Migration, Skills-Biased Technical Change, and Human Capital Accumulation: Evidence From the Great Migration

Migration, Skills-Biased Technical Change, and Human Capital Accumulation: Evidence from the Great Migration

by

Fay Asimakopoulos

Presented to the Department of Applied Mathematics
in partial fulfillment of the requirement
for a Bachelor of Arts degree with Honors

Harvard College
Cambridge, Massachusetts
April 3, 2020

Abstract

In this paper, I show that the mass migration of relatively skilled workers from a developing region can encourage skill formation in their places of origin. I exploit exogenous variation in South-to-North migration during one of the largest domestic migration events in American history, the Great Migration of four million African Americans (1940-1970). I instrument immigrants' location decision relying on pre-existing settlement patterns, which I establish by matching two decades of census records, and exogenous variation in domestic migration induced by World War II. The Great Migration aided in the decline of a backwards agricultural tenancy system, improved Southern black workers' returns to education, increased the relative proportion of those that held high-skill occupations, and decreased occupational segregation as explained by literacy. Throughout, I provide evidence that these observations are consistent with a canonical model of skills-biased technical change.

Acknowledgements

There are simply too many people to thank. First, I thank my advisor Professor Alberto Alesina for his support, advice, and guidance. Thank you for always encouraging my research plans. I am equally thankful to my Economics 985 Seminar Advisor, Dr. Jeffrey Miron, for his actionable feedback, which helped me turn this paper from an idea into a reality. I must also thank Professor Marco Tabellini, whose work not only inspired my interest in the Great Migration, but whose kind and thoughtful responses to my questions played a large role in getting this project to its current form.

My interest in Economics was fuelled by many mentors. I am indebted to Professor Melissa Dell for teaching me an incredible amount and helping me solidify my passion for the discipline through a Research Assistant position last summer. I am also thankful to Professor Karen Dynan, whose Macroeconomics course first planted the desire to be an Economist in me, and whose advice throughout the years has been nothing short of invaluable.

Throughout the process of researching and writing this thesis, I benefited from the company and encouragement of too many friends to fully list. Derek Lee and Hadley DeBello – thank you for your feedback, your delicious food, and for reinstating my faith in the importance of this project. Julia Huesa – many late nights spent in Lamont Library scraping data would have been lonelier without your good spirits. Emilė Radytė – thank you for inspiring me to be a more efficient editor and more interdisciplinary thinker. Camille Falezán – I am grateful to you for making the Harvard Undergraduate Diversity in Economics group possible. Your indefatigable desire to make the discipline more accessible inspires me.

Finally, and on a more personal note, I want to thank my family in Greece and chiefly my parents, Sonia and Steve, who remain my biggest inspiration. Your sacrifices and selflessness push me to be a better person and a clearer thinker. This work is, and always will be, dedicated to you.

Contents

1 Introduction

Between 1940 and 1970, four million African Americans left the rural South of the United States and settled in urban areas in the North and West of the country in search of better economic opportunities (McMillen, 1997). Although emigration is frequently depicted as a problem for developing economies, this landmark historical event coincided with a period of rapid economic and social transformation in the South, which transitioned from a traditionally agrarian economy premised on the harvest of cotton and corn, to an industrial production system not unlike that of the North (Wright, 1986).

The effects of emigration on sending communities are both under-studied and theoretically ambiguous. In this paper, I study whether the Great Migration (henceforth Migration) improved occupational outcomes for the nearly 14 million Southern blacks that remained in the South. Comparing counties in the South that experienced larger or smaller decreases in their black population due to the Migration, I estimate the impact of these changes on average economic outcomes in the period spanning 1940 to 1970. I find that, consistent with the historical literature, the Migration contributed to large-scale farm consolidation, agricultural modernization, and improvements in manufacturing productivity. I also show that it had positive effects on occupational outcomes for African Americans, increasing economic returns to education and the proportion of black workers employed in high-skill jobs, while decreasing some forms of occupational segregation between white and black workers.

My empirical strategy depends on the fact that black migrants during the Great Migration settled where others from their communities had moved, giving rise to specific linkages between Southern counties of origin and Northern destinations (Boustan, 2010; Derenoncourt, 2018; Boustan and Tabellini, 2019). To address omitted factors that may simultaneously prompt increases in black out-migration and changes in Southern economic conditions, I use a modified version of the "shift-share" approach first proposed by Altonji and Card (1991). A "shift-share" instrument combines two sources of information: the non-Southern migration destination choices of black Southern migrants before 1940 and variation in net migration into the same counties from 1940 to 1970. To ensure that my instrument only exploits variation in the composition of black migrants across Southern states over time,

I follow Tabellini (2018) and condition on the share of each county's population that was black in 1940. Following the literature on "shift-share" instrument design (Adão et al. 2019; Borusyak et al. 2018; Pinkham et al., 2018) I also identify and address three threats to identification by modifying the instrument.

First, the characteristics of Southern counties that pushed African Americans to migrate before 1940 might have persisted, confounding the Great Migration's effects on each county's economic trajectory and on later migration patterns. I tackle this in two ways. First, I document that predicted black out-migration is not correlated with two other major factors co-determining Southern economic growth: World War II and New Deal funding. Then, I show that the 1920-1940 change in the black population is not correlated with black out-migration predicted by my instrument after 1940. As an additional check, I provide evidence that the instrument does not predict the 1940-1970 change in the white population.

The second concern is that the instrument could be driven by a small number of specific linkages between Southern origin counties and Northern destinations (Pinkham et al., 2018). I show that the instrument is not sensitive to variation coming from the initial shares of blacks migrating to different Northern states by repeating the first-stage analysis with an additional control for the share of each county's out-migrants that went to a different Northern state between 1920 and 1940. Reassuringly, the strength of the instrument remains unchanged.

Finally, the instrument would not be valid if migration into each Northern county was correlated with Southern push factors that were simultaneously systematically related to the county of origin of pre-1940 black migrants. To address this concern, I predict migration into non-Southern states from 1940 to 1970 using county-level data on World War II spending. An attractive feature of the World War II manufacturing boom is that it dramatically and unexpectedly altered both the number and the composition of Southern migrants, reducing the auto-correlation in migration flows to Northern destinations.

Using this strategy, I show that the Great Migration led to the modernization of the Southern systems of agriculture and manufacturing as well as to the improvement of African Americans' occupational outcomes as measured by an increase in the return to education, an increase in the share of African Americans holding high-skill jobs, and a decrease in occupational segregation in high-segregation counties.

Counties experiencing greater black out-migration during the Migration saw a significantly larger decline in the proportion of their farms that operated under the tenant system, robust to measuring the proportion of tenant farms both as a proportion of overall farms and as a proportion of overall farm acreage. They experienced greater farm consolidation, measured by a larger decline in the number of farms and a larger increase in per-farm acreage. Finally, I present some evidence that high out-migration counties saw a greater degree of mechanization, as measured by greater tractor adoption.

While there is no indication that manufacturing wages rose, I also present evidence that counties that were more affected by the Great Migration experienced a larger increase in per-capita manufacturing value added, implying an increase in the productivity of manufacturing. The same counties also saw a decrease in the share of the population working for manufacturing establishments, relative to counties with fewer emigrants.

Furthermore, counties more affected by the Great Migration reported higher returns to education for black workers, as well as a higher proportion of black workers working in high-skill occupations in 1970. I measure the average relative return to education of Southern blacks by running a standard Mincer wage equation for each county, using the 2 % sample of the 1970 census. My findings indicate that increased black out-migration is simultaneously associated with a flat wage penalty due to race as well as with increased returns to education for African Americans in Southern counties in 1970. I find no evidence of a similar premium when I estimate the Mincer equation across the entire population. I compute the share of whites and blacks working in low- and high-skill occupations, and find that black out-migration is associated with an increase in the share of both black and white workers working in high-skill jobs, but no statistically significant change in the number holding low-skill jobs. These results are consistent with the economic intuition that a period of skill-biased technical change spurred on by mechanization and manufacturing growth incentivized occupational upgrading for African Americans, without finding evidence of improved average work conditions.

Finally, I show that greater black out-migration is associated with lower occupational segregation in 1970, once I condition on the literacy levels of workers. Using census data and two distinct segregation indices, I study the effect of black out-migration on occupational

segregation between black and white workers. I construct three separate forms of segregation: observed occupational segregation, a simulated measure of occupational segregation that randomizes race within literacy levels, and a residual measure of occupation that filters out segregation due to literacy. I also restrict my sample to county groups with an above or below median value of each form of segregation, to test for heterogeneous effects on counties. I find that, regardless of whether a county is high- or low-segregation, the Great Migration is associated with lower "literacy-conditional" levels of segregation, suggesting a segregation-lowering pathway through the increased literacy of Southern blacks. Once I filter out variation due to literacy, the relationship between out-migration and segregation disappears.

This paper relates to three strains of literature. Most immediately, it contributes to a nascent field using historical census data to study the long-run impact of the Great Migration. Derenoncourt (2018) shows that racial composition changes due to South-to-North migration during the peak of the Great Migration (1940-1970) reduced upward mobility in Northern cities, with the largest effects on black men due to increased incarceration rates. Tabellini (2019) shows that an increase in racial heterogeneity following the Great Migration decreased the provision of public goods by receiving cities, primarily due to a decrease in property values resulting from white backlash. In this paper, I provide evidence of positive effects of the Great Migration on the occupational outcomes of African Americans in origin counties and provide suggestive evidence of new intermediate mechanisms that could affect economic success: higher incentives for skills acquisition and lower levels of occupational segregation.

