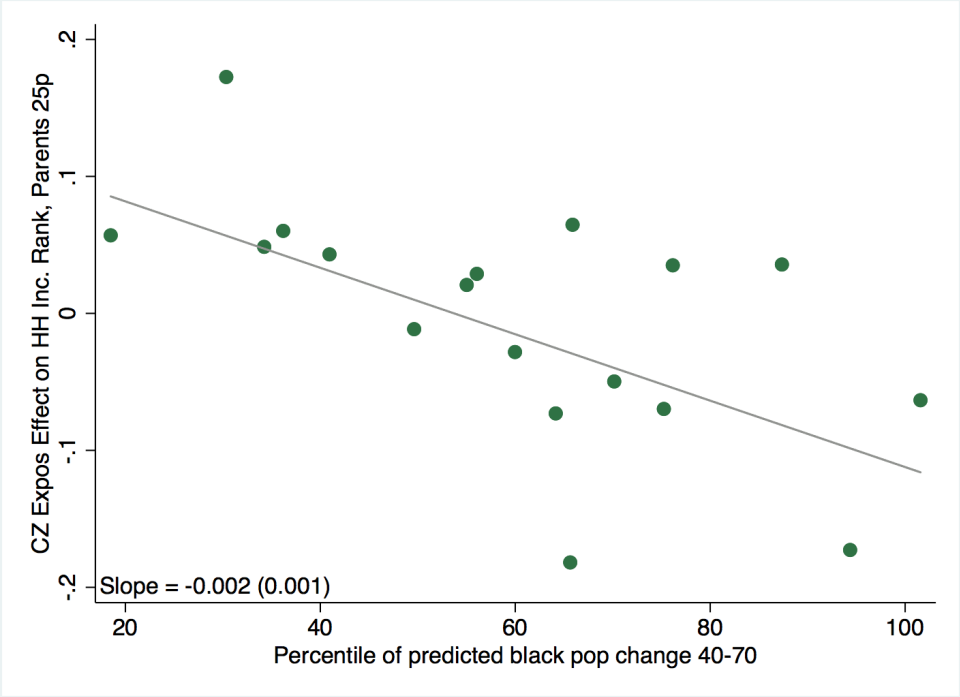
Notes: This binned scatterplot depicts the relationship between the percentile of actual black population increase during the Great Migration (1940 to 1970) for northern commuting zones and the instrument for black population increase over the same period. The instrument is the percentile of predicted 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 US census for 1940; Boustan (2016).

Figure 5: Great Migration reduced upward mobility for low income families in northern commuting zones



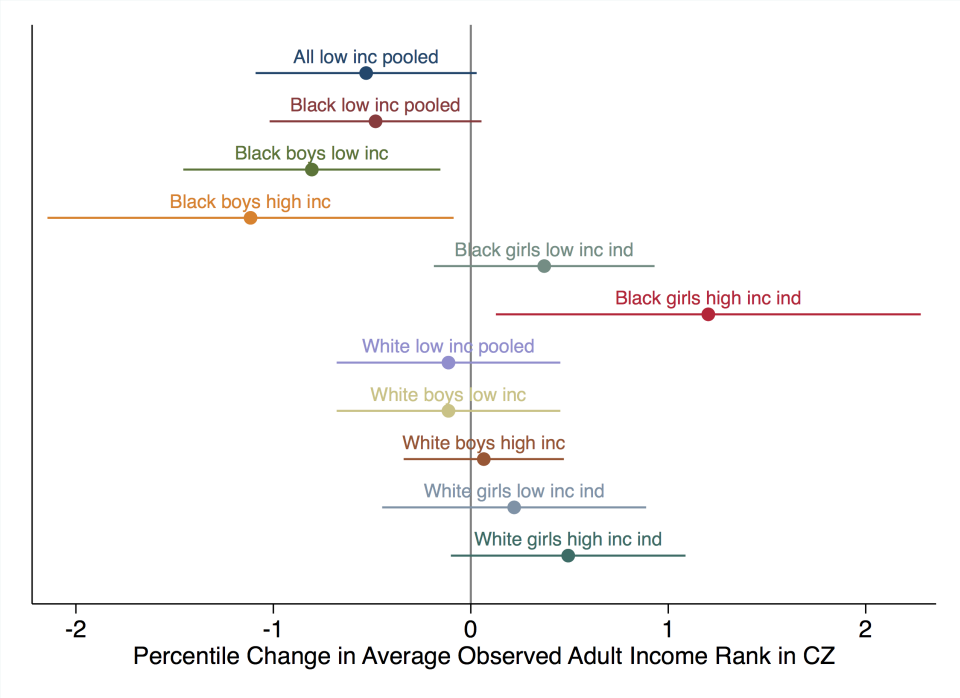
Notes: This binned scatterplot depicts the relationship between average upward mobility in the 2000s for men and women with low income parents and the instrument for black population increases during the Great Migration. The unit of observation is a commuting zone. The right hand side variable is grouped 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 US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

Figure 6: Childhood exposure to Great Migration CZs reduced adult income of children from low income families



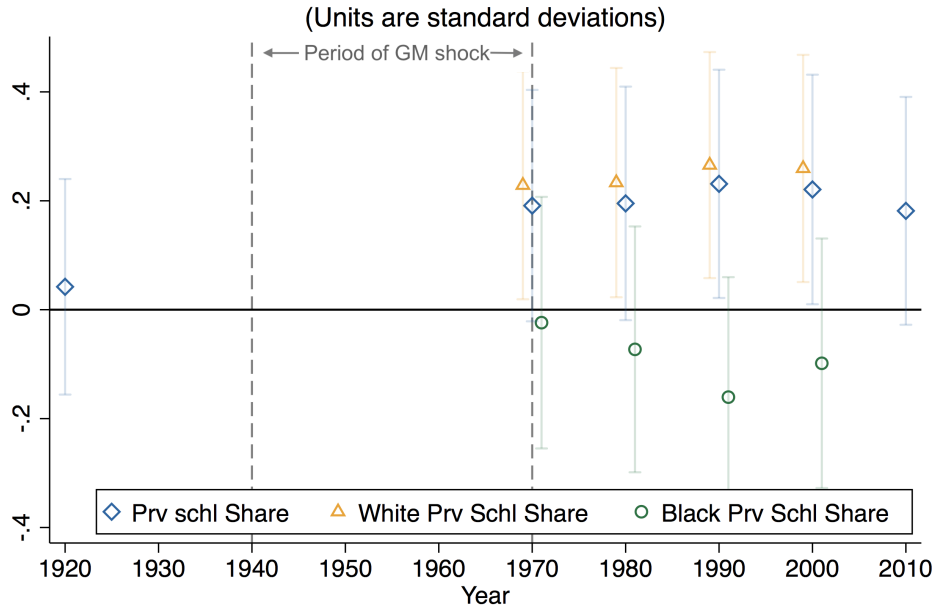
Notes: This binned scatterplot depicts the relationship between commuting zone childhood exposure effects in the 2000s for men and women with low income parents and the instrument for black population increases during the Great Migration. The unit of observation is a commuting zone. The right hand side 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 US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

Figure 7: Race and gender heterogeneity in impact of Great Migration on upward mobility



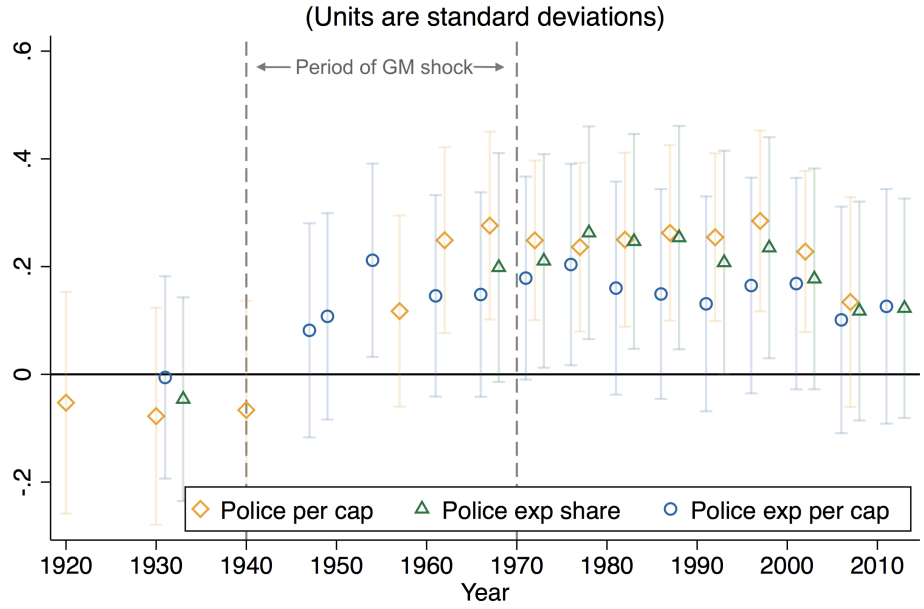
Notes: This figure plots coefficients from regressions of average upward mobility in the 2000s for men and women from low and high income parents on the instrument for black population increases during the Great Migration, in approximately one standard deviation units. The unit of observation is a commuting zone. Upward mobility is defined as expected mean individual or 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 US census for 1940; Boustan (2016); Chetty et al. (2018).

Figure 8: Great Migration impact on private schooling



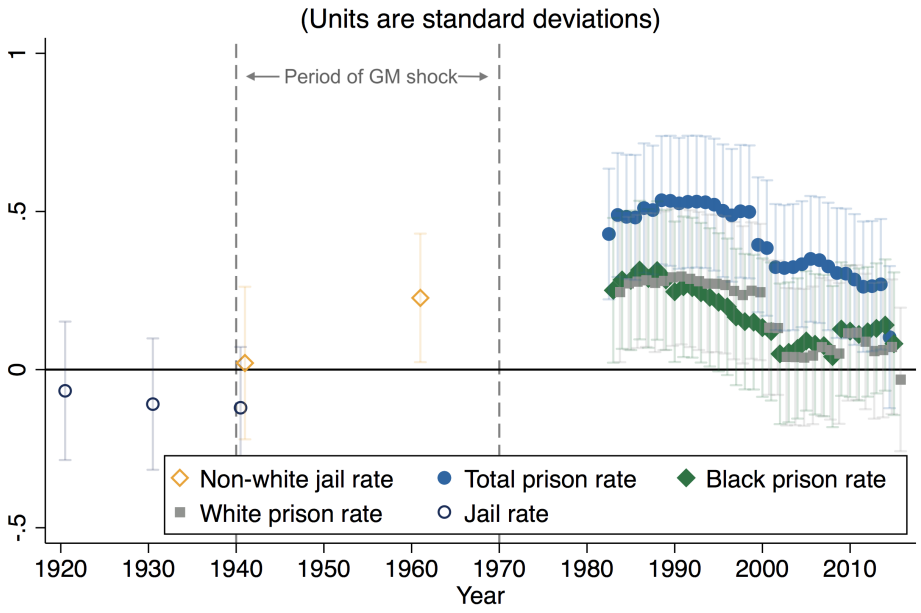
Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions for each year where the dependent variable is private school enrollment rates. The unit of observation is a commuting 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 US census for 1940; Boustan (2016).

Figure 9: Great Migration impact on policing investments



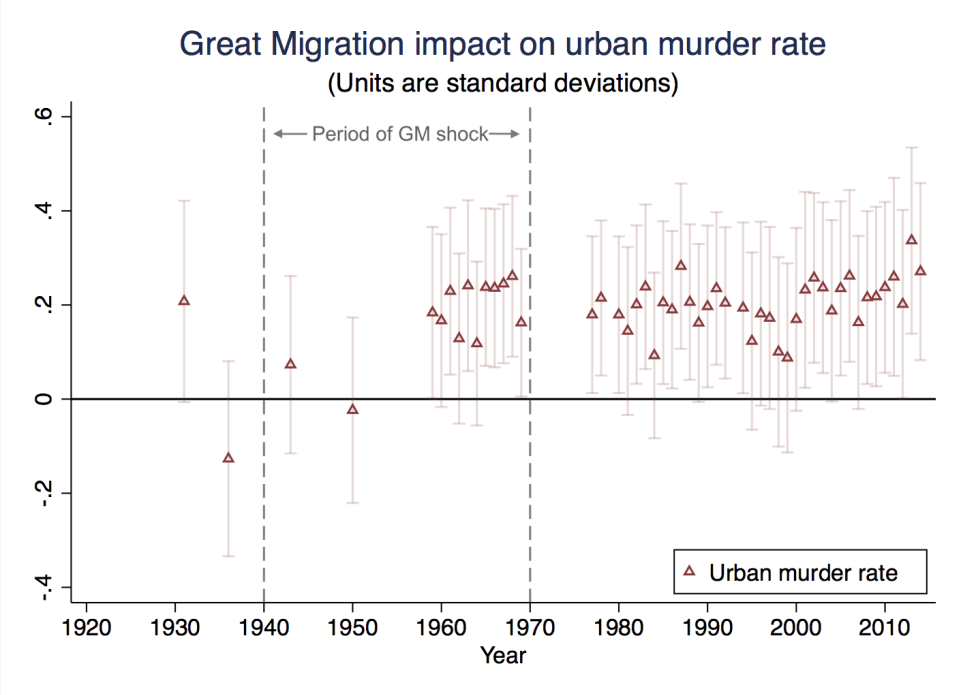
Notes: This figure plots the coefficient on the instrument for black population increases during the Great Migration, in approximately one standard deviation units, in separate regressions for each year where the dependent variable is either the share of local government expenditures on policing, police 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 US census for 1940; Boustan (2016).