Second, my work is one among a number of recent studies that investigate the link between migration on technical change. Andersson et al. (2019) argue that the exodus of Swedes to the United States during the American Age of Mass Migration (1850-1920) induced technological change in sending locations. The authors find that out-migration led to increased adoption of new technologies in both the agricultural and industrial sectors as well as to higher unskilled wages in agriculture, a shift towards employment in the nascent industrial sector, a larger presence of incorporated firms, as well as higher tax revenues. Similarly, Clemens et al. (2018) use the exclusion of Mexican agricultural workers in the United States after the end of the seasonal migration Bracero program to study the effects

of a labor market exclusion policy on wages and technology adoption. They find that the dissolution of the program led employers who were dependent on migrant labor to adopt labor-saving technology, rather than hire more native workers. As far as I know, I am the first to recast these questions in the context of out-migration during the Great Migration.

Finally, my findings relate to theories of labor scarcity-induced directed-technical change (Acemoglu, 2010; Acemoglu and Autor, 2012; Goldin and Katz, 2010). Labor saving technologies – including many of the agricultural and manufacturing innovations introduced during the 20^{th} Century – reduce the marginal product of labor, and promote the adoption of technological advances after periods of labor scarcity. If high and low skill labor are gross complements, then skill augmenting technology increases the skill premium. These findings are consistent with my findings about the mechanization of Southern agriculture, and in line with my discovery that the black skill premium increased parallel with an increase in the adoption and use of skill-intensive technology.

The paper closest to mine is Hornbeck and Naidu (2014), which finds that flooding events in the Mississippi delta that led to a plausibly exogenous migration of African Americans in the 1920s are associated with increased levels of mechanization, as landowners in flooded counties modernized agricultural production and increased its capital intensity relative to landowners in nearby similar non-flooded counties. Still, I posit that my paper improves upon the authors' empirical strategy by eliminating the endogeneity threat that might come from spatial auto-correlation due to the flood being both a push factor and co-determining factor in agricultural growth, and crucially takes a longer-term stance on Southern economic development by using data that spans the entirety of the Migration.

The paper proceeds as follows: Section 2 describes the historical background to the Great Migration, and gives contextual information to situate this large migration episode among other changes occurring in the Southern and Northern economy. Section 3 presents the data used in the study, describes the matching algorithm used to construct migration shares, and discusses any limitations. Section 4 lays out the empirical strategy, constructs the two instruments for black out-migration, and argues that the instrument meets the relevance and exogeneity conditions. Section 5 summarizes the main empirical results for the effect of black out-migration on the tenancy system, manufacturing growth, and black Southerners'

occupational and human capital outcomes. Section 6 concludes, and outlines directions for future research.

2 Background

2.1 Background to the Southern Economy

At the beginning of the 20^{th} century, white Southern planters economically dominated areas with concentrated African American populations in a feudal caste system that touched every corner of daily life. Many African Americans in the South were employed in the farm sector, particularly in the production of cotton. Under systems of share-cropping and share-tenancy, black Southerners provided the labor that enabled Southern farms to continue to exist, but were rarely paid their marginal product. An absence of economic and social insurance meant that they could easily lose their farms or their status as cash or share tenants because of crop failures, low cotton prices, ill health, soil exhaustion, excessive interest rates, or an inability to compete with tenant labor, since under this transitory system of land tenure the landlord was required to provide nothing but the land. This equilibrium proved difficult to break out of as a constant surplus of tenants meant that at the end of the crop year landlords could easily recruit new renters, often on terms even more favorable to them. By some estimates, in 1920 two-thirds of all tenants moved from one farm to another (Conrad, 1965).

In the middle decades of the 20^{th} century, during a period that coincided with massive out-migration from the region, farms were consolidated and the system of sharecropping transitioned toward capital-intensive (rather than previously labor-intensive) production techniques. Multiple historical accounts have suggested that black out-migration, by inducing labor shortage, might have spurred this economic transformation (Mandle, 1992). Raper (1946) suggests that a decline in low-skill labor following the mass migration of black farmers lowered the relative cost of switching to more capital-intensive farming processes. Another channel might have been through political economy: Margo (1991) and Boustan and Tabellini (2019) argue that out-migration increased the bargaining power of remaining black residents due to the threat of further departures, thereby incentivizing the passing of

more inclusive policy.

2.2 Background to Black Emigration

Starting in the 1910s, African-Americans left the South in such large numbers that by the 1970s, the black population of the region had halved (McMillen, 1997). This Northern movement was motivated by an array of long-run push and pull factors, and activated by the stress of post-World War I events. Under Jim Crow laws in the South, black Americans – although no-longer enslaved – faced severe limitations on their political, social, and economic freedoms. Right as the Great Depression decreased demand for cheap labor in the South, World War II bolstered industrial growth in the North, prompting increasing numbers of black migrants to seek a better fortune through migration.

Despite the long-run alignment of pull and push factors, several conditions delayed the large-scale out-migration by Southern blacks until the beginning of the 20^{th} Century. One such reason was the poor education quality and mismatched agrarian skills of black workers. Even if unskilled non-farm labor did not require extensive schooling, literacy was important in learning about and taking advantage opportunities in different regions (Margo, 1991). A second important factor was competition for Northern jobs from low-skilled European immigrants, who poured into American cities during the decades leading up to the First World War and faced less employment discrimination than African Americans due to their relative proximity to whiteness (Calderon et al., 2018). For a long time, the possibility of higher Northern wages was thus offset by a lower probability of employment for Southern black immigrants (Collins, 1997). A sequence of economic shocks, combined with the 1917 and 1924 Immigration Acts that barred a majority of low-skilled Southern and Eastern European migrants from entering the country, the Boll-Weevil pest that devastated agriculture, and generous New Deal packages that weakened ties to agriculture were needed to kickstart the Migration.

Although black workers were treated with hostility and subjected to legal discrimination, the Southern economy was deeply dependent on them as croppers, cotton pickers, and factory workers. At the beginning of the Great Migration, white Southern elites might have not yet realized their dependence on black labor. Many appeared to be unconcerned by the prospect

of a mass black exodus, with some industrialists and cotton planters seeing the possibility of black out-migration as an opportunity to reduce surplus labor (Reich, 2014). As the migration picked up, Southern elites began to realize that a prolonged out-migration might bankrupt the South. As a result, some Southern employers increased their wages to match those on offer in the North, with a fraction going as far as to oppose aspects of Jim Crow laws.

On the other hand, some employers began to act violently in an attempt to coerce workers to remain in the South. At the same time that efforts were made to restrict bus and train access for blacks, agents were stationed in Northern cities to report on unionization and the rise of black nationalism, and newspapers were pressured to make their coverage of life in the Northern more negative (Reich, 2014). Still, there is some evidence that as the mechanization of agriculture in the late 1930s had resulted in another labor surplus, Southern planters put up less resistance in later waves of the Great Migration.

3 Data

In this section, I review this paper's major data sources as well as the construction of its main economic variables and controls.

Data on Black Out-Migration: In order to track South-to-North migrants' locations, I match Southern-born men across the 1920 and 1940 censuses. More precisely, I use the restricted[1] full-count version of the 1920 and 1940 census in conjunction with per-county net-migration counts computed by Winkler et al. (2013) and Gardner and Cohen (1992), and made available by the Inter-university Consortium for Political and Social Research (henceforth ICPSR). In order to construct my instrument (see Section 4.1), I identify men[2] who lived in the South in 1920 and who had moved to the North in 1940 using the iterative matching algorithm first proposed by Ferrie (1996) and fully outlined and automated by Abramitzky, Boustan, and Eriksson (2012, 2014). This procedure entails matching census

[1]The restricted version includes full names for all enumerated individuals.

[2]During this era, it was typical for women to adopt their husbands' names upon getting married. Therefore, this matching strategy would work less well for women.

records according to the NYSIIS-standardized version of their name, as well as their age, race and state of birth.[3] In order to avoid over-stating migration counts, since most false matches will be coded erroneously as migrants, I require that all successful matches are unique by place of birth and exact on a 5-year age band.

A concern with false matches over-stating migration counts is that this might produce a biased image of Southern migrants' pre-period migration patterns. This would be particularly problematic, since my instrument relies on these patterns as its source of variation in migration locations. To show that this effect does not significantly bias my estimates, I perform an additional robustness check for the above matching technique using the full-count 1940 census to calculate a "1935-1940 migration matrix" of individuals who lived in some Southern county in 1935, but had moved to a Northern county by 1940. The 1940 Census is the first census to ask individuals to report on their locations 5 years before. The two estimates are highly correlated, and the first stage remains robust to using the migration matrix to construct the shift-share instrument (see Table 1). However, the 1920-1940 matched-sample matrix remains my preferred specification, as its larger number of South-to-North linkages (over 20,000) taken over an extended period of time (20 vs. 5 years) gives a more detailed picture of the initial distribution of Southern immigrants in the North.[4] By contrast, the 1935-1940 matrix captures only a fraction of the pre-period migration trends, which might over- or under-state the pre-period settlement patterns due to year-specific trends.

I obtain age-specific net migration estimates by decade for US counties from 1950 to 1970 from Winkler et al. (2013) and from 1930 to 1950 from Gardner and Cohen (1992). Both files are made available by the ICPSR. These data include estimates of net migration for each decade from US counties by five year age group, sex, and race. The underlying mi-

[3]The NYSIIS standardization strategy analyzes words and creates group keys from letters that represent word pronunciation groups, therefore minimizing the possibility that individuals are not matched on names due to common spelling mistakes.

[4]This plays a crucial role in the shift-share instrument's construction. Both Adão et al. (2018) and Goldsmith-Pinkham et al. (2018) find that the shift-share's identifying assumptions are more likely to be met given a greater the number of shares. For a more thorough discussion of the shift-share literature, consult section 4.2.

gration numbers are estimated by comparing the population in each age-sex-race cohort at the beginning and end of a Census period and attributing the difference in population count to net migration, after adjusting for births and mortality. Any net inflow of immigration from abroad would be captured in this measure as an increase in the county's rate of net in-migration. This method has become standard practice to estimate internal migration in the United States, as originated by Kuznets et al. (1957). Following Boustan et al. (2012), I divide estimated net migration to or from the county from decade t to $t+10$ by population at time t to calculate a migration rate.

Data on County-Level Outcomes: I source a variety of agricultural and manufacturing data at the county level for years 1910-1950 from the data set digitized by Haines et al. (2016), and made available by the ICPSR. Among other numbers, I make use of per-county data on agricultural output, tenancy, farm capital, manufacturing wages and output, population, density and urbanization.

Data on Plantation Counties: If counties that were more or less suitable for agriculture were on a different economic trajectory to those that were not, we might be concerned that this underlying effect is biasing the paper's estimates. In order to control for economic divergence due to a county's exposure to the plantation system, I use data from Brannen (1924) to distinguish between plantation and non-plantation economies.[5] Brannen's data comes from a since-lost agricultural census that records which Southern counties were "plantation counties" in 1910. His data records 270 plantation counties that contained a population of 7,195,600 in 1910, with over 50 percent of the population being black. By contrast, the 298 non-plantation counties contained a population of 6,288,076, with less than 30 percent of it being black (Mandle, 1992). The digitized data is made available online in the program for "When the Levee Breaks: Black Migration and Economic Development in the American South" by Hornbeck and Naidu (2015).

[5]A plantation county is *"a county that is practically always naturally fertile or capable of being made highly productive by the use of commercial fertilizers and manures or by crop rotation"* (Brannen, 1924)

Data on Agroclimatic Conditions: For an additional layer of checks, I also control for a county's broader suitability for cotton agriculture, as well as for its average terrain ruggedness. Land that is more suitable for cotton agriculture might face a delay in the transition from farming to manufacturing due to the higher profit margins associated with farming. Similarly, more rugged terrain presents challenges for operating machinery such as tractors and mules relative to more even terrain that might provide obstacles to the successful mechanization of agriculture. Crop suitability is measured to reflect the maximum potential yield of that crop, as calculated by the Food and Agriculture Organization (FAO) using data on climate, soil type, and ideal growing conditions for that crop. Terrain ruggedness is measured as the standard deviation in altitude across county points, calculated from the USGS National Elevation Dataset.

Data on New Deal Spending: New Deal agricultural policies might have created an incentive for landowners to displace croppers and employ wage labor instead in order to collect agricultural subsidy payments (Whatley, 1983; Sundstrom, 2013). I control for the confounding effect that New Deal Spending might have had on the economic trajectory of Southern states by including county-level data New Deal Spending by Fischback and Liu (2018) on five categories of New Deal spending: Public Works Spending, Agricultural Adjustment Act Spending, Relief Spending, New Deal Loans, and Mortgage and Home Improvement Loans guaranteed by the Federal Housing Administration. All five values represent mean spending from March 1933 through June 1939 divided by population in 1930.

Data on WWII Spending: Similar to the New Deal, World War II military spending ushered an era of economic growth across the United States that likely also contributed to the post-Great Migration economic transformation of the South. I digitize data on the location of investment in structures from from the archives of the War Manufacturing Facilities (U.S. War Production Board 1945). These data provide the most comprehensive view of individual investment projects during mobilization for World War II. I also make use of this data for a modified version of my instrument (Section 4.3).

Finally, I also make use of three decades of digitized US Census data from the Integrated

Public Use Microdata Series (henceforth IPUMS) throughout the paper.

4 Empirical Strategy

In this section, I introduce my study's baseline estimating equation (Section 4.1), construct my main instrument for immigration (Section 4.2), create a modified instrument for immigration (Section 4.3) and report first stage results (Section 4.4).

4.1 Estimating Equation

My empirical analysis contains two parts. In the first one, I estimate the effects of the Great Migration on the agricultural and manufacturing sectors; in the second, I look at the effect of out-migration on various human capital measures for African Americans that remained in the South.

To investigate the effect on agricultural and manufacturing sectors, I start off with a model that relates the share of a Southern county's population that is black to that county's labor market outcomes, controlling for county (α_s) and year (δ_t) fixed effects:

$$y_{st} = \alpha_s + \delta_t + \beta_0 X_{st} + \beta_1 Black_{st} + \epsilon_{st} \tag{1}$$

The regressor $Black_{st}$ is the share of the county's population that is black, so in this case β_1 is the coefficient that tracks the effect of an increase in the black share on our desired outcome.

Then, I stack data for the three decades between 1940 and 1970. Taking the first difference cancels out county- and year-fixed effects and yields the following equation:

$$\Delta y_{st} = \delta_{it} + \beta_1 \Delta Black_{st} + \beta_2 X_{st} + u_{st} \tag{2}$$

The coefficient δ_{it} refers to state-by-year fixed effects. X_{st} is the interaction of year dummies and 1940 county characteristics, and u_{st} is an error term clustered at the county level.

The regressor of interest $\Delta Black_{st}$ the share of the county's population that is black, and so β_1 is the coefficient that tracks the effect of a change in the black share on the evolution of our desired outcome. In my preferred specification, X_{st} is a vector of initial

county characteristics is interacted with year dummies. These initial characteristics include the share of each county's population that was black in 1940, agroclimatic controls, a dummy indicating whether a county was a "Plantation County", as well as New Deal and WWII spending figures. The coefficient of interest, β_1, tracks the effect of changes in black share within the same county over time as compared to other counties in the same state in a given period.

For the second part of the analysis, I look at how black departures impacted the economic return to education, labor market status, and occupational segregation for those African Americans that stayed behind. Due to data limitations, I aggregate data at the 1970 county group level g and estimate the following regression:[6]

$$y_{gt} = \delta_{it} + \beta_1 \Delta Black_{gt} + \beta_2 X_{gt} + u_{gt} \tag{3}$$

As above, the coefficient δ_{it} refers to state-by-year fixed effects, and X_{qt} to the interaction of year dummies and 1940 county group characteristics. u_{qt} is an error term clustered at the county group level.

For this part of my paper, I turn to time-invariant measures of black labor market performance. This means that I am no longer exploiting variation in outcomes across counties and over time, but merely county-level variation near the end of the Great Migration, in 1970. The justification behind this is simply one of data constraints. The 1950, 1960 and 1970 censuses offer varying (and frequently mismatched) levels of geographical granularity. While the 1940 Census identifies respondents' county, the 1950 US Census only identifies counties with a population of at least 100,000. Respectively, the 1960 and 1970 censuses aggregate locations at the "PUMAMINI" and 1970 county group levels. Since there is no simple and straight-forward cross-walk connecting these three overlapping definitions of county groups, I was concerned about possible loss of information due to changing and/or imprecisely calculated boundaries. As more US Census data becomes declassified and precisely georeferenced

[6] According to IPUMS, county groups are geographically contiguous groups of counties with a population of at least 250,000. Most county groups are contained within a single state, though a few contain counties in different states. When faced with such county groups, I assign them to the state most in "common" among their counties.

over time, researching changes in these outcomes over time might become more feasible.

4.2 Shift-Share Design

We might expect that black Southerners' migration choices are correlated with economic conditions in their origin counties. They might be more likely to leave counties with unfavorable economic conditions, such as slower growth or industrialization, as well as counties with favorable economic conditions that are becoming less affordable. In either case, out-migration could be significantly correlated with confounding factors driving positive or negative economic changes across Southern areas. If that is the case, running an OLS regression of the change in black net migration flows on our economic outcomes would lead to biased results and represent an instance in which exogeneity is threatened.[7] To overcome these and similar concerns, I will predict black outflows from Southern county s during decade τ using a version of the shift-share instrument commonly adopted in the immigration literature. In Section 4.3, I will also estimate a further-modified versions of the instrument, using exogenous World War 2 industrial investment and the effect of the 1924 immigration ban to isolate pull factors prompting Southerners to migrate North.

I construct a modified version of the classic immigration shift-share instrument (Altonji and Card, 1991), following Boustan and Tabellini (2019).[8] I use the pre-Great Migration settlement patterns of black Southern migrants in combination with their net-migration into Northern counties to isolate a plausibly exogenous measure of migration. Specifically, this instrument predicts the number of African-Americans moving out of Southern county s between 1940 and decade τ, $m_s^{1940-\tau}$ by interacting the share of African-Americans living in Southern county s in 1920 that had migrated to Northern county n in 1940, ω_{sn} with the rate of African-American migration into each Northern county n for the same time period, $bl_{n\tau}$, scaled by the total population of the county at the beginning of the decade τ. The full

[7] Formally, we can say that the error term is correlated with our outcome of interest, $E(\epsilon_{st}|\Delta Black_{st}, \delta_{st}, X_{st}) \neq 0$.

[8] While the original Altonji and Card (1991) instrument predicts migration *into* a county by interacting the share of individuals that from each county that had immigrated into a given county with net yearly immigration into the county, I interact the share of individuals that had emigrated out of a specific county into another county with net immigrants into the county of out-migration.

equation is:

$$m_s^{1940-\tau} = \sum_n \omega_{sn}^{1920-1940} \times bl_{n\tau} \tag{4}$$

Since we are interested in the share of black migrants (rather than the absolute number of black migrants), I construct the initial share of migrants $\omega_{sn}^{1920-1940}$ as the number of black migrants from each Southern county s living in each Northern county n normalized by the total number African-Americans born in Southern county s that had migrated from the county (including to other counties within the South) according to the 1940 US Census. Formally, I estimate:

$$\omega_{sn}^{1920-1940} = \frac{bl_{sn}}{bl_s} \tag{5}$$

Here, bl_{sn} is number of black migrants from each Southern county s living in each Northern county n, and bl_s is the number of total out-migrants from s. The intuition for the instrument is as follows: from the 1940s onward, the North saw a surge in Southern black migration, during an event typically referred to as the "Second Great Migration." There is strong evidence that these immigrants chose which Northern destinations to migrate to by following individuals from their geographical community, social or familial circles – meaning that their settlements were highly persistent due to social networks and family ties.