Figure 10: 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 US census for 1940; Boustan (2016).

Figure 11: 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 US census for 1940; Boustan (2016).

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

	School Attendance			Median education
	1920	1930	1940	1940
GM Shock	-0.006	-0.007	0.005	-0.011
	(0.023)	(0.024)	(0.011)	(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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: IPUMS complete count US census for 1940; Boustan (2016).

Table 2: Robustness of results on childhood exposure effects

$\hat{GM}$	-0.00144***	-0.00266***	-0.00242***	-0.00204***	-0.00233***	-0.00220***	-0.00261***	-0.00246***
	(0.000468)	(0.000506)	(0.000628)	(0.000637)	(0.000673)	(0.000632)	(0.000645)	(0.000619)
R-squared	0.0684	0.227	0.249	0.283	0.320	0.291	0.258	0.278
N	130	130	130	130	130	130	130	130
Precision Wt	Y	Y	Y	Y	Y	Y	Y	Y
Census Div FE	N	N	Y	Y	Y	Y	Y	Y
Baseline Controls	N	Y	Y	Y	Y	Y	Y	Y
Pop 1940	N	N	N	Y	N	N	N	N
Black Shares 1920-1940	N	N	N	N	Y	N	N	N
1940 Black Share Quartile FEs	N	N	N	N	N	Y	N	N
Recent WS Mig Share 1940	N	N	N	N	N	N	Y	N
FB White Share 1940	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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: IPUMS complete count US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

Table 3: Robustness of results on black men's upward mobility

GM Shock	-0.000380***	-0.000366***	-0.000248**	-0.000292**	-0.000266**	-0.000160	-0.000290**	-0.000263**
	(0.0000884)	(0.0000927)	(0.000113)	(0.000114)	(0.000122)	(0.000111)	(0.000118)	(0.000107)
R-squared	0.127	0.138	0.185	0.208	0.187	0.258	0.196	0.275
N	129	129	129	129	129	129	129	129
Precision Wt	Y	Y	Y	Y	Y	Y	Y	Y
Baseline Controls	N	N	Y	Y	Y	Y	Y	Y
Census Div FE	N	Y	Y	Y	Y	Y	Y	Y
Pop 1940	N	N	N	Y	N	N	N	N
Black Shares 1920-1940	N	N	N	N	Y	N	N	N
Black Share Rank 1940	N	N	N	N	N	Y	N	N
Recent WS Mig Share 1940	N	N	N	N	N	N	Y	N
FB White Share 1940	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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: IPUMS complete count US census for 1940; Boustan (2016); Chetty et al. (2018).

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

	Ordinary Least Squares					
	Pooled HH Lower Inc	Women HH Lower Inc	Men HH Lower Inc	Pooled Ind Lower Inc	Women Ind Lower Inc	Men Ind Lower Inc
GM	-0.0557***	-0.0455***	-0.0660***	-0.0203*	0.0128	-0.0526***
	(0.0100)	(0.0104)	(0.0103)	(0.0111)	(0.0143)	(0.0108)
R-squared	0.563	0.503	0.595	0.307	0.184	0.493
	Reduced Form					
	Pooled HH Lower Inc	Women HH Lower Inc	Men HH Lower Inc	Pooled Ind Lower Inc	Women Ind Lower Inc	Men Ind Lower Inc
GM Shock	-0.0299***	-0.0223**	-0.0374***	-0.0184*	0.0000493	-0.0371***
	(0.00952)	(0.00964)	(0.00996)	(0.00978)	(0.0127)	(0.00984)
R-squared	0.493	0.449	0.515	0.308	0.178	0.457
	Two-stage least squares					
	Pooled HH Lower Inc	Women HH Lower Inc	Men HH Lower Inc	Pooled Ind Lower Inc	Women Ind Lower Inc	Men Ind Lower Inc
GM	-0.0981***	-0.0733**	-0.123***	-0.0606*	0.000162	-0.122***
	(0.0301)	(0.0300)	(0.0324)	(0.0328)	(0.0402)	(0.0350)
R-squared	0.498	0.474	0.494	0.232	0.178	0.321
N	130	130	130	130	130	130
Outcome Mean	45.79	47.04	44.55	45.54	42.74	48.29
Outcome SD	3.379	3.283	3.617	2.972	3.527	3.375

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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

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

	Ordinary Least Squares					
	Pooled HH Higher Inc	Women HH Higher Inc	Men HH Higher Inc	Pooled Ind Higher Inc	Women Ind Higher Inc	Men Ind Higher Inc
GM	-0.0340***	-0.0293***	-0.0387***	-0.00658	0.0116	-0.0239***
	(0.00792)	(0.00793)	(0.00822)	(0.00816)	(0.0106)	(0.00810)
R-squared	0.528	0.513	0.533	0.474	0.425	0.465
	Reduced Form					
	Pooled HH Higher Inc	Women HH Higher Inc	Men HH Higher Inc	Pooled Ind Higher Inc	Women Ind Higher Inc	Men Ind Higher Inc
GM Shock	-0.0109	-0.00807	-0.0138*	-0.000766	0.00961	-0.0109
	(0.00743)	(0.00734)	(0.00778)	(0.00722)	(0.00936)	(0.00733)
R-squared	0.466	0.463	0.463	0.471	0.424	0.437
	Two-stage least squares					
	Pooled HH Higher Inc	Women HH Higher Inc	Men HH Higher Inc	Pooled Ind Higher Inc	Women Ind Higher Inc	Men Ind Higher Inc
GM	-0.0359	-0.0265	-0.0452*	-0.00252	0.0316	-0.0359
	(0.0222)	(0.0223)	(0.0231)	(0.0229)	(0.0302)	(0.0229)
R-squared	0.528	0.512	0.531	0.473	0.408	0.455
N	130	130	130	130	130	130
Outcome Mean	58.82	60.40	57.28	57.95	55.39	60.44
Outcome SD	2.570	2.533	2.684	2.510	3.118	2.470

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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

Table 6: Great Migration impact on average upward mobility for black families in the 2000s

	First Stage				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
$\hat{GM}$	0.304***	0.304***	0.304***	0.304***	0.304***
	(0.0750)	(0.0750)	(0.0750)	(0.0750)	(0.0750)
F-Stat	23.32	23.32	23.32	23.32	23.32
	Ordinary Least Squares				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM	-0.0532***	-0.0386***	-0.0723***	-0.0437**	-0.101***
	(0.00959)	(0.0113)	(0.0113)	(0.0215)	(0.0174)
R-squared	0.416	0.274	0.445	0.203	0.357
	Reduced Form				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
$\hat{GM}$	-0.0165*	-0.0114	-0.0247**	-0.00476	-0.0393**
	(0.00921)	(0.0102)	(0.0112)	(0.0190)	(0.0167)
R-squared	0.286	0.212	0.288	0.176	0.214
	Two-stage least squares				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM	-0.0516**	-0.0358	-0.0773***	-0.0149	-0.123***
	(0.0253)	(0.0298)	(0.0299)	(0.0572)	(0.0463)
R-squared	0.416	0.273	0.444	0.191	0.349
N	129	129	129	129	129
Outcome Mean	33.19	35.22	31.21	46.67	44.21
Outcome SD	2.747	2.900	3.335	5.281	4.757

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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty et al. (2018).

Table 7: Great Migration impact on average upward mobility for white families in the 2000s

	First Stage				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
$\hat{GM}$	0.304***	0.304***	0.304***	0.304***	0.304***
	(0.0750)	(0.0750)	(0.0750)	(0.0750)	(0.0750)
F-Stat	23.32	23.32	23.32	23.32	23.32
	Ordinary Least Squares				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM	0.000760	0.00678	-0.00356	-0.00678	-0.0137
	(0.0127)	(0.0140)	(0.0117)	(0.00888)	(0.00835)
R-squared	0.197	0.186	0.209	0.241	0.312
	Reduced Form				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
$\hat{GM}$	0.00483	0.00748	0.00305	0.00881	0.00341
	(0.0112)	(0.0123)	(0.0103)	(0.00781)	(0.00744)
R-squared	0.198	0.187	0.209	0.245	0.298
	Two-stage least squares				
	Pooled Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM	0.0159	0.0246	0.0100	0.0289	0.0112
	(0.0359)	(0.0395)	(0.0329)	(0.0265)	(0.0243)
R-squared	0.187	0.175	0.200	0.141	0.262
N	130	130	130	130	130
Outcome Mean	45.23	46.27	44.27	61.78	59.58
Outcome SD	3.163	3.459	2.926	2.273	2.246

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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty et al. (2018).

Table 8: Great Migration impact on childhood exposure effects in the 2000s for low income families

	First Stage on GM					
GM Shock	0.292***	0.289***	0.294***	0.290***	0.288***	0.294***
	(0.0652)	(0.0652)	(0.0658)	(0.0651)	(0.0655)	(0.0656)
F-Stat	27.46	26.82	26.97	27.72	27.19	27.22
	Household Income Rank			Individual Income Rank		
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.00282***	-0.00203**	-0.00404***	-0.00167***	0.000591	-0.00336***
	(0.000582)	(0.000861)	(0.000867)	(0.000590)	(0.000857)	(0.000920)
R-squared	0.206	0.106	0.227	0.176	0.0190	0.198
	Reduced Form					
GM Shock	-0.00242***	-0.00218**	-0.00341***	-0.00203***	-0.000927	-0.00315***
	(0.000628)	(0.000928)	(0.000963)	(0.000641)	(0.000939)	(0.00102)
R-squared	0.249	0.137	0.229	0.213	0.0381	0.196
	Two-stage least squares					
GM	-0.00830***	-0.00756**	-0.0116***	-0.00701***	-0.00322	-0.0107***
	(0.00250)	(0.00343)	(0.00357)	(0.00245)	(0.00330)	(0.00363)
R-squared	-0.0774	-0.0486	0.0250	-0.0259	-0.0490	0.0709
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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

Table 9: Great Migration impact on childhood exposure effects in the 2000s for high income families

	First Stage on GM					
GM Shock	0.302***	0.300***	0.302***	0.301***	0.301***	0.302***
	(0.0632)	(0.0633)	(0.0633)	(0.0631)	(0.0635)	(0.0632)
F-Stat	27.69	27.32	27.41	28.00	27.74	27.69
	Household Income Rank			Individual Income Rank		
	Pooled	Women	Men	Pooled	Women	Men
	Ordinary Least Squares					
GM	-0.00132**	-0.00110	-0.00188**	-0.000128	0.000958	-0.00115
	(0.000587)	(0.000808)	(0.000770)	(0.000550)	(0.000830)	(0.000784)
R-squared	0.252	0.196	0.133	0.466	0.378	0.175
	Reduced Form					
GM Shock	-0.00127**	-0.000742	-0.00144*	-0.00192***	-0.00131	-0.00226***
	(0.000599)	(0.000829)	(0.000823)	(0.000564)	(0.000890)	(0.000822)
R-squared	0.324	0.260	0.141	0.515	0.385	0.214
	Two-stage least squares					
GM	-0.00421**	-0.00247	-0.00476*	-0.00638***	-0.00434	-0.00747***
	(0.00205)	(0.00274)	(0.00266)	(0.00229)	(0.00317)	(0.00289)
R-squared	0.232	0.220	0.125	0.228	0.246	0.0582
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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

Table 10: Great Migration impact on 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	0.00266***	0.188***	1.026***	0.0255***	0.175***	0.0103***
	(0.000426)	(0.0243)	(0.119)	(0.00381)	(0.0254)	(0.00151)
R-squared	0.296	0.397	0.583	0.290	0.423	0.292
	Reduced Form					
GM Shock	0.00117**	0.0661**	0.446***	0.0155***	0.0991***	0.00654***
	(0.000531)	(0.0282)	(0.138)	(0.00523)	(0.0311)	(0.00207)
R-squared	0.307	0.484	0.646	0.150	0.454	0.162
	Two-stage least squares					
GM	0.00300**	0.169**	1.138***	0.0395***	0.253***	0.0167***
	(0.00131)	(0.0702)	(0.351)	(0.0116)	(0.0784)	(0.00461)
R-squared	0.310	0.477	0.627	0.317	0.430	0.317
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, 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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: Collins and Margo (2007). Great Migration data sources: CCDB, IPUMS complete count US census for 1940; Boustan (2016).

References

Adão, R., Kolesár, M., and Morales, E. (2018). Shift-share designs: Theory and inference. Working Paper 24944, National Bureau of Economic Research.

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

Altonji, J. G. and Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pages 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., Casey, M., Hardy, B. L., and Logan, T. D. (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., Chernozhukov, V., and Hansen, C. (2011). Lasso methods for gaussian instrumental variables models.

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

Bhattacharya, D. and Mazumder, B. (2011). A nonparametric analysis of black–white differences in intergenerational income mobility in the united states. *Quantitative Economics*, 2(3):335–379.

Black, D. A., Sanders, S. G., Taylor, E. J., and Taylor, L. J. (2015). The impact of the Great Migration on mortality of African Americans: Evidence from the Deep South. *American Economic Review*, 105(2):477–503.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *Quarterly Journal of Economics*, 118(4):1335–1374.

Borjas, G. J., Freeman, R. B., and Katz, L. F. (1996). Searching for the effect of immigration on the labor market. Working paper, National Bureau of Economic Research.

Borusyak, K., Hull, P., and Jaravel, X. (2018). Quasi-experimental shift-share research designs. Working Paper 24997, National Bureau of Economic Research.

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.

Burchardi, K. B., Chaney, T., and Hassan, T. A. (2016). Migrants, ancestors, and investments. Working paper, National Bureau of Economic Research.

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64.