As discussed in Boustan (2010), Dereroncourt (2018) and Calderon et al. (2019) among others, as the first African Americans started to move to the North, migration patterns were influenced by the newly constructed railroad network – such as the Illinois Central, which connected various Mississippi counties to Chicago explains why black migrants from Mississippi were disproportionately concentrated in Chicago or St.Louis. The stability of these community enclaves was further reinforced by the process of chain migration during the "First Great Migration"(1915 - 1930), during which over a million blacks migrated to Northern and western cities, often moving to areas with a larger share of individuals from their home state or county. These migrants largely located according to historic settlement patterns–or in other words, settled close to their neighbors. Thus, the standard instrument predicts the location of migrants at the national level ("shift") with historic settlement

patterns ("shares"). 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.

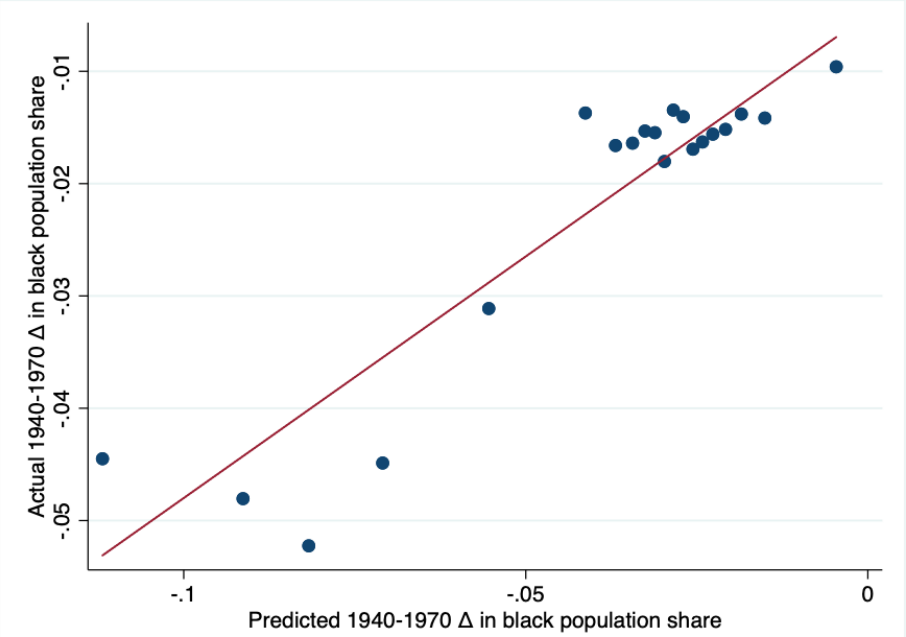
I can further illustrate this with a highly stylized example that illuminates the migration predictions of the model. Say that half of the migrants from New Orleans Parish (or county) who migrated to the North before the 1940s went to Chicago, and the other half went elsewhere. Say also that Chicago received no other black Southerners during this period. For the years in which New Orleans black migrants went North after the 1940s, the instrument would predict that half would go to Chicago and half elsewhere. The half that would go to Chicago would make up the entire predicted inflow of migrants in Chicago for that year. In years where no New Orleans residents decided to make the journey North, the instrument would thus predict no inflow of Southern migrants to Chicago.

Direct measures of county-level in-migration and out-migration are not available for this time period, so I use net migration estimates produced by the ICPSR using forward-census methods, as explained in the Data section. As discussed by Dereroncourt (2018), some minor complications may emerge because the only available figures that may be calculated using the forward-census methods are net migration figures, and some Southern counties experienced positive net migration (in-migration) as opposed to negative (out-migration), despite having a large number of their black population contemporaneously migrate Northern. 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.

Finally, as detailed in the Data section, my shift-share instrument does not appear to be significantly biased by errors in Census matching when estimating the "migration matrix," ω_{sn}. I show this by re-calculating the "migration matrix" using information on the 1940 census about individuals' location in 1935. We can see that the instrument remains effectively unchanged when re-estimated in this way by referencing Table 1, column (5) in Section 4.4.

Figure 1: Instrument Strength

(a) Correlation between instrumented and actual change in black population (for Full Sample)



(b) Correlation between instrumented and actual change in black population (for a Winsorized Sample)

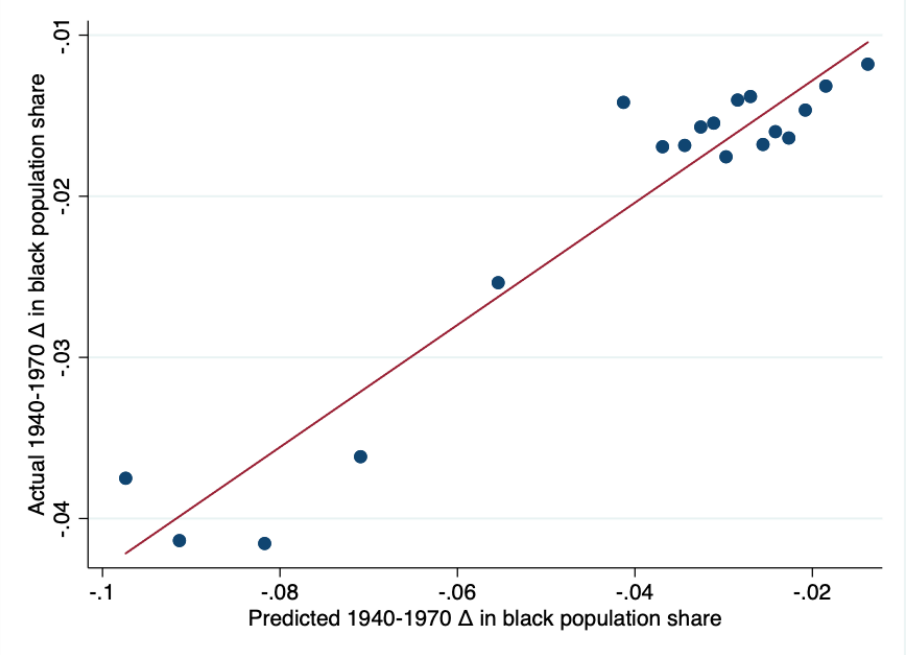
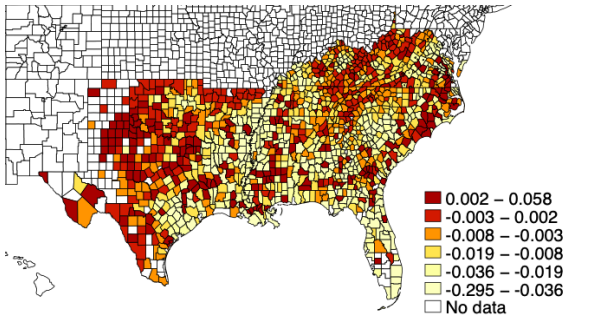
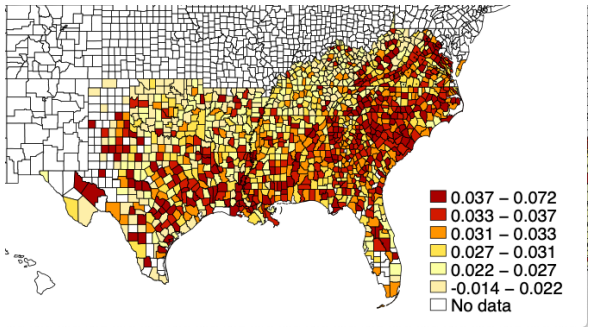


Figure 2: Instrument Strength

(a) Actual Change in Black Share



(b) Predicted Change in Black Share



Source: Data from ICPSR and IPUMS. Calculations are author's own.

4.2.1 Identifying Assumptions and Instrument Validity

The two key identifying assumptions behind the instrument are that, conditional on county and state by-year fixed effects, the economic trajectories of Southern-born African Americans that remained in their county of origin after 1940 must not be correlated with (*i*) the distribution and mix of African Americans that settled in the North between the years 1920-1940 (Goldsmith-Pinkham et al., 2018) and (*ii*) cross-county pull factors systematically related to immigration into different Northern counties in 1940 (Borusyak et al., 2018). In this Section, I will focus on addressing the first concern by conducting a battery of placebo and other pre-trend tests. In Section 4.3, I will ease the second concern by modifying the shift-share instrument such that net migration into Northern counties is predicted solely through variation in the of investment each county received during World War.

First, I show that the instrument is uncorrelated with per-county World War II investment and New Deal spending across four distinct categories: Public Works Spending, Agricultural Adjustment Act Spending, Relief Spending, and New Deal Loans. The regression coefficients, which are universally small and have large standard errors, are reported in Appendix Table A3. This lack of correlation is reassuring, since by most historical accounts, these two events represented the largest harbingers of economic investment and transformation in the South. While we cannot rule out other spurious correlations between the shift-share and events affecting the economic trajectories of Southern counties, this is a reassuring sign of exogeneity with regards to those two events.

Second, I more explicitly address the first concern about the mix of immigrants by showing that the pre-period (1920-1940) change in the black share is uncorrelated with changes in black population predicted by the shift-share instrument. As noted by Jaeger et al. (2018), one potential threat to shift-share instruments for the contemporaneous period is the high persistence of migration between periods. Appendix Figure A1 and A2 and Table A2 illustrate, respectively through scatter plots and reported regression outputs, that there was a break between the two periods. Figure A1 and Table A2 show that is no statistically significant correlation between Northern migration between 1920 and 1940 and migration between 1940 and 1970 as predicted by the "shift-share." Figure A2 additionally shows that there appears to be no relationship between black and white migration during the Great Migration, reassuring us that the "shift-share" captures just black, rather than broader, migration patterns.