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

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

Case, A. C. and Katz, L. F. (1991). The company you keep: The effects of family and neighborhood on disadvantaged youths. Working paper, National Bureau of Economic Research.

Chay, K. and Munshi, K. (2012). Black networks after emancipation: evidence from reconstruction and the Great Migration. *Working Paper*.

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

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

Chetty, R., Hendren, N., Jones, M. R., and Porter, S. R. (2018). Race and economic opportunity in the united states: An intergenerational perspective. Working paper, National Bureau of Economic Research.

Chetty, R., Hendren, N., and Katz, L. F. (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.

Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? the geography of intergenerational mobility in the united states. *Quarterly Journal of Economics*, 129(4):1553–1623.

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 Margo, R. A. (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 Wanamaker, M. H. (2014). Selection and economic gains in the Great Migration of African Americans: New evidence from linked census data. *American Economic Journal: Applied Economics*, 6(1):220–252.

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

Cutler, D. M., Glaeser, E. L., and Vigdor, J. L. (1999). The rise and decline of the American Ghetto. *Journal of Political Economy*, 107(3):455–506.

Damm, A. P. and Dustmann, C. (2014). Does growing up in a high crime neighborhood affect youth criminal behavior? *American Economic Review*, 104(6):1806–32.

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

Dobbie, W., Grönqvist, H., Niknami, S., Palme, M., and Priks, M. (2018). The intergenerational effects of parental incarceration. Working paper, National Bureau of Economic Research.

Epple, D. and Romano, R. E. (1996). Ends against the middle: Determining public service provision when there are private alternatives. *Journal of Public Economics*, 62(3):297–325.

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

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

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

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

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2018). Bartik instruments: What, when, why, and how. Technical report, National Bureau of Economic Research.

Graham, B. (2016). Identifying and estimating neighborhood effects. NBER Working Papers 22575, National Bureau of Economic Research.

Gyourko, J., Saiz, A., and Summers, A. (2008). A new measure of the local regulatory environment for housing markets: The Wharton Residential Land Use Regulatory Index. *Urban Studies*, 45(3):693–729.

Heller, S. B., Shah, A. K., Guryan, J., Ludwig, J., Mullainathan, S., and Pollack, H. A. (2017). Thinking, fast and slow? Some field experiments to reduce crime and dropout in Chicago. *The Quarterly Journal of Economics*, 132(1):1–54.

Hilger, N. G. (2015). The Great Escape: Intergenerational Mobility in the United States since 1940. Working paper, National Bureau of Economic Research.

Hornbeck, R. and Moretti, E. (2018). Who Benefits From Productivity Growth? Direct and Indirect Effects of Local TFP Growth on Wages, Rents, and Inequality. Working paper, National Bureau of Economic Research.

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

Jackson, C. K., Johnson, R. C., and Persico, C. (2015). The effects of school spending on educational and economic outcomes: Evidence from school finance reforms. *Quarterly Journal of Economics*, 131(1):157–218.

Johnson, R. (2009). Ever-increasing levels of parental incarceration and the consequences for children. *Do prisons make us safer? The benefits and costs of the prison boom*, pages 177–206.

Kasy, M. (2015). Identification in a model of sorting with social externalities and the causes of urban segregation. *Journal of Urban Economics*, 85:16–33.

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

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

Liu, S. (2018). Incarceration of African American Men and the Impacts on Women and Children. Working paper.

Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R., and Sanbonmatsu, L. (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. (1991). Segregated Schools and the Mobility Hypothesis: A Model of Local Government Discrimination. *Quarterly Journal of Economics*, 106(1):61–73.

Massey, D. S. and Denton, N. A. (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 post-bellum US South. *Journal of Labor Economics*, 28(2):413–445.

Nix, E. and Qian, N. (2015). The fluidity of race: "Passing" in the United States, 1880-1940. Working paper, National Bureau of Economic Research.

Rothstein, J. (2018). Inequality of educational opportunity? schools as mediators of the intergenerational transmission of income. Working paper, National Bureau of Economic Research.

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

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

Shertzer, A. and Walsh, R. P. (2016). Racial sorting and the emergence of segregation in American cities. Working paper, National Bureau of Economic Research.

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

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

Sviatschi, M. M. (2018). Making a Narco: childhood exposure to illegal labor markets and criminal life paths. Working paper.

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

Tiebout, C. M. (1956). A pure theory of local expenditures. *Journal of Political Economy*, 64(5):416–424.

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. J. (2011). *When work disappears: The world of the new urban poor*. Vintage.

Wilson, W. J. (2012). *The truly disadvantaged: The inner city, the underclass, and public policy*. University of Chicago Press.

Appendix A The changing geography of black upward mobility

In this appendix, I present descriptive evidence on the shift in regional patterns of black upward mobility in the US over the 20th century. As late as 1940, upward mobility rates were much lower in the South while the North offered a clear pathway to intergenerational progress. Today, the geography of upward mobility for black children looks very different. Some of the best locations for upward mobility in 1940 are among the worst today. I argue that a portion of this change is due to black upward mobility rates declining in the North.[39] In this section, I document these changes using historical census data and contemporary measures of upward mobility by race. The descriptive results in this section motivate the main analysis of the causal impact the Great Migration had on upward mobility.

There are two parts to this descriptive analysis. First, I construct a series of children's outcomes by race and region that dates from 1880-2010 in order to show the evolution of racial gaps in the North versus the South over a long time span. This series reveals the following: greater racial equality in the North than in the South between 1880 and 1940; improvements in relative black outcomes in both regions over this same period; and the *re-emergence* of racial gaps in the North after 1940. In the second part of this section, I compare commuting zone level measures of upward mobility by race and gender in 1940 and 2015. I find complete convergence in regional differences in upward mobility for black Americans since 1940. Historical rates of black upward mobility are uncorrelated with today's. By contrast, these two measures are positively correlated for white children.

A.1 Racial gaps in children's educational attainment, 1880-2010

One of the key limitations in constructing a long time series of upward mobility in the US is the availability of data that captures both the outcomes of children and the economic status of their parents. Prior to 1940, measures of income and educational attainment are not available in the US Census, though literacy rates are.[40] Beginning in 1940 and after,

[39] Black children's outcomes in the South improve dramatically over this time period, also contributing to the changing regional pattern in upward mobility.

[40] Matched samples of fathers and sons can be constructed from pre-1950 censuses using first and last name and state of birth, allowing one to construct measures of the intergenerational transmission of occupation score or literacy prior to 1940 and income and educational transmission in 1940. However, matching techniques typically do not allow daughters to be matched across censuses, and match

income becomes available, and more detailed information on educational attainment supplants indicators of literacy, reflecting real progress in population education levels.[41]

One measure of educational attainment that can be constructed from 1880 onwards, allows for conditioning on parent economic status, and remains economically significant today is teen school attendance.[42] Figure A1 shows black and white teen school attendance, separately by region. Each series shows uninterrupted increases over the 1880 to 1940 period. From 1940 to 1970, however, the growth in black teen attendance rates slows substantially. This stagnation has implications for the racial gap in teen school attendance rates by region.

Figure A2 shows the unadjusted white-black school attendance gap for teenagers aged 14-17 from 1880 to 2010,[43] estimated from the linear probability model in equation 11 below:

$$\text{Attend}_{irzt} = \sum_{1880}^{2010} \beta_t \text{North}_i \times \text{Black}_i \times \delta_t + \delta_r + \delta_z + \varepsilon_{irzt} \tag{11}$$

The outcome Attend_{irzt} is a binary indicator of school attendance for teen i of race r in region z (where $z = \{\text{North}, \text{South}\}$), and decade t.

During the first part of the 20th century, racial gaps in educational attainment are much higher in the South than in the North, though this measure likely understates the true gap.[44] Using this procedure, I estimate the gap in teen school attendance to be about 20 percentage points in the south and 10 percentage points in the North, thus gaps are about half as large in the North. In 1900, this difference in racial inequality peaks, with the north having a 15 percentage point smaller gap in teen school attendance. The next several decades show both continual improvements in the racial teen school attendance gap and substantial southern catch-up.[45] By 1940, the racial gap in teen school attendance in the North has disappeared and the southern gap has shrunk to about 8 percentage points.

rates are notoriously worse for African Americans. For work documenting the changing geography of intergenerational occupation score mobility for white fathers and sons, see Tan (2018).

[41]The literacy rate for black Americans aged 14 and older improved from 30% in 1880 to 84% by 1930. See "120 Years of Literacy," `https://nces.ed.gov/naal/lit_history.asp`.

[42]These measures can be constructed for 1870 as well, but estimates of black school attendance by region for these years are noisy due to the very small number of black children living outside the south.

[43]These figures show attendance rates for boys and girls together. The patterns are similar when estimated for boys and girls separately.

[44]In the first part of the 20th century, de jure segregation of southern schools and the much lower quality of black schools generated large gaps in the outcomes of black and white children (Margo, 1990).

[45]The trend in the unadjusted gap shows a continued improvement in the south, which experienced major improvements in educational access and quality for black children over this period.

The patterns in racial inequality after 1940 are of great interest. Between 1940 and 1960, the apparent northern premium in black educational attainment relative to the South vanished. Although improvements in relative black outcomes in the South over this period play a role, what is remarkable is the re-emergence of the racial gap in teen school attendance in the North. This re-emergence holds true even when conditioning on parent socioeconomic status. Including household head occupation score and the state of birth of both parents and children does not alter the basic post 1940 upward trend in racial inequality in teen school attendance in the North.[46]

A.2 Long run change, 1940-2015

I provide an additional piece of evidence in changing social mobility patterns for black children. I correlate a more detailed measure 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 black teenagers from households in which the household head has 5-8 years of schooling[47] who obtain at least 9 years of education. I then correlate this measure with a measure of income upward mobility in 2015 for children from different racial backgrounds. The measure 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.[48] Section 3.1 describes these data in much greater detail.

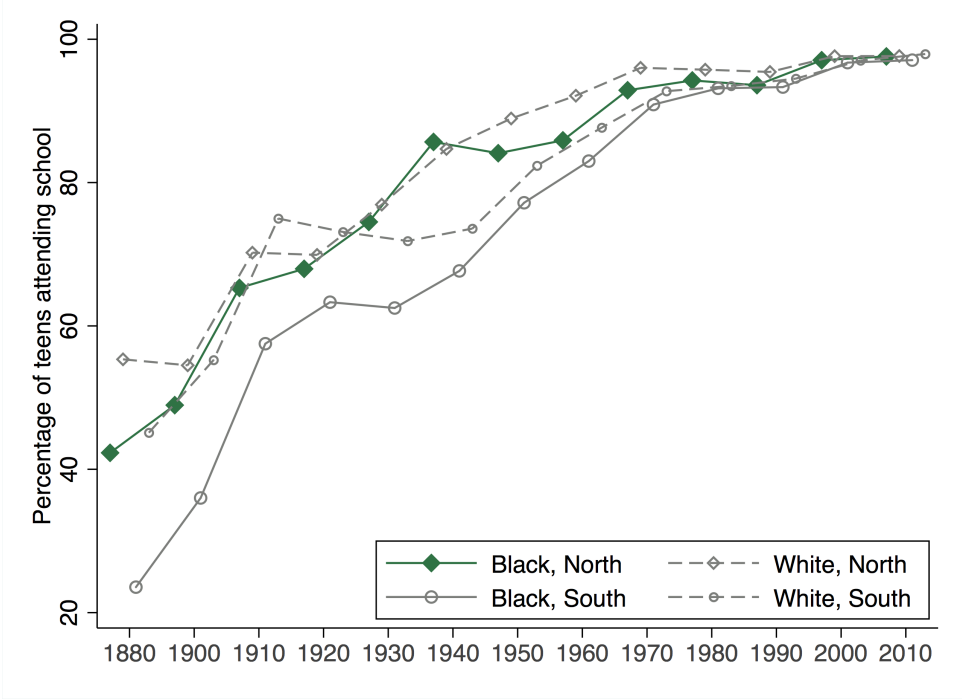
Figure 1b illustrates these two measures on a map of US commuting zones. The top panel shows the historical measure while the bottom panel shows the contemporary measure. Darker green represents higher upward mobility while lighter green represents lower upward mobility. In Table A1, I report the correlation coefficients between historical and contemporary upward mobility measures separately by race and gender. For white men and women, historically 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 A4 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.

[46] The regional gaps in racial inequality resulting from this alternative estimation strategy are smaller, but the broad pattern of increasing gaps in the North after 1940 remains true. The adjusted gaps show increases in racial inequality in both the South and the North after 1940, but as early as 1950, the gap in the North exceeds that of the South.

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

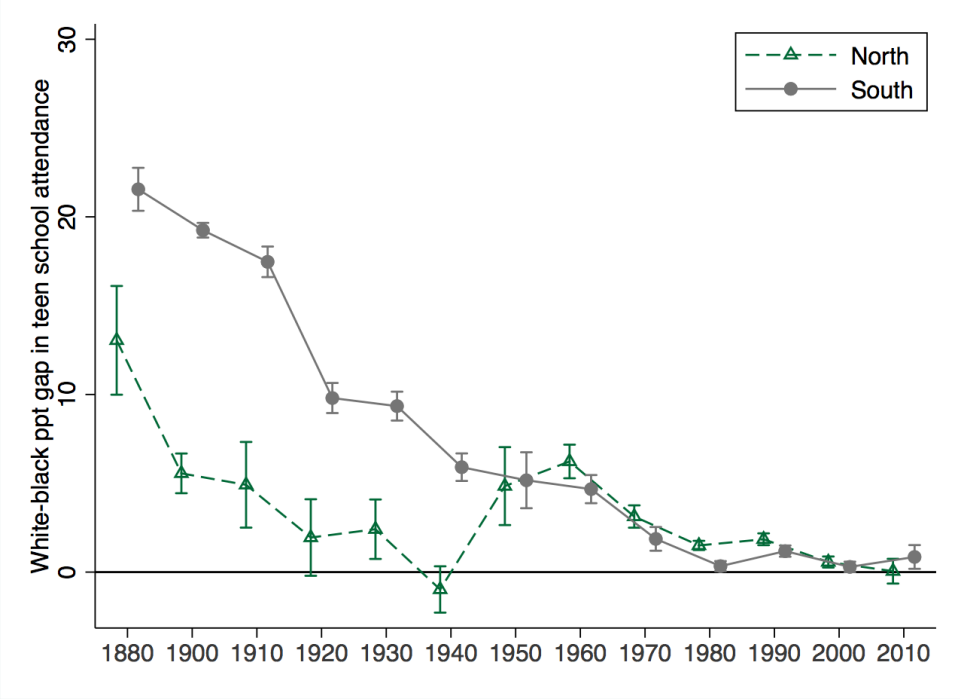
[48] The children come from 1980s birth cohorts[49].

Figure A1: Teen school attendance by subgroup



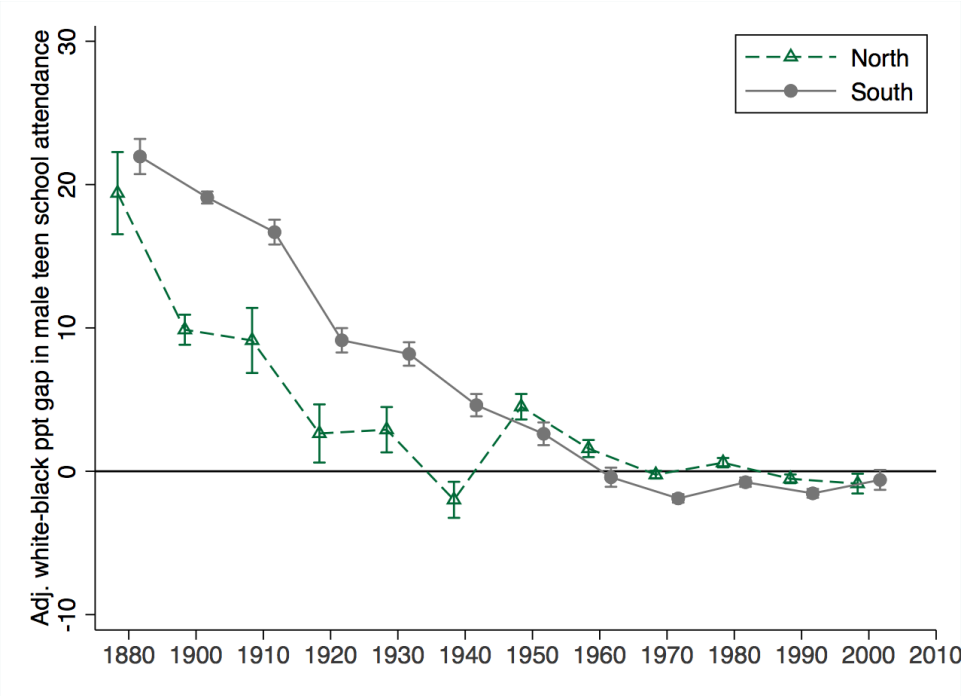
Notes: Teen school attendance by region (South and Non-South ("North")). The sample is 14-17 year-old boys and girls living in households. *Source*: IPUMS.

Figure A2: White-Black Gap in Teen School Attendance



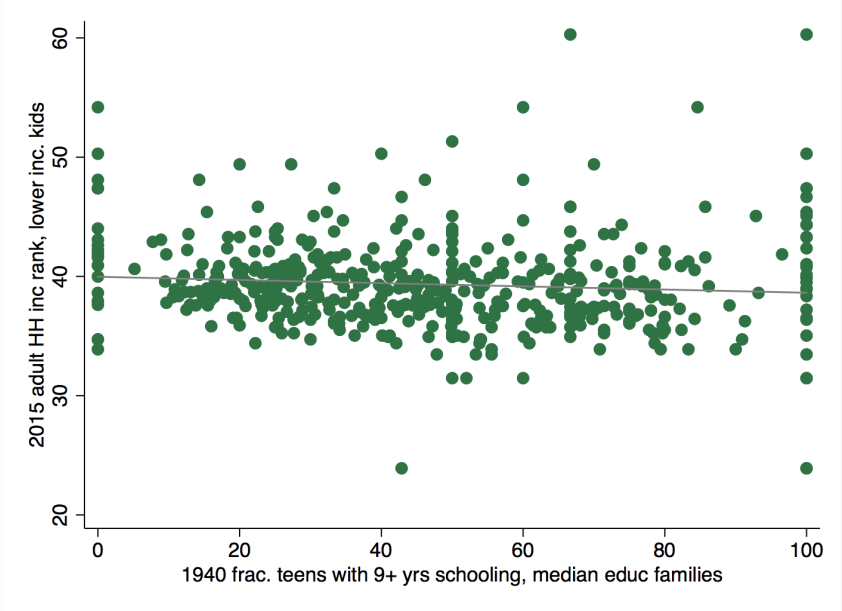
Notes: This figure shows the unadjusted racial gap in teen school attendance by region (South and Non-South ("North")). Plotted above are yearly regression coefficients from estimating equation 11 with the mean effect of black added in and multiplied by -1 to generate the white-black gap in attendance. The sample is 14-17 year-old boys and girls living in households. Appendix Figure A3 depicts the series adjusted for household head occupation score and birth state as well as teen birth state. Appendix Figure A1 shows the raw black and white teen attendance rates separately by gender. *Data sources*: IPUMS.

Figure A3: White-black gap in teen school attendance, adjusted



Notes: Racial gap in teen school attendance by region (South and Non-South ("North")) adjusted for household head occupation score and birth state as well as teen birth state. The sample is 14-17 year-olds living in households. *Source*: IPUMS.

Figure A4: Correlation of 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 between 1940 educational upward mobility defined as fraction of 14-17 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 in 2015 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. The correlations between historical and contemporary upward mobility are reported for black and white women in Appendix Table A1. *Data sources*: IPUMS for 1940 measure and Chetty, Hendren, Jones, and Porter (2018) for 2015 measures.

Table A1: 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.

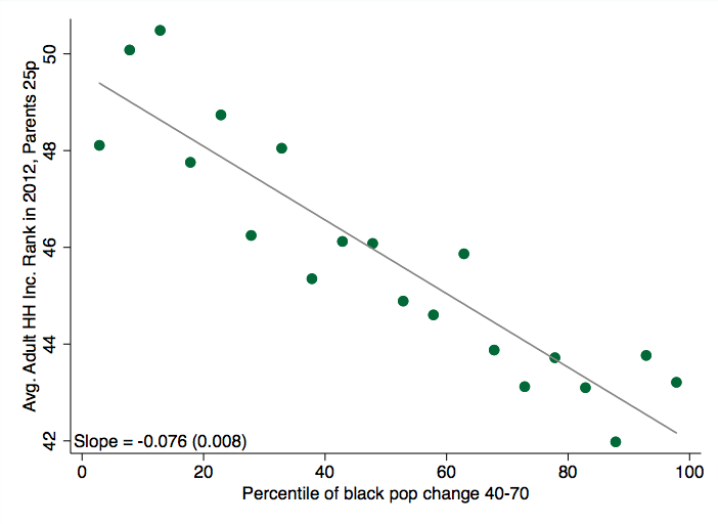
Appendix B Commuting Zones in Sample

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	

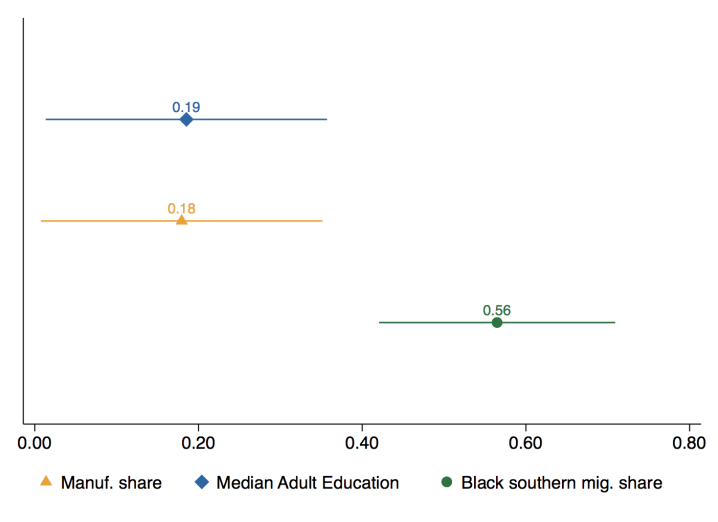
Notes: Name refers to largest city in the commuting zone.

B.1 Descriptive statistics

FIGURE B1: DESCRIPTIVE EVIDENCE AND SUMMARY STATISTICS



(a) Relationship between 1940-1970 black population change and upward mobility in 2012



(b) Correlation between black population change and baseline 1940 covariates

Notes: Panel (a) shows a binned scatterplot of the relationship between the percentile of actual black population increase between 1940 and 1970 and upward mobility in 2012. The right hand side variable is grouped into 20 bins (5 percentiles each). The unit of observation is a commuting zone. Panel (b) plots the correlation coefficient between percentile of black population increase between 1940 and 1970 and three baseline 1940 covariates: the share of the labor force in manufacturing, educational upward mobility in 1940 described in Section 3, and the share of the 1940 urban population made up of recent black southern migrants. *Data sources*: IPUMS complete count US census for 1940; CCDB.

Appendix C Shift-share instrument for black population increases during the Great Migration

To estimate the causal impact of the Great Migration on upward mobility in cities in the North today, I instrument for 1940-1970 urban black population increases (normalized by the 1940 urban population) in northern commuting zones using variation in southern black migration patterns. Specifically, I interact variation in location choices of black southern migrants who moved prior to 1940 with variation in net-migration from southern counties between 1940 and 1970 predicted using southern economic variables. 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.

C.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:

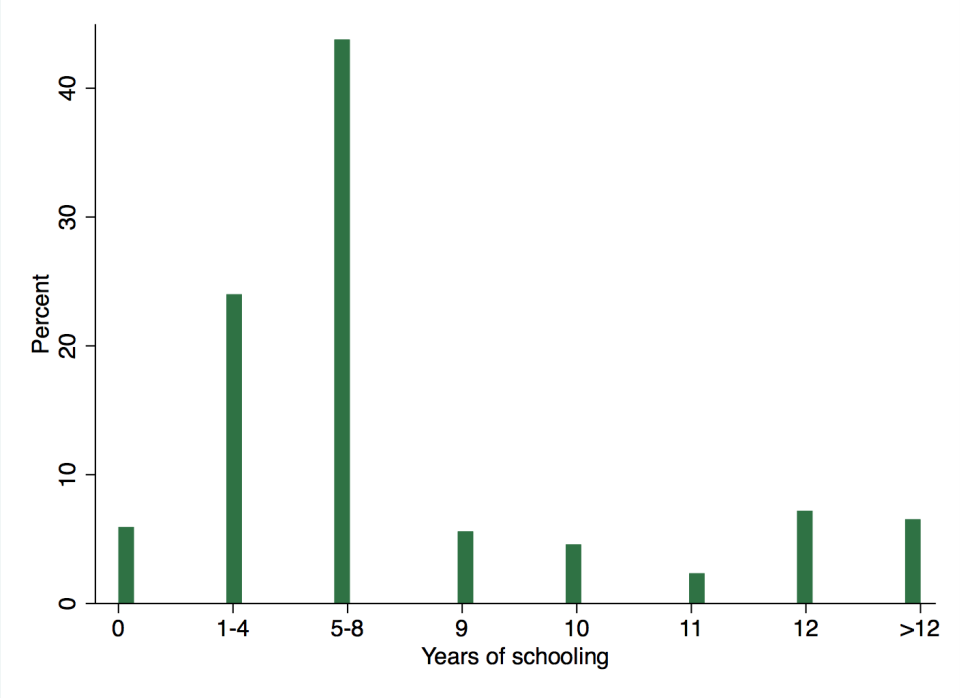
$$\omega_{jc}^{1935-1940} = \frac{b_{cj}}{b_j} \tag{12}$$

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 C3 depicts $\omega_{jc}^{1935-1940}$ for a select group of cities and southern counties. De-

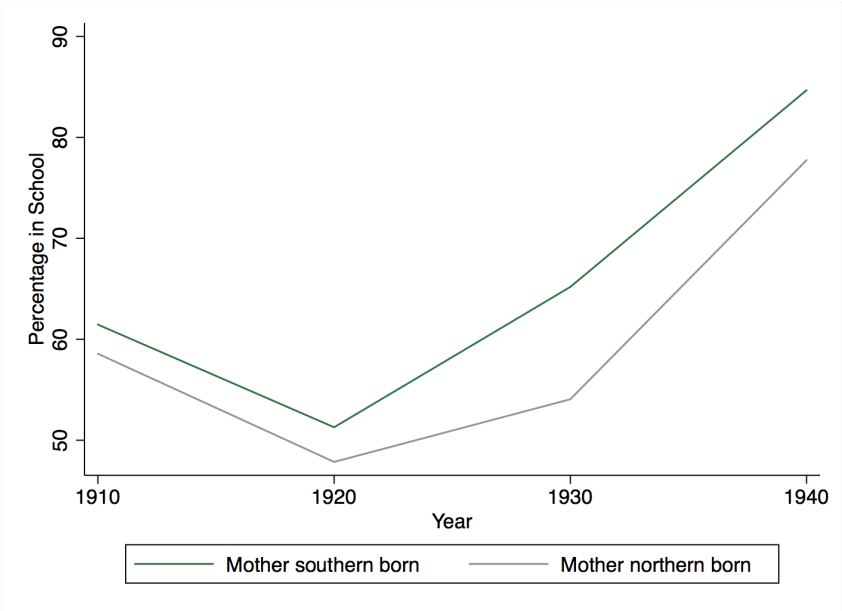
picted is the share of 1935-1940 black migrants from the largest sending county for each southern state who settled in the following cities: Baltimore, Boston, Chicago, Los Angeles, New York, Philadelphia, and Salt Lake City. The figure captures the immense heterogeneity in settlement patterns across and volume of migration into the cities in question. Figure C1 shows the educational distribution for 1935-1940 black southern migrants aged 25 plus.