These findings complement existing historical evidence that the distribution of black migrants changed significantly between the First (1920-1940) and Second Wave (1940-1970) of the Great Migration, as a series of large scale immigration-related events, including the end of World War I, the passage of the Immigration Acts of 1928, and the Bracero Agricultural program lowered the serial correlation in migration networks and flows across the United States.

Finally, following Goldsmith-Pinkham et al. (2018), Derenoncourt (2019) and Calderon et al. (2019), I construct a version of the "leave-one-out" instrument. I replicate the first-stage analysis by interacting the 1920-1940 Southern immigrant mix [9] ($\omega_{sn}^{1920-1940}$ in Equation (1) in Section 4.2) with year dummies. This test is intended to check whether results are driven by some specific Southern group that happened to settle in specific counties before 1940 and is responsible for a large component of the variation in immigration over time (Goldsmith-Pinkham et al., 2018). Reassuringly, the first-stage results, which are visualized in Appendix Figure A3, remain strong and precisely estimated in the presence of this test.

[9] I aggregate $\omega_{sn}^{1920-1940}$ at the Northern state level, meaning that I test for the share of migrants from different Southern counties settling in different Northern states. Unfortunately, given over 20,000 combinations of Southern origin counties and Northern destination counties, a fully county-level matrix was computationally expensive to compute, as well as difficult to visualize. I also dropped any state that received fewer than 50 migrants, such a small number of observations threatened to bias my estimates.

4.3 Modified Shift-Share: Using Northern Pull Factors

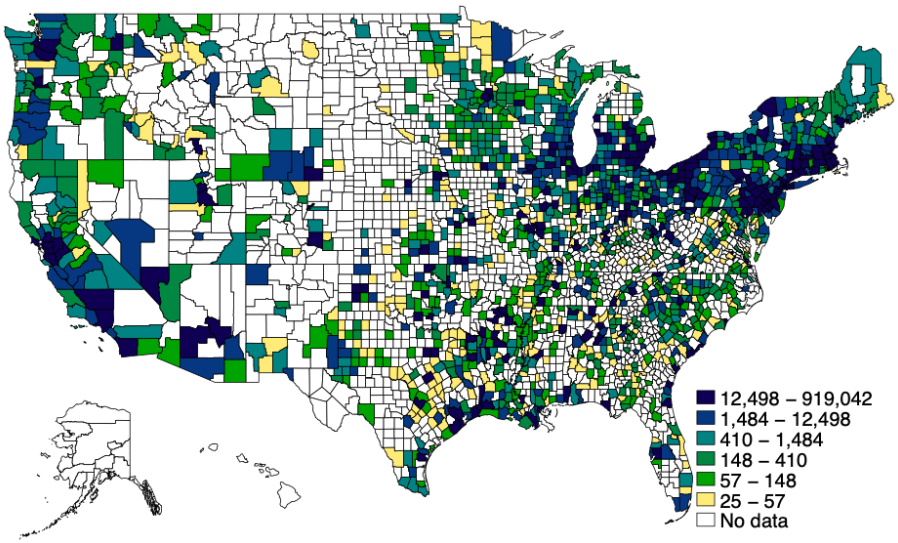
The Second World War saw "approximately 1.6 million civilians [leave] the South for other parts of the country." (McMillen, 1997). The pull of military service and the opportunity for lucrative employment in the defense industry were some of the dominant factors motivating the Northward migration of Southern African-Americans. This wartime boom was heavily dependent on the government's industrial stimulus package, which financed supply contracts and investment in new facilities across the country – and particularly so in the North.

In this section, I will use self-digitized data on the aggregate value of wartime investment to address the concern that post-1940 migration into Northern counties might be correlated with local shocks in Southern counties of origin. If such shocks were in turn correlated with the pre-1940 distribution of Southern born African Americans across Northern counties, then, the identifying assumption would be violated (Borusyak et al., 2018). Using the pull of World War II as a motivating factor, I will construct a modified version of the instrument that predicts migration into Northern counties on the basis of World War II investment.

Figure 3 maps spatial variation in county-level World War II investment, based on the self-digitized data from the U.S. War Production Board's 1945 archives, showing that most of the funding was concentrated in areas outside of the South, such as the Northeast and Midwest. These once-classified data provide the most comprehensive view of individual investment projects during mobilization for World War II (Jaworski, 2017).

To be clear, I do not need to claim that World War II spending was entirely independent of local conditions in order for my identification assumptions to hold. Rather, I require that the conditions that attracted migrants to counties in the North – in this case, World War II-induced manufacturing growth – were not systematically related to the economic trajectories of migrants' counties of origin. As long as, say, more World War II funding wasn't allocated to Northern counties that hosted Southern migrants from poorer states, this should not be a concern. Although a discussion of the mobilization program by Koistinen (2004) suggests the location of new facilities was motivated by the "production of standardized and relatively high quality products," rather than by economic development objectives, it is still possible that these objectives were correlated with growth potential, and thus codetermined economic

Figure 3: Variation in World War II investment



Source: Data from U.S. War Production Board 1945. Calculations are author's own.

growth in counties that received funding. In some instances, lobbying by local communities may have even led to the placement of a war-related plant (Schulman, 1991). However, those idiosyncratic conditions and funding decisions seem like edge cases, and there is no evidence that the placement of these plants across the North would have impacted aggregate economic conditions in the South in a statistically significant way.

Following Derenoncourt (2018), I predict net migration into each Northern county by fitting the following regression:

$$\widehat{bl_{n\tau}} = \beta_0 + \beta_\tau \Delta WW2_n + \epsilon_{n\tau} \tag{6}$$

$$\widehat{bl_{n\tau}} = bl_{n\tau} + \epsilon_{n\tau} \tag{7}$$

In the above equation, $bl_{n\tau}$ is the black net migration rate in Northern county n during decade τ, and $\Delta WW2_n$ is the aggregate value of investment in 1940 dollars in each county, digitized from War Manufacturing Facilities (U.S. War Production Board 1945). $\Delta WW2_n$ is normalized by per-county manufacturing value-added in the prior decade from Haines (2010) in order to account for the concern that the same dollar amount of investment could induce a differential economic effect depending on the baseline size of the economy. First-stage results

from using equation (6) to instrument net-migration are reported on column 4 of Table 1.

After estimating equation (6), I use the predicted rate of migration into Northern counties (in this case, $\widehat{bl_{n\tau}}$) in the place of the observed rate of migration, $bl_{n\tau}$. I then aggregate these (predicted) flows to obtain the predicted number of black migrants from each state in each decade – $m_s^{1940-\tau}$ – as in equation (4). This finally enables me to construct a modified version of the shift-share instrument in equation (4).

First stage results from equation (6) are reported in column 4 of Table 1. The sign and direction of the first stage regression is consistent with all other instruments.

4.4 First Stage Results

Table 1 presents first stage results for the relationship between actual and predicted immigration, after controlling for county- and state-by-year fixed effects. In column 1, the dependent variable is the fraction of immigrants over actual city population, and the regressor of interest is the baseline instrument constructed in equation (2). Column 2 replicates column 1 by adding state-by-year effects interacted with the per-county population in 1940. Column 3 presents a "winsorized" version of the instrument, in which values outside of the 5^{th} and 95^{th} percentile are replaced with values closer to the rest of the set, while column 4 reports the first stage for the instrument based on Northern pull factors. Finally, column 5 re-estimates the instrument using the 1935-1940 matrix to apportion the initial shares. In all cases, the F-stat is very high, and there is a strong and significant relationship between the change in the black share among the population and the instrument. Figures 1 (a) and (b) report the graphical analogue of columns 2 and 4, respectively plotting the non-"winsorized" and "winsorized" relationship between the decline in the black share immigrants and the instrument for immigration.

5 Results

This section presents my study's main results. Section 5.1 presents evidence that black out-migration was associated with a decline in farm tenancy, with farm consolidation, and with the mechanization of agricultural production. Section 5.2 documents the Great Mi-

Table 1: First Stage Results

Dependent Variable:	Change in the Black Share				
	(1)	(2)	(3)	(4)	(5)
$\Delta Black_{\tau}$	0.117**	0.137**	0.0711**	0.223***	0.142**
	(0.049)	(0.051)	(0.024)	(0.041)	(0.016)
F	58.00	58.00	58.24	83.37	53.2
State by year FEs	Yes	Yes	Yes	Yes	Yes
1940 Black Share	No	Yes	Yes	Yes	Yes
Specification	Stacked FD	Stacked FD	Stacked FD	Stacked FD	Stacked FD
N	3,309	3,309	3,309	3,309	3,309
Outcome Mean	0.0402	0.0402	0.0402	0.0402	0.0402

Notes: The sample includes a panel of 1,103 Southern US counties (see Table A.1) that had more than one black resident in 1940. The table reports five stacked first difference regressions. The dependent variable is the decad--al change in the black share, defined as the number of blacks divided by total population in the county. The ma--in regressor of interest is the instrument constructed in the above section (see equation 4). Columns 1 to 5 control for interactions between state dummies and period dummies. Columns 2 to 5 add interactions between period dummies and the 1940 black share. Column 3 winterizes the sample and column 4 recalculates the instrument using WW2 as a pull factor (see equation 6).

Robust standard errors clustered at county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

gration's mixed effect on the manufacturing sector, as measured by manufacturing value added, the share of employment in manufacturing, and manufacturing wages. Section 5.3 investigates the effects on black human capital accumulation, and finds a link between black out-migration and increased economic returns to education for African Americans in the South. Finally, Section 5.4 links the Great Migration to the improving labor market status of African Americans and offers suggestive evidence that the Migration might have led to a decline in a form of literacy-conditional occupational segregation.

5.1 Tenancy and Mechanization

Table 2 presents my main results for the effect of the Great Migration on the agricultural tenancy system across three different categories: the proportion of tenant farms, farm consolidation, and farm mechanization. There is a negative [10] and statistically significant relationship between out-migration and the change in the proportion of tenant farms. This relationship is robust to measuring the proportion of tenant farms in two different ways: as the change in the number of tenant farms as a share of total farms, and as the change in the number of tenant farms per acre of harvested farm.