FIGURE C1: 1935-1940 BLACK SOUTHERN MIGRANT EDUCATIONAL ATTAINMENT

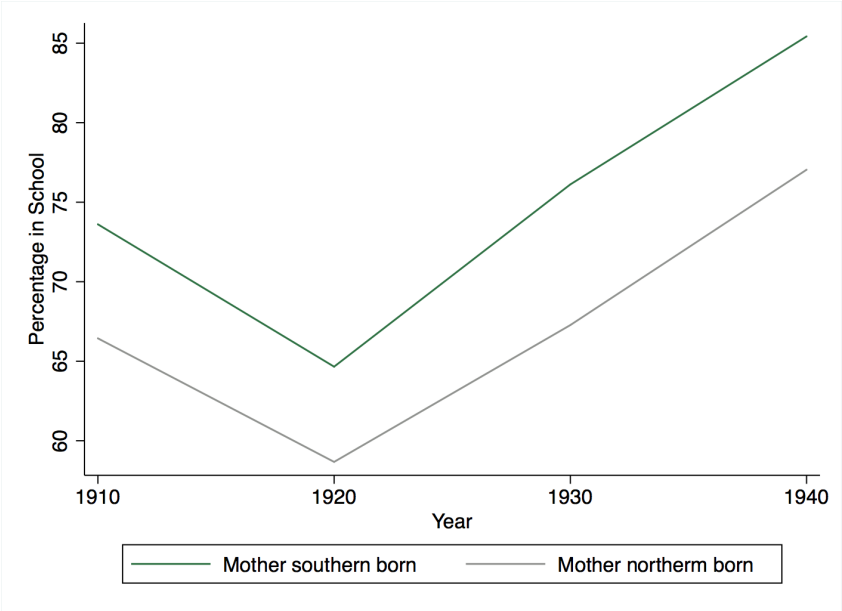


Notes: Histogram of years of schooling for 1935-1940 black southern migrants aged 25 plus. *Data sources*: IPUMS Complete Count 1940 US Census.

Figure C2: School attendance rates of black teens in the North with southern- vs. northern-born mothers



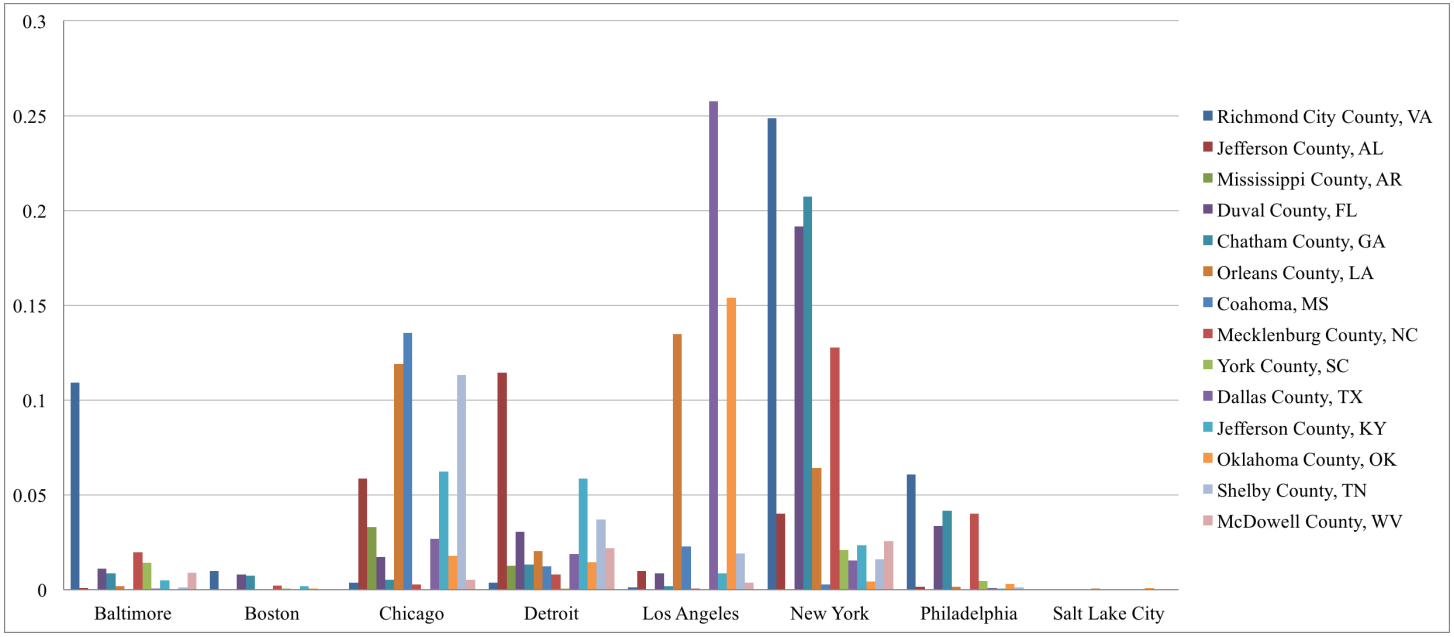
(a) Black teens with illiterate mothers



(b) Black teens with low occupation score fathers

Notes: 1910-1940 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.

Figure C3: Settlement patterns of 1935-1940 black southern migrants



Notes: Share of black individuals reporting the indicated counties (legend to the right of chart) as their county of residence in 1935 who were enumerated in the cities displayed in 1940. The counties included are those with the largest net migration from the state out of all other counties in the state. *Data sources*: IPUMS Complete Count 1940 US Census.

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

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

$$m_{jt} = \beta_0 + \mathbb{Z}'_{jt-10}\beta_1 + \varepsilon_{jt},$$
$$\hat{m}_{jt} = m_{jt} - \varepsilon_{jt}$$

where m_{jt} is net migration for southern county j between decade $t-10$ and t, $\mathbb{Z}'_{jt-10}$ is the set of predictors measured in decade $t-10$, and $\hat{m}_{jt}$ is predicted net migration from county j.[50] I then generate predicted migration into northern cities by multiplying the share of pre-1940 migrants from each county by the predicted number of migrants leaving that county between 1940 and 1970:[51]

$$\hat{m}_{ct} = \sum_{j=1,..,1223} (\omega_{cj}^{1935-40} \cdot \hat{m}_{jt}) \tag{13}$$

where $\omega_{cj}^{1935-40}$ is the share of black migrants from southern county j living in city c.

$$\sum_{j \in S} \sum_{t=1950}^{1970} \hat{m}_{jt \text{ urban},CZ} \tag{14}$$

Specifically, let m_{ct} be historical black in-migration into city c in decade t, and let ω_{cj} be the share of county j's outmigrants between 1935 and 1939 who reside in city c by 1940. Predicted in-migration $\hat{m}_{ct}$ is the sum of predicted outmigration from southern counties, weighted by ω_{cj}:

$$\hat{m}_{ct} = \sum_{j=1,..,1386} (\omega_{cj}^{1935-40} \cdot \hat{m}_{jt}) \text{ and}$$
$$\hat{m}_{c,t+10} = \hat{b}_{c,t} + \hat{m}_{c,t+10} \ \forall t > 1940.$$
$$\text{For } t = 1940, \ \hat{b}_{c,1940} = b_{c,1940}.$$

Under the assumption that county-level variation in southern economic indicators from 1940-1970 is uncorrelated with northern destination city characteristics for migrants

[50] 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).

[51] Because the available figures are net migration figures, and some southern counties experienced positive net migration (in-migration) as opposed to negative (in-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.

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.[52]

I choose the set of predictors by applying the Least Absolute Shrinkage and Selection Operator ("LASSO") algorithm to datasets of southern county net black migration estimates in 1950, 1960, 1970. For each decade, I use 5-fold cross-validation to choose the tuning parameter on the penalty term, the absolute value of the sum of the coefficients on each southern county covariate. I begin with a the set of predictors used in Boustan (2010) of net black migration rates from southern counties, where each predictor is measured in t and used to predict migration in $t + 10$. For migration from 1940-1950, 1950-1960, and 1960-1970, predictors are measured in 1940, 1950, and 1960, respectively.

Boustan (2010) uses the following variables in each year: the percent acreage in cotton; percent tenant farms; share of the labor force in agriculture; indicator 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.

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

[52] 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.

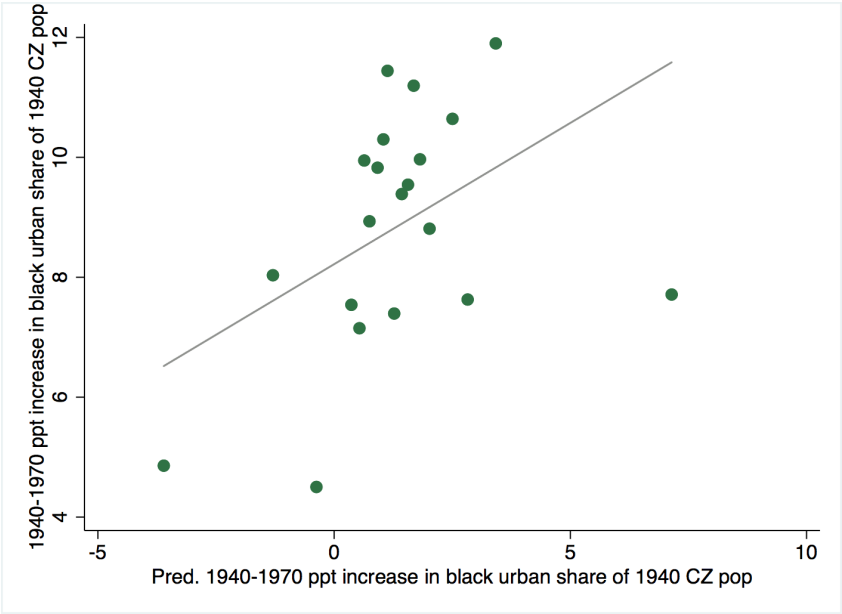
- Share of the labor force in agriculture
- WWII spending per capita
- Percent acreage in cotton
- Percent acreage in tobacco
- Indicator for mining state
- Indicator for mining state × Share of the labor force in mining
- Share of the labor force in mining

Variables selected in 1960:

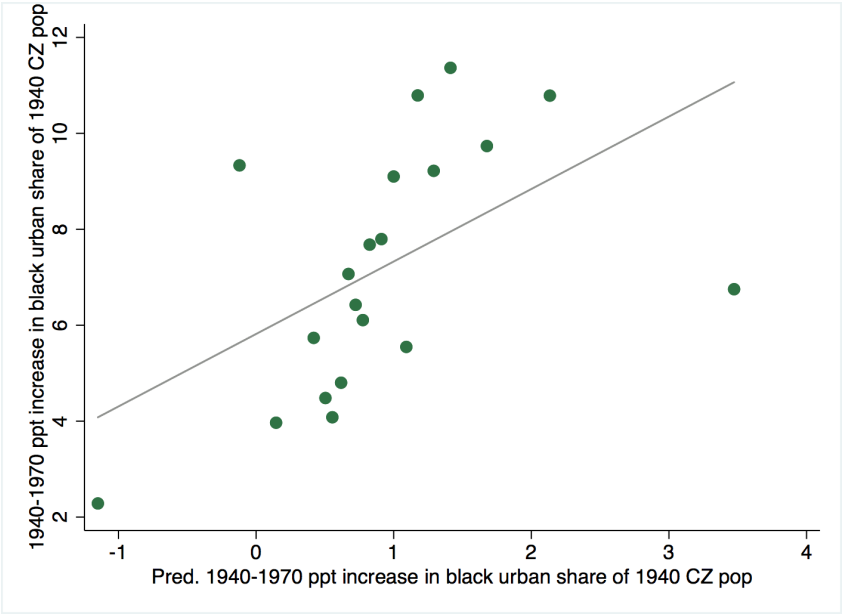
- Percent tenant farms
- Share of the labor force in agriculture
- Indicator for tobacco growing state
- Share of the labor force in agriculture × Tobacco growing state
- Percent acreage in cotton
- Indicator for mining state
- Indicator for mining state × Share of the labor force in mining
- Share of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

FIGURE C4: FIRST STAGE IN LEVELS OF 1940-1970 BLACK POPULATION INCREASE



(a) Full sample



(b) Windsorized

Notes: Panel (a) shows the relationship between actual urban black population increase from 1940-1970 as a share of the initial 1940 population and predicted increase of the same. Panel (b) shows this relationship after windsorizing the sample based on the 5th and 95th percentile of increases. *Source*: Complete count 1940 census; CCDB; Boustan (2016).

Appendix D Appendix to upward mobility results

Table D1: Upward mobility results summary

	Average	Expos. Effects	Black, p25	Black, p75
GM Shock	-0.0299***	-0.00242***	-0.0248**	-0.0357**
	(0.00952)	(0.000628)	(0.0113)	(0.0167)
Edu. Upward Mobility 1940	-0.0735*	0.0000909	-0.0144	-0.0772
	(0.0416)	(0.00268)	(0.0492)	(0.0728)
LF in manuf. 1940	-0.144***	-0.00321*	-0.0816**	-0.0000935
	(0.0270)	(0.00192)	(0.0319)	(0.0472)
Black Southern Mig 1935-1940	-5.099***	-0.0720	-0.567	-2.208
	(1.387)	(0.0654)	(1.649)	(2.443)
Midwest	-0.261	0.100***	-1.391*	-0.656
	(0.610)	(0.0357)	(0.722)	(1.069)
South	-2.054	0.170**	-0.306	1.418
	(1.291)	(0.0752)	(1.527)	(2.262)
West	-1.756*	-0.103*	-1.379	-1.004
	(0.958)	(0.0546)	(1.143)	(1.693)
R-squared	0.493	0.249	0.185	0.116

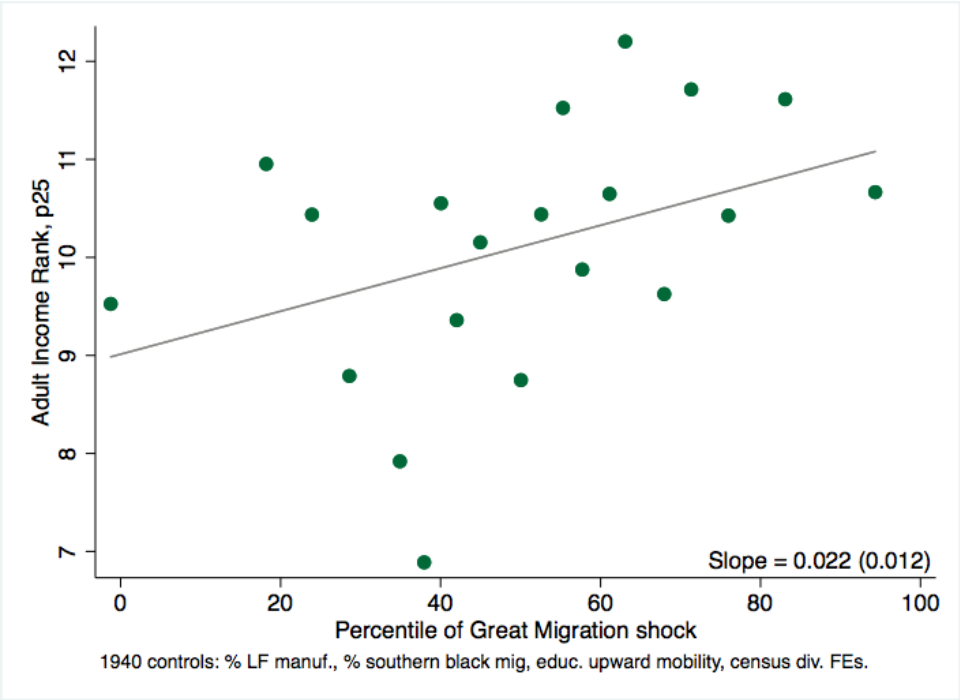
Dependent variable is mean individual income rank, where income is measured from IRS tax returns for cohorts born between 1978 and 1983. Pooled income refers to household income, as opposed to individual income. Independent variable is predicted change in black population share between 1940 and 1970. 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 from Chetty and Hendren (2018).

TABLE D2: GREAT MIGRATION IMPACT ON AVERAGE UPWARD MOBILITY FOR BLACK FAMILIES IN THE 2000S (INDIVIDUAL EARNINGS)

	First Stage				
	Pooled HH Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM Shock	0.304***	0.304***	0.304***	0.304***	0.304***
	(0.0750)	(0.0750)	(0.0750)	(0.0750)	(0.0750)
F-Stat	23.32	23.32	23.32	23.32	23.32
	Ordinary Least Squares				
	Pooled HH Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM	-0.0480***	0.0250***	-0.0538***	0.0307**	-0.0710***
	(0.00711)	(0.00795)	(0.00875)	(0.0155)	(0.0128)
R-squared	0.354	0.200	0.256	0.0872	0.213
	Reduced Form				
	Pooled HH Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM Shock	-0.0301***	0.0132	-0.0366***	0.0250	-0.0464***
	(0.00772)	(0.00804)	(0.00927)	(0.0153)	(0.0134)
R-squared	0.213	0.155	0.138	0.0782	0.104
	Two-stage least squares				
	Pooled HH Lower Inc	Women Lower Inc	Men Lower Inc	Women Higher Inc	Men Higher Inc
GM	-0.0516**	0.0316	-0.0777**	0.112*	-0.112**
	(0.0253)	(0.0292)	(0.0318)	(0.0587)	(0.0480)
R-squared	0.416	0.228	0.299	0.0649	0.209
N	129	129	129	129	129
Outcome Mean	33.19	40.33	38.88	49.32	51.46
Outcome SD	2.747	2.763	3.151	5.036	4.481

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 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 in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty et al. (2018).

FIGURE D1: GM IMPACT ON RACIAL GAP BETWEEN MEN, LOW INCOME PARENTS



Notes: This binned scatterplot depicts the relationship between the racial gap in upward mobility among men in 2015 and the percentile of predicted black population change between 1940 and 1970 (in units of 30 percentile points). The unit of observation is a commuting zone. 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. 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 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).

D.1 Childhood exposure effects

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

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

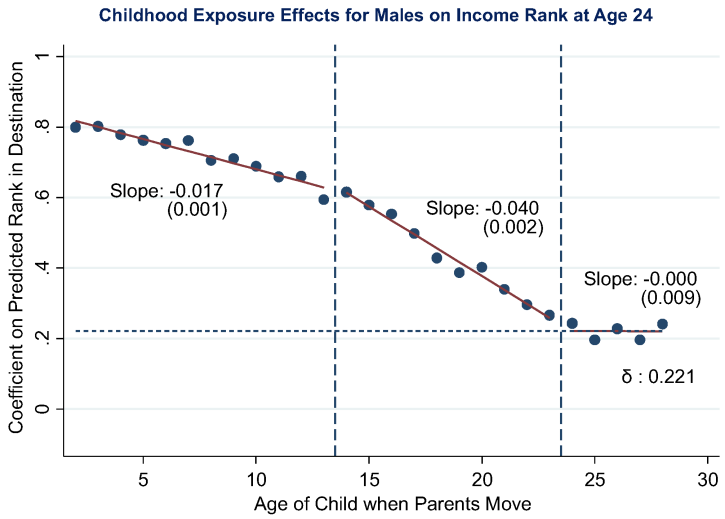
α_c is an unbiased estimate of effect of additional year of childhood exposure to location c on adult outcome y_i.

D.1.1 Scaling childhood exposure effects

Assuming muted effect for early years according to Figure D2, the effect of full childhood exposure for 23 years should be adjusted in the following manner:

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

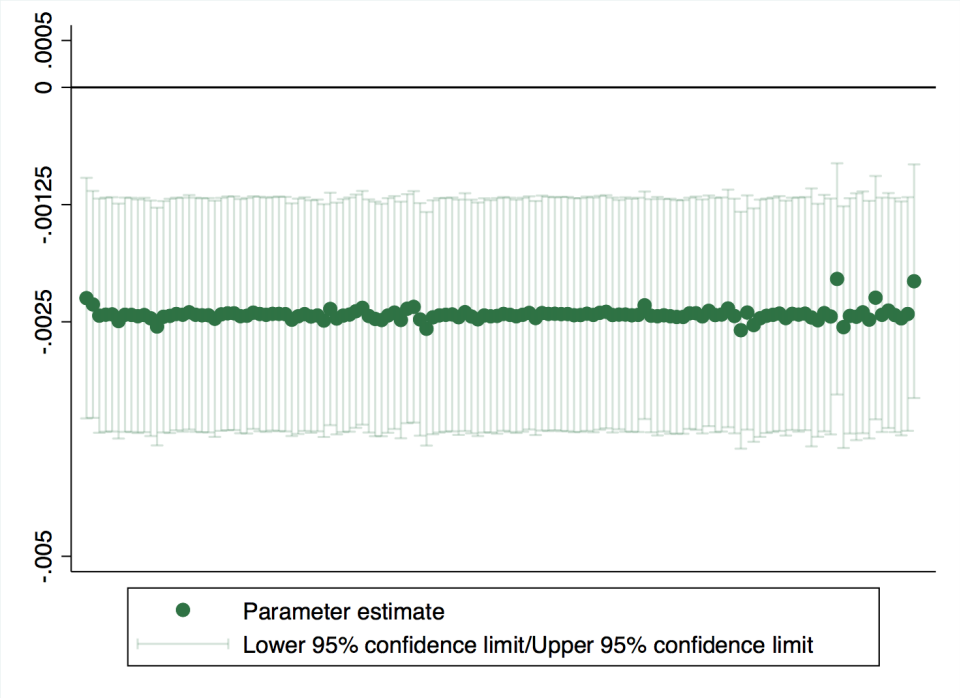
Figure D2: Heterogeneity in childhood exposure effects by age of child (Chetty et al., 2018)



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.

D.2 Additional robustness checks

FIGURE D3: 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 US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

TABLE D3: RESULTS ON CHILDHOOD EXPOSURE EFFECTS USING SOUTHERN-STATE-OF-BIRTH TO DEFINE GREAT MIGRATION SHOCKS

	First Stage on GM					
	1880	1910	1920	1930	1940	1935-40
GM Shock	0.131	0.224**	0.609***	0.455***	0.327***	0.292***
	(0.0907)	(0.0956)	(0.113)	(0.106)	(0.0984)	(0.0652)
F-Stat	4.203	6.638	16.51	24.45	26.98	27.46
	Ordinary Least Squares					
GM	-0.00262***	-0.00277***	-0.00292***	-0.00310***	-0.00303***	-0.00265***
	(0.000610)	(0.000633)	(0.000691)	(0.000797)	(0.000844)	(0.000823)
R-squared	0.226	0.229	0.229	0.229	0.227	0.223
	Reduced Form					
GM Shock	-0.00108*	-0.00161**	-0.00301***	-0.00316***	-0.00255**	-0.00242***
	(0.000654)	(0.000721)	(0.000992)	(0.00103)	(0.000980)	(0.000628)
R-squared	0.128	0.143	0.179	0.196	0.190	0.249
	Two-stage least squares					
GM	-0.00830	-0.00718**	-0.00495***	-0.00696***	-0.00780**	-0.00830***
	(0.00597)	(0.00350)	(0.00158)	(0.00233)	(0.00319)	(0.00250)
R-squared	-0.324	-0.0778	0.175	0.0815	0.0233	-0.0774
N	130	130	130	130	130	130
Precision Wt	Y	Y	Y	Y	Y	Y
Mean Expos FX	-0.0160	-0.0160	-0.0160	-0.0160	-0.0160	-0.0160
SD Expos FX	0.172	0.172	0.172	0.172	0.172	0.172
SD GM	24.82	24.82	24.82	24.82	24.82	24.82

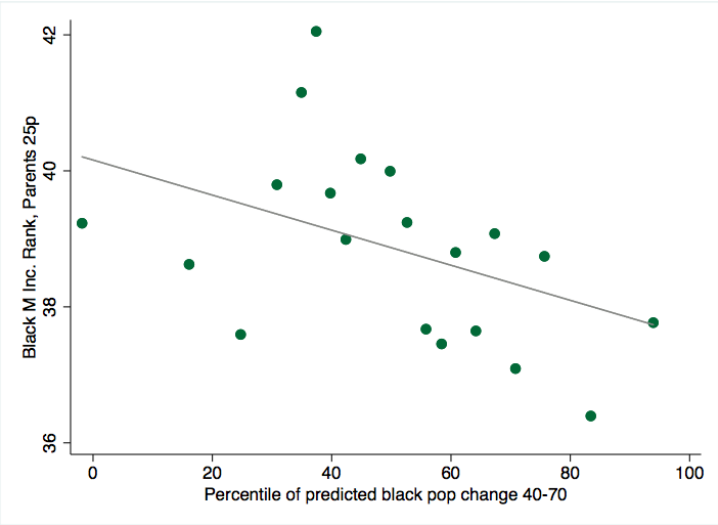
Notes: This table reports results on commuting zone childhood exposure effects using 1880-1940 southern-state-of-birth to construct the instrument for black population changes. The unit of observation is a commuting zone. Dependent variable is the estimated causal impact of one additional year of childhood 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. 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 migrant shares and post-1940 outflows of migrants as predicted by southern economic factors alone. Columns 1-5 use southern state of birth of the black population in northern cities in 1880, 1910, 1920, 1930 and 1940, respectively, to define black southern migrant shares. 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 the year indicated. Column 6 uses 1935-1940 black migration patterns to define the shares and assigns predicted southern county outflows to northern cities based on these shares. Baseline 1940 controls include share of urban population made up of 1880, 1910, 1920, 1930, 1940, or 1935-1940 black southern migrants respectively; educational upward mobility; share of labor force in manufacturing; and census division fixed effects. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