We note that 2SLS coefficients in columns (1) and (2) are an order of magnitude larger than their OLS counterparts, although both are positive and statistically significant. This implies that black migrants might have endogenously migrated out of areas in which the decline in tenancy was less pronounced. At first, this might seem puzzling: historical accounts stress that the tenancy system was a form of Southern agricultural "backwardness" that harmed African Americans. The unpredictability of tenancy farming could do as much as bankrupt black tenants all the while denying them adequate insurance. However, we must also remember that, for some farmers, the tenancy system was the only form of farm labor they knew. Therefore, it would not be unsurprising if some perceived an initial decline in tenancy as a form of economic downturn, and were intimidated by the prospect of having to enter a different (and possibly less favorable) agricultural arrangement.

There is also a positive and statistically significant relationship between out-migration and farm consolidation. Specifically, out-migration is associated with a relative decline in the number of farms, as well as with a relative increase in the size of farms (as measured in acres per farm). As with the tenancy results, the relative size of the 2SLS estimates implies that black migrants endogenously sorted out of areas in which farm consolidation was more present. This is a reasonable result, if we believe that farm consolidation implied a threat to black farmers' agricultural employment. Although farm consolidation was encouraged by the New Deal's acreage-reduction initiatives (McMillen, 1997) with the end goal of reforming Southern agriculture, it is likely that black migrants felt more threatened by the immediate

[10] Since $\Delta Bl\tau$ measures migration, a positive coefficient indicates a negative relationship between out-migration and our variable of interest.

threat of getting fired due to down-sizing.

Finally, there is also a positive relationship between out-migration and farm mechanization. Out-migration is associated with a larger change in the per-capita value of farm machinery, and with an increase in per-farm tractor adoption. While there the instrumental variables result of out-migration on the change in per capita machinery value is not statistically significant, this could reflect the fact that the per-capita value of machinery is a noisy measure of mechanization since it (i) captures a wide variety of a farm's capital and is (ii) vulnerable to fluctuations in the cost of materials. By contrast, both OLS and 2SLS estimates on tractor adoption are strong and statistically significant. Tractors are a labor-saving innovation that was influential in American agricultural development (Hornbeck and Naidu, 2012; Olmstead and Rhode 2001; Gardner 2002; Steckel and White 2012), but whose adoption lagged in the South. Interestingly, Raper (1946), a contemporary of the Migration, viewed them as one of the predecessors to the successful mechanization of agriculture, and stressed that wherever tractors went, other inventions – such as the mechanic cotton spinner – would follow.

The relative strength of the 2SLS estimate compared to the OLS estimate (in this case, the 2SLS estimate is over 6 times bigger) once again implies that African Americans were more likely to migrate out of counties with lower levels of machinery adoption. This is also puzzling if we believe that mechanization presented a large perceived threat to black farmers' agricultural employment prospects. However, historical accounts have cast doubt on the extent to which black farmers migrated in response to mechanization (Sundstrom, 2013). Another possible reason is that areas that mechanized faster were those in which farmers had a higher amount of reserve capital, since the initial investment necessary to purchase machines would have been quite high. These regions might have also been wealthier, which might plausibly have a negative relationship with emigration.

In Appendix Table A3, I replicate the regressions using the World War II pull-factor instrument from Section 4.3. The direction and significance of the first tenancy and farm size estimates remain consistent with coefficients, if anything, rising in magnitude. This is an affirmation of the robustness of my measure of tenancy and farm consolidation. However, the second tenancy estimate, the number of farms, and both mechanization measures fall

Table 2: Black Out-migration and Agriculture

Change in:	Prop. Tenancy		Farm Consolidation		Mechanization	
	(1) tenant farms	(2) tenant acres	(3) farms	(4) acres/farm	(5) machinery	(6) tractors
A. OLS						
$\Delta Black_\tau$	0.630***	0.482***	1.59***	-2.006***	-0.629**	-6.593***
	(4.47)	(0.349)	(0.270)	(0.234)	(0.301)	(0.571)
B. 2SLS (Shift-Share)						
$\Delta Black_\tau$	9.57***	8.53**	14.78***	-7.173**	-9.963	-41.912***
	(1.15)	(3.85)	(6.59)	(3.27)	(14.08)	(2.084)
C. First Stage						
ΔBl_τ	0.267***	0.182**	0.0572***	0.0572***	0.061*	0.141***
	(0.1045)	(3.85)	(0.0171)	(0.0171)	(0.0291)	(0.006)
State × year FEs	Yes	Yes	Yes	Yes	Yes	Yes
1940 black share × year	Yes	Yes	Yes	Yes	Yes	Yes
N	3,309	1,103	3,309	3,309	1,097	3,259

Notes: The sample includes a panel of 1,103 Southern US counties (see Table A.1) that had more than one black resident in 1940. The table reports six stacked first difference OLS and 2SLS regressions. The dependent variables are the decadal change in (1) the proportion of total farms that are tenant farms, (2) the number of farms that are tenant farms per acre of farmland, (3) the total number of farms, (4) average farmland acre per farm, (5) total value of machinery per farm, and (6) number of tractors per farm. Panel A reports OLS estimates. Panel B instruments the change in the black share with a "shift-share" instument (Equation 4). Panel C reports the First Stage relationship between the actual and instrumented change in the black share.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

out of significance, casting some doubts on the stability of this section's estimates.

5.2 Manufacturing

Models of labor scarcity-induced technical change (Acemoglu, 2012; Clemens et al., 2018) predict that labor scarcity encourages labor-saving technological adoption. Under specified non-homothetic preferences, some theoretical models predict that improvements in agricultural productivity release farm labor to other sectors (Jung, 2019). If the mechanization of agriculture during the Great Migration led to a greater intensity of capital (and lower intensity of labor, respectively) in cotton production as implied by the model we might wonder to what extent (i) it was followed by the sectoral reallocation of some low-skilled agricul-

tural workers into manufacturing and (*ii*) whether there were spillovers that led to greater productivity or capital intensity in that sector.

In this section, I presents results that are consistent with the Great Migration ushering an increase in labor-saving technology in an adjacent manufacturing sector but, unlike Jung (2019), I do not find significant evidence of labor reallocation into manufacturing. In Table 3,I outline the effect of black out-migration on three manufacturing outcomes: the change in per-firm average manufacturing value added, the change in the share of workers employed in manufacturing, and the change in those workers' manufacturing wages. There is a strong relationship between black out-migration and an increase in per-firm average manufacturing value added, which only gets larger in magnitude when instrumented. This result implies that labor productivity in manufacturing increased as a result of the Great Migration and the effect has persisted throughout the Great Migration. Appendix Table A5 similarly reports a positive, albeit imprecisely estimated effect using the pull-factor instrument.

By contrast, the relationship between black out-migration and the change in the share of workers employed in manufacturing is ambiguous. My preferred 2SLS estimates in column (2) imply that a higher rate of black out-migration is linked to an average decline in the share employed in manufacturing; however, re-estimating the equation using the pull-factor instrument in Appendix Table a5 yields a positive and statistically significant relationship. Although the OLS estimates indicate a similarly positive relationship, this is likely biased by the fact that black migrants were more likely to leave counties with a larger share of the population in agriculture (and, by extension, on average a lower share in manufacturing). Although we cannot rule out a reallocation of workers into manufacturing in the years immediately following the onset of the Great Migration, it seems like over the course of the Great Migration, we do not observe a statistically significant reallocation of agricultural workers into manufacturing.

Finally, there is no statistically significant relationship between black out-migration and the change in per-worker average manufacturing wages at the time. Neither the OLS estimate, nor any of my 2SLS coefficients are statistically significant. Although the 2SLS coefficient is large and implies a positive relationship between out-migration and wage growth, it is also imprecisely estimated. We can interpret this in two different ways. On one hand,

assuming that wages are non-sticky proxies of labor productivity, this estimate can cast doubt on the robustness of the claim that black out-migration is associated with an increase in productivity. However, if the productivity increase came from an increase in technology that was capital-intensive and labor saving, it is possible that the demand for labor in manufacturing permanently went down, therefore reconciling an increase in productivity with a decrease in labor demand.

5.3 Economic Return to Education

In order to clarify the discussion about manufacturing productivity in the previous section, I will explore another channel in the form of directed technical change. Specifically, I look at whether greater black out-migration is associated with an increase in the return to education for black workers that remained in the South. Theories of directed technical change predict that, given sufficient elasticity of substitution, technical change will be biased toward more abundant factors (Acemoglu, 1998, 2002). Clemens et al. (2018) find that local labor markets responded to immigration shocks by adjusting the labor-intensity of production technologies. An analogous pattern is observed in a historical context as well. According to Hornbeck and Naidu (2014), counties that were affected by the Great Mississippi Flood in 1927 developed modernized and capital-intensive agricultural technologies following the outflow of the black population. At the same time, formal models of migration show that increased migration options can affect the skill base of the origin country by strengthening incentives to invest in learning and skill acquisition (Dustmann, 2011).

Under certain fixed assumptions, the canonical economics models would predict an increase in the skill premium. Historical sources contemporary to the Great Migration classified the adoption of mechanized forms of agricultural as a form of directed technical change that favored skilled labor. Raper (1946) raised the concern that "[a]s mechanization of cotton production develops, there will be local work only for those who know best how to operate tractors and repair machinery, and who can most readily read and understand instructions." According to the canonical model of skill-biased technical change (Acemoglu and Autor, 2012; Goldin and Katz, 2010; Tinbergen, 1974), if high and low skill labor are gross complements, a relative improvements in a high skill augmenting technology increases the skill

Table 3: Effect of Black Emigration on Manufacturing

Change in:	(1) Mfg. Val. Added	(2) Share in Mfg.	(3) Mfg. Wages
Panel A. OLS			
ΔBl_τ	-2796.736***	-0.0481**	71.710
	(735.53)	(0.0268)	(435.81)
Panel B. 2SLS (Shift-Share)			
ΔBl_τ	-7496.83*	1.158**	-1245.4
	(4549.2)	(0.522)	(3495.9)
Panel C. First Stage			
ΔBl_τ	0.348***	0.0619***	0.0716**
	(0.0575)	(0.0240)	(0.0281)
State FEs	Yes	Yes	Yes
Controls	Yes	Yes	Yes
N	1,494	2,774	2,306

The sample includes a panel of 1,103 Southern US counties (see Table A.1) that had more than one black resident in 1940. The table reports three stacked first difference OLS and 2SLS regressions. The dependent variables are the decadal change in (1) the manufacturing value added, (2) the share of total workers employed in manufacturing, and (3) the per-worker average manufacturing wages. Panel A reports OLS estimates. Panel B instruments the change in the black share with a "shift-share" instrument (Equation 4). Panel C reports the First Stage relationship between the actual and instrumented change in the black share.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

premium.

In this section, I will empirically explore the extent to which out-migration episode could have affected the return to human capital. To assess the impact of out-migration on return to human capital, I first estimate the relative return to education of blacks for each county, using the 2% sample[11] of the 1970 Census. Employing estimated wage returns as outcome variables, I test whether the return to education of blacks decreased relative to whites in proportion to the out-migration of African Americans from the following standard Mincer wage equation

$$w_{sj} = \alpha_c + \beta_{c0} Edu_{sj} + \beta_{c1} Black_{sj} + \beta_{c2} Black_{sj} \times Edu_{sj} + \gamma_{c0} Exp_{sj} + \gamma_{c1} Exp_{sj}^2 + X'_{sj}\delta_{sj} + \epsilon_{sj} \quad (8)$$

where for each county group s[12], w_{sj} and Edu_{sj} are log weekly wage[13] and years of schooling for individual j,[14] and, as is typical in the Mincer literature, Exp_{sj} is $max(0, Age_{sj} - 6 - Edu_{sj})$. The dummy variable $Black_{sj}$ is equal to 1 if individual j is black and 0 otherwise, and X_{sj} is a vector of other individual controls including dummy variables for sex, household head status, and marital status.[15] This equation is estimated separately for each Southern county group s. After estimating the Mincer equation, I adopt the measured coefficients as outcome variables using the same shift-share strategy as employed by the rest of the paper.

I chose to estimate the return to education using a Mincer equation, due to the model's recent reappraisal. As one of the first formal models in labor economics to realize that

[11]Technically there is no 2% sample in IPUMS but two 1% samples of 1970 at the neighborhood level, which do not overlap. I use an aggregation of both samples and refer to it throughout as a 2% sample.

[12]The most granular geographical unit in the 1970 census is the "county group" which is an aggregation of neighboring counties that form an economically contiguous area with at least 250,000 residents. There are 404 county groups, with a surjective mapping between counties and county groups.

[13]The existing evidence generally supports the log-linear specification. For a more extensive overview of the debate, see Card (1999), Grossbard (2006), and Heckman (2008).

[14]I only include workers working more than 29 hours a week, in order to compare individuals with a similar willingness to utilize human capital in the labor market. After constructing the sample, I exclude county groups with fewer than 20 observations.

[15]Although Kniesner, Padilla and Polachek (1978, 1980) show that aggregate economic conditions affect measured rates of return to schooling, and excluding them might bias the Mincer estimates, we are already controlling for labor-market heterogeneity this by running the equation separately in each county group (which is constructed with the intention of providing a contiguous commuting/economic zone).

"choices among alternative [work options] differing in the probability distribution of the income they promise" (Polachek, 2007), the Mincer equation is arguably less sophisticated than newer labor economics techniques: economists have criticized the model for failing to account for the non-linearity of education's effect on earnings[16] (Card, 1999), as well as for modeling that the percentage increase in earnings attributable to schooling need not be independent of an individual's school or experience (Heckman, Lochner, and Todd, 2003). However, the model has been reappraised in a number of recent papers (Polachek, 2007), due to the attractiveness of its central features. Not only is it based on a micro-founded structural model of investment in human capital, but it also fits the data remarkably well in most contexts (Lemieux, 2006).

In particular, given that the return to education is estimated for 1970 – a period of time that predates the widespread attainment of post-secondary education, the dot-com boom, as well as the opening-up of the American economy to international exports – we have reason to believe that the "sheepskin effect" exerts less of a discrete impact on individuals' earnings potential. Moreover, my research design is such that the analysis is conducted on a county-by-county level – thereby assuaging the two critiques of the standard modified Mincer wage regression as applied to the analysis of race wage disparities (Neal and Johnson, 1996) – that (i) blacks and whites typically have different levels of human capital, even conditional on observed years of schooling, and, in any event (ii) completed years of schooling is an endogenous choice variable that will depend on any number of factors, including the quality of schooling to which a young person has been exposed.

Table 4 indicates that increased black out-migration is simultaneously associated with a flat wage penalty due to race (β_{c1}) as well as with increased returns to education (β_{c2}) for African-Americans in Southern counties in 1970. At the same time, there is no statistically significant association with return to education when the Mincer equation (8) is re-estimated for the average of the entire population, without specifying black and black $\times$ education dummies. This implies a statistically insignificant effect (or otherwise statistically different

[16]This effect is commonly known as the "sheepskin effect," referring to the existence of separate wage premiums for fulfilling the final years of elementary school, high school, or college versus merely completing an extra year of schooling.

Table 4: Effect of Black out-migration on Return to Skills

	Educ.		Black (β_{c1})		Black × Educ. (β_{c2})	
Dependent Variable	(1) OLS	(2) IV	(3) OLS	(4) IV	(5) OLS	(6) IV
$\Delta Black$	2.907	-0.557	1.323*	23.17**	-0.142*	-2.050**
	(4.47)	(0.349)	(2.33)	(3.36)	(-2.65)	(-3.07)
Cons.	(0.339)	(0.353)	-0.0718	-0.453*	-0.0409*	-0.00497
	(0.1045)	(0.114)	(-0.42)	(-2.58)	(-2.54)	(-0.29)
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	414	414	414	414	414	414

The sample includes a panel of 138 Southern US county groups that had more than one black. resident in 1940. The OLS column reports OLS estimates while the IV column estimates 2SLS estimates using the instrumented "shift-share" decline in the black share.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

effect) on the return to education for other racial groups. The coefficients β_{c1} and β_{c2} respectively estimate average wage loss among African Americans due to race, and through the differential return to education for black as opposed to non-black individuals in the sample. All regressions include a number of controls including state fixed effects, a plantation county dummy, the pre- and post-Great Migration share of the population that was black and urban, geoclimatic controls for suitability to cotton agriculture and mechanization, and New Deal Spending.

These results are consistent with the theory that a period of skill-biased technical change spurred on by mechanization and manufacturing growth incentivized occupational upgrading for African Americans, without finding evidence of other improved labor force conditions on average. Note that, unlike the results presented in Sections 5.1 and 5.2, the estimates here suffer from a smaller number of observations,since results are aggregated to the county

group level.[17]. Moreover, the absence of time-variation means that these regressions are less econometrically robust than those in Sections 5.1 and 5.2. Still, they provide suggestive evidence in favor of this mechanism.

As a counterpoint, we might posit that a large part of this variation could have been instead driven by post-desegregation improvements in the Southern education system.[18] I try to address this concern by restricting the equation for individuals between the ages of 25 and 65, which implicitly drops individuals born before desegregation who would have seen the greatest improvement in their education quality. Additionally, I explore the relationship between Southern out-migration and educational quality, using data on the number of segregation-era black teachers, black schools, and per-capita educational spending as proxies. Although data is only available for a set of 384 Southern counties, making it difficult to rule-out alternative explanations, the regressions indicate that is no statistically significant association between black out-migration and educational quality before desegregation. Fitted scatter plots illustrating the absence of a relationship are reported in Appendix Figure A4.

5.4 Occupational Outcomes

5.4.1 Labor Market Status of African Americans

We might naturally wonder to what extent the increased return to education for African Americans in high out-migration counties led to occupational upgrading. Although the higher rewards to skill acquisition should imply that, ceteris paribus, skilled blacks in high out-migration counties earned more, it is possible that structural barriers to education and employment provided a large enough number of black workers from realizing an education

[17]Our observations include 138 unique Southern county groups, compared to a panel of over 1,000 Southern counties observed over time for Sections 5.1 and 5.2)

[18]There is a rich literature showing that an improvement in Southern schooling quality following desegregation improved African Americans' labor market performance. For example, Collins and Yoon (2005) identify Southern individuals born in the 1920s and 1930s from 1970 Census data to show "that Southern-born blacks who finished their schooling just before effective desegregation occurred in the South fared poorly compared to Southern-born blacks who followed behind them in school by just a few years."

to change the aggregate distribution of skilled laborers. Moreover, the existence of a causal effect might be of great social interest since, in the pre-Migration Southern economy dominated by the tenancy system, blacks undoubtedly occupied a lower average rung than whites. Among Southern farm operators in 1920, approximately 60 per cent of whites were owners, compared with only 22 per cent of blacks; conversely, only 11 per cent of whites were croppers, compared with over 37 per cent of black operators (Wright, 1986).

Using the 2 % sample of the 1970 Census, and restricting my attention to the number of respondents between age 25 and 65 whose occupation category is identified, I compute the share of whites and blacks working in low- and high-skill occupations respectively.[19] In short, I am looking at how out-migration might have affected the within-race distribution of employment, rather than the relative racial composition within low- and high-skill occupations.