TABLE D4: RESULTS ON BLACK MEN'S UPWARD MOBILITY USING SOUTHERN-STATE-OF-BIRTH TO DEFINE GREAT MIGRATION SHOCKS

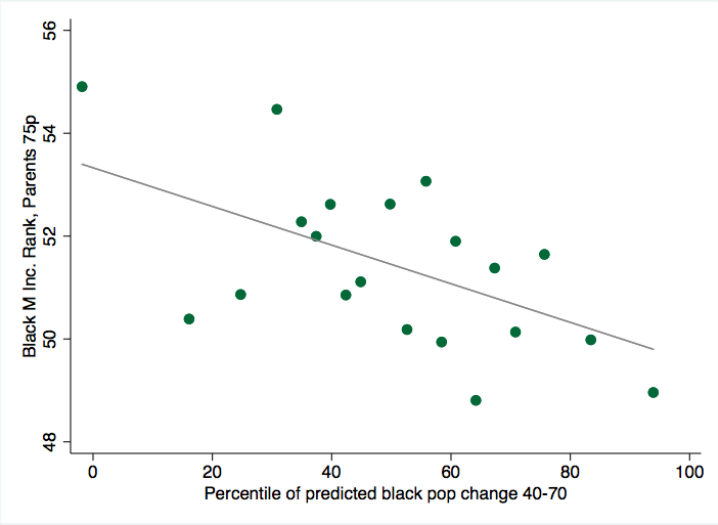
	First Stage on GM					
	1880	1910	1920	1930	1940	1935-40
GM Shock	0.133	0.140	0.353***	0.451***	0.403***	0.304***
	(0.0822)	(0.0970)	(0.103)	(0.0966)	(0.0935)	(0.0750)
F-Stat	5.631	7.580	13.70	22.65	24.84	23.32
	Ordinary Least Squares					
GM	-0.0474***	-0.0486***	-0.0507***	-0.0578***	-0.0643***	-0.0625***
	(0.00964)	(0.0101)	(0.0109)	(0.0118)	(0.0121)	(0.0120)
R-squared	0.289	0.287	0.288	0.298	0.312	0.309
	Reduced Form					
GM Shock	-0.0143	-0.0184	-0.0315**	-0.0585***	-0.0565***	-0.0248**
	(0.00942)	(0.0116)	(0.0134)	(0.0136)	(0.0137)	(0.0113)
R-squared	0.163	0.167	0.196	0.270	0.258	0.185
	Two-stage least squares					
GM	-0.112	-0.114	-0.0862**	-0.127***	-0.137***	-0.0777**
	(0.0767)	(0.0743)	(0.0349)	(0.0318)	(0.0351)	(0.0318)
R-squared	0.0296	0.0424	0.225	0.0993	0.110	0.299
N	129	129	129	129	129	129
Mean Dep Var	38.88	38.88	38.88	38.88	38.88	38.88
SD Dep Var	3.151	3.151	3.151	3.151	3.151	3.151
SD GM	28.80	28.80	28.80	28.80	28.80	28.80

Notes: This table reports results on black men's upward mobility using 1880-1940 southern-state-of-birth to construct the instrument for black population changes. 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 migrant shares and post-1940 outflows of migrants as predicted by southern economic factors alone. Columns 1-5 use southern state of birth of the black population in northern cities in 1880, 1910, 1920, 1930 and 1940, respectively, to define black southern migrant shares. 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 the year indicated. Column 6 uses 1935-1940 black migration patterns to define the shares and assigns predicted southern county outflows to northern cities based on these shares. Baseline 1940 controls include share of urban population made up of 1880, 1910, 1920, 1930, 1940, or 1935-1940 black southern migrants respectively; educational upward mobility; share of labor force in manufacturing; and census division fixed effects. Standard errors in parentheses: * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$. *Data sources*: CCDB, IPUMS complete count US census for 1940; Boustan (2016); Chetty and Hendren (2018b).

Figure D4: Great Migration impact on black men's outcomes, residualizing southern county net-migration on state FEs



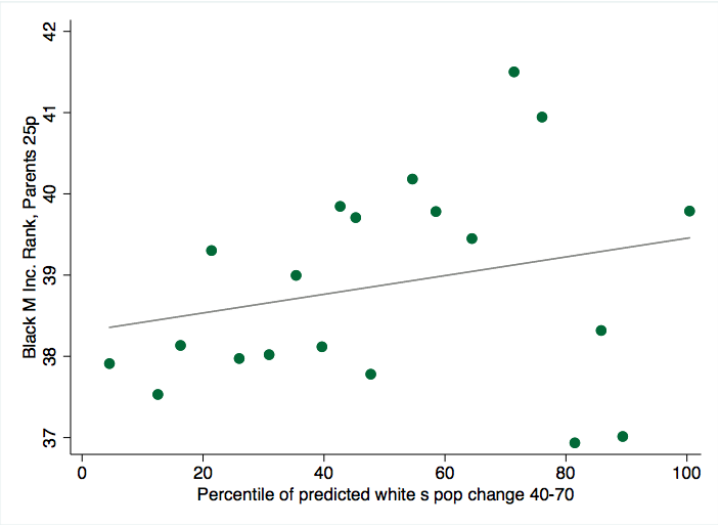
(a) 25th percentile



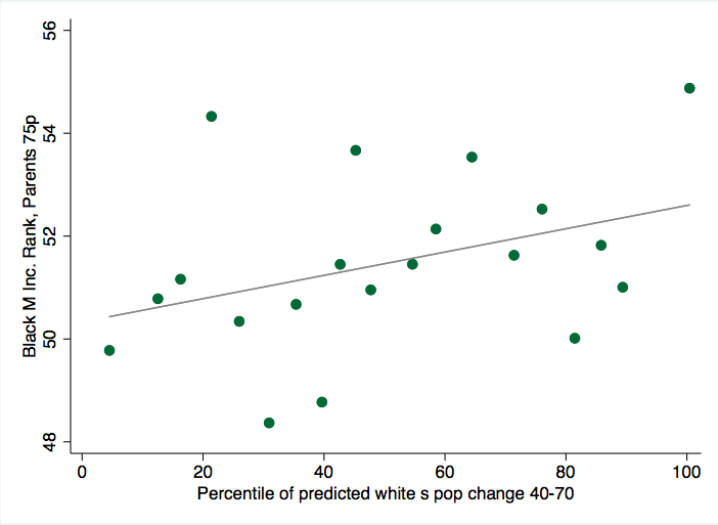
(b) 75th percentile

Notes: Panel (a) depicts a binned scatterplot of the relationship between the percentile of predicted black in-migration and black men's upward mobility (25th percentile of parent income distribution). In this definition of the Great Migration shock, southern county net-migration rates have first been residualized on southern state fixed effects. 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 D5: White southern migration impact on black men's outcomes



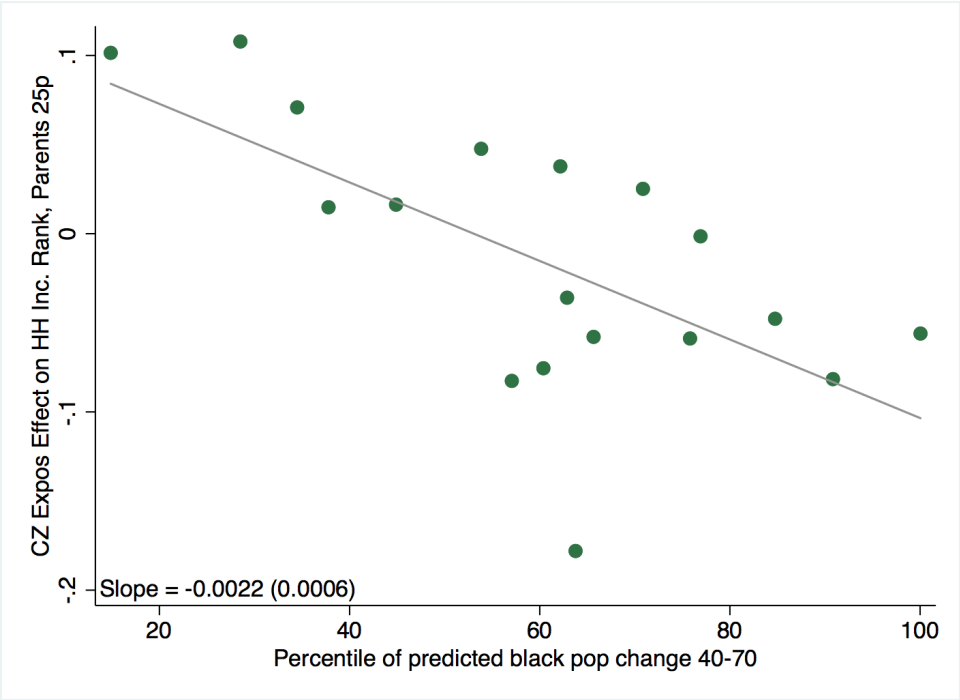
(a) 25th percentile



(b) 75th percentile

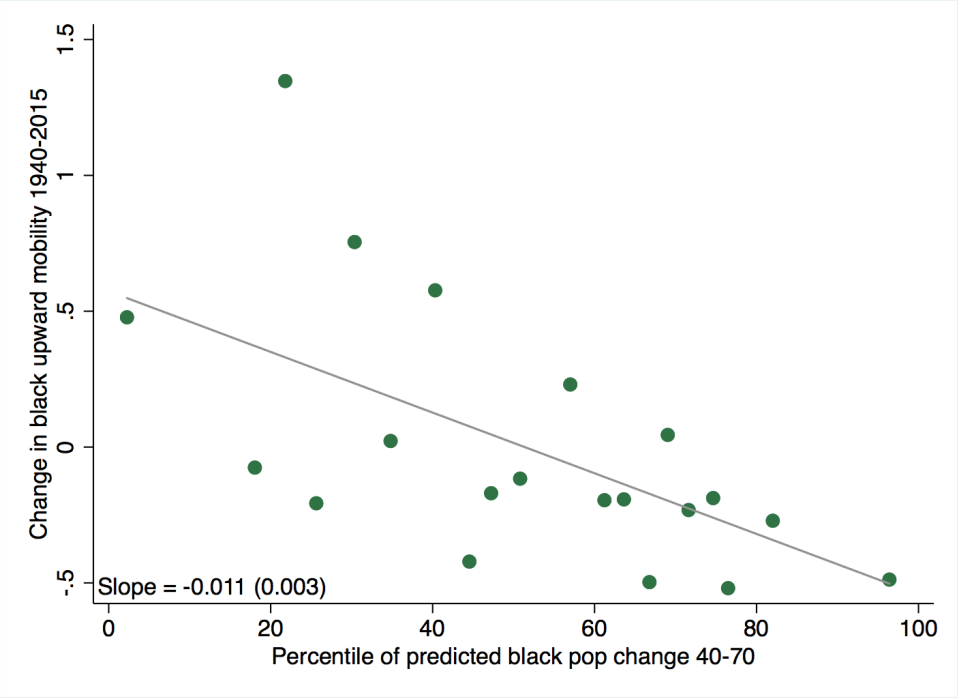
Notes: Panel (a) depicts a binned scatterplot of the relationship between the percentile of predicted white southern in-migration and black men's upward mobility (25th percentile of parent income distribution). Panel (b) depicts the same for black men from the 75th percentile of parent income distribution. 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 D6: Great Migration impact on CZ exposure effects, controlling for quartiles of black population share in 1940



Notes: This figure depicts a binned scatterplot of the relationship between the percentile of predicted black population increase and the 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).

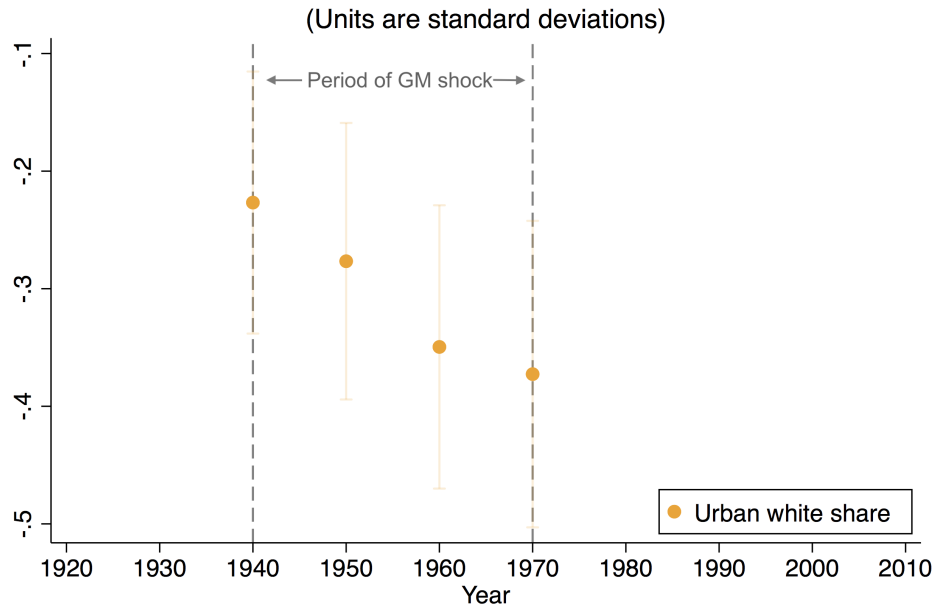
Figure D7: Great Migration impact on change in black men's upward mobility, 1940-2015



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 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).

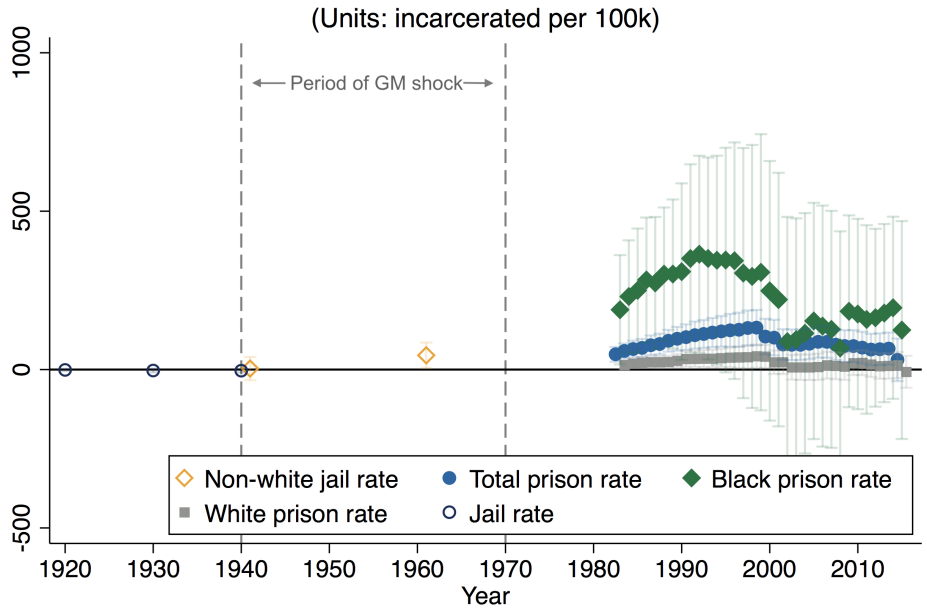
Appendix E Additional results on local mechanisms

Figure E1: Great Migration impact on urban white share



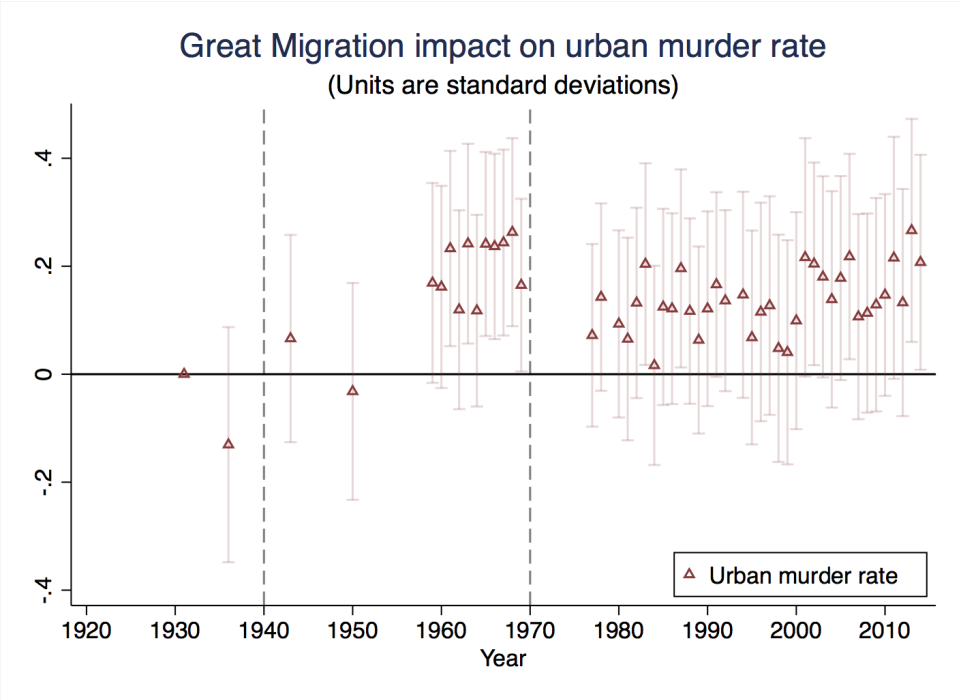
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is the urban white population share. The unit of observation is a commuting 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*: County Data Books 1947-1977.

Figure E2: Great Migration impact on incarceration rates, levels



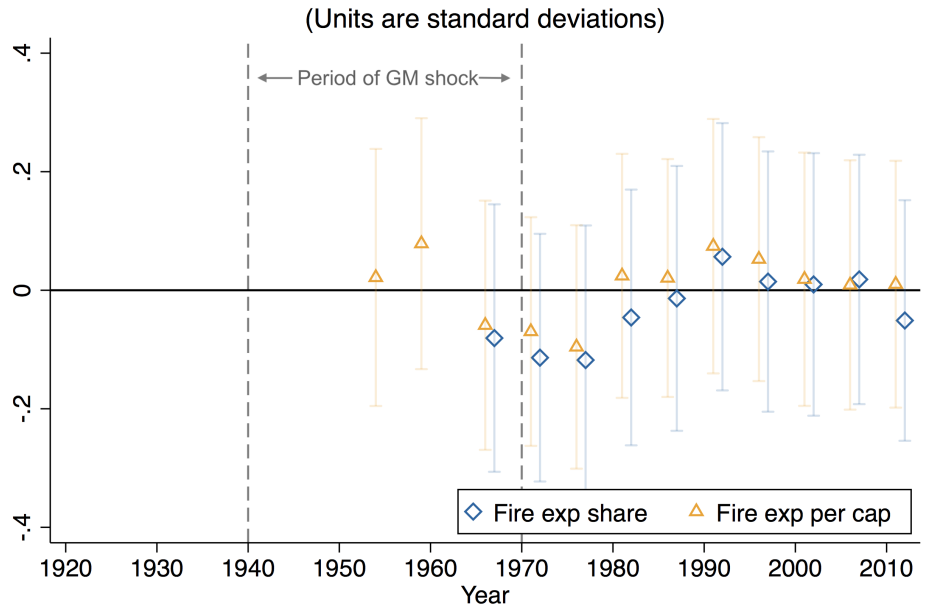
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is county jail population per 100,000 (1940 and 1960) or federal and state 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 E3: Great Migration impact on murder rates with pre-period control



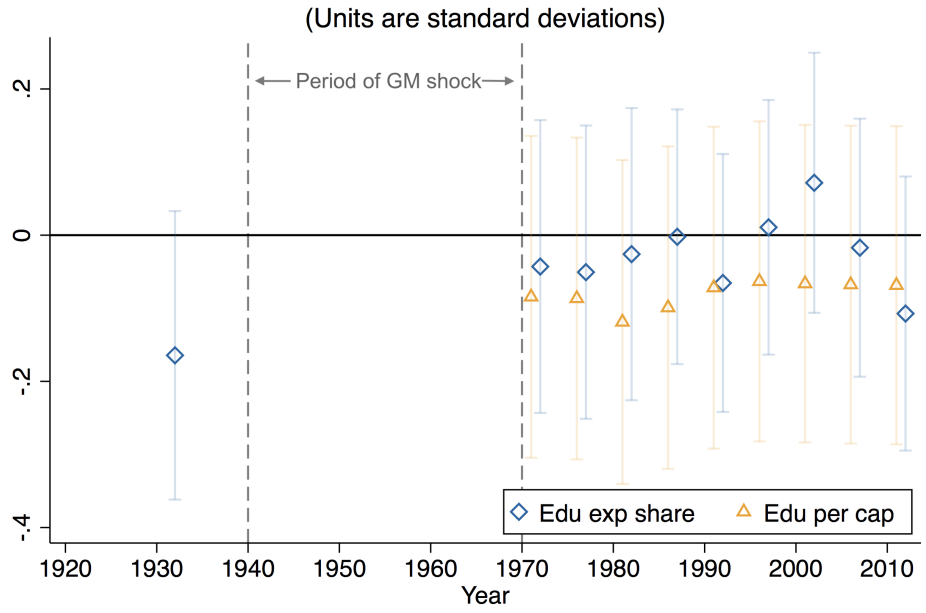
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is urban murder rates per 100,000 in commuting zones. All regressions include controls for the 1931 murder rate. 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 source*: Uniform Crime Reports.

Figure E4: Great Migration impact on fire-fighting investments



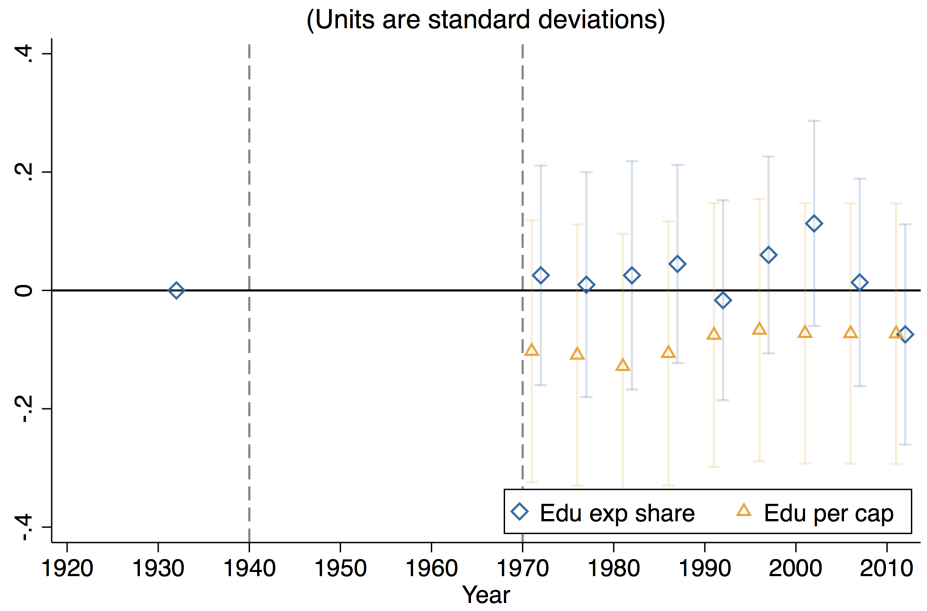
Notes: This figure plots the coefficient on percentile of predicted black population change (units are 30 percentile points, approximately 1 standard deviation) in separate regressions for each year where the dependent variable is either the share of local government expenditures on fire-fighting or 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 E5: 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 E6: 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).

Table E1: Great Migration impact on housing supply restrictions

	Ordinary Least Squares											
	LPPI	SPII	SCII	LZAI	LPAI	LAI	DRI	OSI	EI	SRI	ADI	WRLURI
GM	0.00550***	-0.000658	-0.00524**	-0.00124	0.00367**	-0.000553	0.000782	0.00189*	0.000150	0.00175	0.0296***	0.00656***
	(0.00141)	(0.00270)	(0.00221)	(0.00163)	(0.00155)	(0.000357)	(0.000704)	(0.000984)	(0.000641)	(0.00110)	(0.00585)	(0.00182)
R-squared	0.225	0.428	0.176	0.121	0.142	0.211	0.161	0.150	0.321	0.102	0.485	0.514
	Reduced Form											
GM Shock	0.00406**	-0.00203	-0.00807***	-0.00152	0.00246	-0.000804*	-0.000932	0.000996	0.000640	0.00361***	0.00598	0.00208
	(0.00171)	(0.00313)	(0.00254)	(0.00191)	(0.00187)	(0.000418)	(0.000841)	(0.00118)	(0.000759)	(0.00129)	(0.00718)	(0.00221)
R-squared	0.211	0.467	0.241	0.160	0.135	0.251	0.170	0.148	0.340	0.145	0.462	0.499
	Two-stage least squares											
GM	0.0114**	-0.00570	-0.0227***	-0.00427	0.00690	-0.00226**	-0.00262	0.00280	0.00180	0.0102***	0.0168	0.00584
	(0.00466)	(0.00855)	(0.00793)	(0.00536)	(0.00502)	(0.00115)	(0.00240)	(0.00318)	(0.00208)	(0.00388)	(0.0189)	(0.00588)
R-squared	0.206	0.460	-0.00491	0.101	0.148	0.234	0.0849	0.162	0.326	-0.0505	0.491	0.521
N	120	120	120	120	120	120	120	120	120	120	120	120
SD GM	28.11	28.11	28.11	28.11	28.11	28.11	28.11	28.11	28.11	28.11	28.11	28.11

Notes: LPPI = Local Political Pressure Index; SPII = State Political Involvement Index; SCII = State Court Involvement Index; LZAI = Local Zoning Approval Index; LPAI = Local Project Approval Index; LAI = Local Assembly Index; DRI = Density Restrictions Index; OSI = Open Space Index; EI = Exactions Index; SRI = Supply Restrictions Index; ADI = Approval Delay Index; WRLURI = Wharton Residential Land Use Regulation Index. *Source*: Wharton Land Regulation Data.

Table E2: Impact of Great Migration on neighborhood quality in 2000

	Standardized variables measured in 2000			
	Murder Rate	Racial Segregation (Theil)	Percent Single Parent	Segregation of Poverty
GM Shock	0.269**	0.598***	0.418***	0.626***
	(0.126)	(0.157)	(0.137)	(0.160)
Baseline mean	3.963	0.192	0.205	0.065
Std Dev	3.219	0.101	0.027	0.026
Observations	125	129	129	129
State FEs	Y	Y	Y	Y
CZ 1940 Pop	Y	Y	Y	Y

This table reports the coefficient on percentile of predicted black population change in regressions where the dependent variable is the urban murder rate per 100,000; the Theil index of racial segregation across census tracts; the Theil index of segregation of below median income families across census tracts; and the percentage of single-parent households.

Appendix F Public Finance and Neighborhoods Database, 1920-2015

F.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. In 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 25,000 or more from the Uniform Crime Reporting publications ("UCR") of the FBI in 1931, 1943, and 1950.[53] 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.[54]

Local government expenditures

Data on local government expenditures come from surveys of state and sub-state 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 or types of government are responsible for

[53]Some large cities did not report to the FBI UCR series in these years. A notable case is New York City. For these cities, I supplement using data generously shared by Price Fishback.

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

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 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 F1 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 F1: 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 3 shows total direct government expenditures by government type. Starting with Column 4, 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).