Table 5 shows that higher levels of out-migration are associated with a higher proportion of both black and white workers in high-skill occupations. Both the OLS and 2SLS estimates point to a positive and statistically significant relationship. For both white and black workers, the 2SLS estimates are higher than their OLS counterparts, implying that black workers endogenously emigrated out of communities that were seeing relative increases in the share of whites and blacks occupying high-end jobs. This is not necessarily surprising: while these regressions indicate relative improvements for the tail end of the occupational distribution, they do not characterize the rest of the occupational distribution. They give us no information about whether there was a relative improvement in the share of middle-skill blacks. Moreover, increases in the shares of very high-skilled workers are often accompanied by stark increases in the prices of rent and amenities, in a process known as "gentrification" (Edlund et al., 2015).

While the OLS estimates imply a decrease in the share of those of either race holding low-skill occupations, the 2SLS estimates are insignificant and imprecisely estimated. Since the

[19]Skill-level of occupations are defined by *edscor50* variable which "indicates the percentage of people in the respondent's occupational category who had completed one or more years of college". Thresholds of the *edscor50* variable for low- and high-skill occupations are determined at 25% and 75% percentiles of the entire South sample. In both black and white samples, the majority of low-skill workers are farm laborers.

OLS estimates are likely to suffer from endogeneity as blacks might have been more likely to leave counties in which they held more low-skill occupations, we cannot infer anything about the relationship. Therefore, I fail to find evidence of a significant effect of out-migration on the share of low-skill blacks and whites.

Table 5: Effect of Black out-migration on Black Labor Market Integration

	White		Black	
Share in:	high-skill occup.	low-skill occup.	high-skil occup.	low-skill occup.
Panel A. OLS				
$\Delta Black_\tau$	-0.187***	0.215**	-0.0703**	0.138**
	(0.0500)	(0.0677)	(0.0281)	(0.0640)
Panel B. 2SLS				
$\Delta Black_\tau$	-0.965**	0.323	-0.707**	0.0730
	(0.426)	(0.414)	(0.334)	(0.807)
Panel C. First Stage				
$\Delta Black_\tau$	0.414**	0.414**	0.414**	0.414**
	(0.192)	(0.192)	(0.192)	(0.192)
State FEs	Yes	Yes	Yes	Yes
County Controls	Yes	Yes	Yes	Yes
N	402	402	402	402

The sample includes a panel of 134 Southern US county groups that had more than one black resident in 1940. The dependent variables are (1) the share of full-time working white workers in high-skill and (2) low-skill occupations respectively. Columns (3) and (4) repreat those estimations for black workers. Panel A reports OLS estimates. Panel B instruments the change in the black share with a "shift-share" inst--rument (equation (4)). Panel C reports the first stage relationship.

Robust standard errors clustered at county group level are shown in the parentheses.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

5.4.2 Intuition from a Model of Occupational Segregation

Occupational segregation on the basis of race was a notorious and distinguishing feature of the Southern economy in the early 20^{th} Century, reflecting labor-market discrimination by employers or co-workers due to costly interracial contact, external economies associated with racial homogeneity in industrial labor markets, and the self-reinforcing effect of racial stereotypes (Sundstrom, 2013).

In the following sections, I hope to provide a two-tiered approach to how black out-migration affected segregation. First, I will outline the intuition behind a straight-forward model that formalizes intuition about the relationship between out-migration, skill differentiation, and the changing cost of occupational segregation. For the full model, please consult Appendix A1. Then, I want to investigate the empirical relationship between out-migration and observed occupational segregation as measured by the dissimilarity and Gini indices. I deploy Monte Carlo methods that allow me to shuffle race within occupational categories in order to decompose occupational segregation as occurring either due to discrepancies in black vs. white literacy levels, or due to residual factors that persist after conditioning on variation in literacy.

Prior to the Great Migration, non-agricultural labor markets in the South were racially segregated along both industrial and occupational lines. Although black farmers overwhelmingly grew the cotton used in mills, the workers who spun and wove it into cloth were far more likely to be white. When blacks were employed in the cotton mills at all, they almost never operated textile machinery, "but served as unskilled laborers tasked with cleaning the plant and moving materials" (Sundstrom, 2013). This is in contrast to other disproportionately black industries, such as "sawmills, tobacco processing, and iron and steel production" (Sundstrom, 2013) that saw more equal division of labor. These racial patterns were neither a reflection of the geographical dispersion of the races, nor a consequence of gaps in education. All-white factories and industries could still be found in areas with large numbers of black workers (Wright, 1986), and operating a cotton-mill was not a skill-intensive job. By all reliable estimates, white cotton-mill operatives most needed no more formal schooling to do their jobs than did black foundry workers. Similarly, occupational segregation was

another permanent fixture of economic life. Among Southern farm operators in 1920, some 60 per cent of whites were owners, compared with only 22 per cent of blacks; on the contrary, only 11 per cent of whites were croppers, compared with 37 per cent of blacks (Wright, 1986; Sundstrom, 2013).

Segregation of individual business firms or establishments is fully consistent with the predictions of standard economic models of labor-market discrimination by employers or co-workers (Becker, 1971; Arrow, 1973). If interracial contact is costly, it is profitable to minimize that contact by employing only blacks or only whites in any given workplace. But segregation by entire industries is more puzzling: why was it not the case that some factories were all-white and others all-black, rather than an entire industry?

This segregation might reflect employers' "distaste" for employing black workers as well as external, network-based economies associated with racial homogeneity in low-skill labor markets. Although employer learning could overturn inaccurate stereotypes about the suitability of black workers (Bohren et al., 2019), this might have been more challenging in highly segregated markets, as employers would not be as easily able to track and compare whites' and blacks' performance in similar tasks over time. Even with experience, racial stereotypes could be self-reinforcing if they resulted in feedback effects on the effort and skill acquisition of workers.

At the same time, in many low-skill industries and occupations, workers learned of job opportunities through relative and acquaintance networks, which, due to a low rate of interracial marriage, would have been racially segregated. If search and hire costs decreased with the size of one's network, firms within given industries might have found it more profitable to focus their search on workers within one racial group. A far more likely scenario is that segregation was further reinforced by segregated training patterns. While operating a spinning machine or power loom did not require a formal education, it did require a period of learning in the presence of other experienced workers and supervisors, who would presumably be of the same race. For a black worker to get started in an all-white industry would thus have required finding a segregated, all-black plant with a trained work force (Sundstrom, 2013).

The Great Migration could have affected the occupational segregation of black and white workers by decreasing taste-based discrimination or by decreasing the size of local networks.

In the model in Appendix Section A, I highlight that, given a deficit of skilled workers, processes that favor the acquisition of skills among a discriminated group can make them more attractive to a marginal, discriminatory employer, and therefore decrease workplace segregation. Such a mechanism would reinforce our observations about the Great Migration and skill acquisition among Southern blacks that remained behind.

5.4.3 An Empirical Analysis of Occupational Segregation

To measure the effect of the Great Migration on occupational segregation between black and white workers, I use a 2% sample of the 1970 Census to compute (i) the Gini index and (ii) dissimilarity index proposed by Duncan and Duncan (1955). The Gini index is a measure of statistical dispersion that is a cornerstone of the inequality literature, and the Duncan dissimilarity index is defined by the following formula:

$$D_{obs} = \sum_{k} |\frac{b_k}{B} - \frac{w_k}{W}| \tag{9}$$

where b_k and w_k are the number of blacks and whites in each three-digit occupational category k, classified according to the *occ*1950 variable provided by IPUMS-USA.[20] The figures B and W respectively denote the total number of blacks and whites in the sample that are in the labor market. A simple intuition behind the index is that it measures the share of black workers that would need to change their occupational category for the distribution of black and white workers across the occupations to be equal. So, while $D_{obs} = 0$ would indicate that black and white workers are identically distributed across the occupations, $D_{obs} = 1$ would denote that they are fully segregated.

While both the Gini and dissimilarity indices provides a robust measure of observed occupational segregation[21], I am interested in separating out the extent of segregation explained by the human capital distribution versus other race-related factors (such as taste-based, or stastistical discrimination). I apply the methodology of Jung (2019) and Hellerstein and

[20]The *occ1950* variable classifies occupations according to the US Census' *Alphabetic Index of Occupations and Industries: 1950*.

[21]Both indices are scale invariant, meaning that they are insensitive to the proportions of each group in the workforce. This is an attractive feature because it allows for a degree of consistency when comparing different county groups with widely varying workforce numbers.

Neumark (2008) to separate out these two drivers behind racial segregation. First, I divide workers into eight cells, according to their age (younger or older than 30), their educational achievement (have or have not graduated from high school), and their assigned sex (male or female). I then re-shuffle whether an individual is black or not within each of these eight cells, and compute the segregation measure from the newly generated sample. As in Jung (2019), I repeat the simulation 100 times for each county group and take the average of each county group, D_c. Because D_c is the average of random segregation measures conditional on human capital, the normalized difference between D_{obs} and D_c can be said to represent the extent of racial segregation which is not explained by variation in literacy. This means that if $D_{obs} = D_c$, occupational segregation between the races is entirely explained by literacy of workers. To estimate the extent of segregation which is not accounted for by literacy, I construct the following effective segregation measure

$$D_* = \frac{D_{obs} - D_c}{1 - D_c} \tag{10}$$

which is normalized to take 1 as the maximum value. To ensure that the results of the estimation are robust to the choice of segregation measure, I apply an identical process in computing the Gini segregation index.

In Table 6, I show that out-migration is associated with a decline in the form of segregation that is accounted for by variation in literacy (D_c). That is, once I randomize for race within literacy brackets, out-migration had a statistically significant negative effect[22] on occupational segregation. Assuming that the results are econometrically robust, the finding suggests that the mechanism through which the Great Migration affected segregation, was at least partially tied to variation in black workers' literacy levels. This is in line with previous findings about the co-terminal increase in black workers' education premium and high-occupation positions. Curiously, there appears to be no significant heterogeneity based on whether a county had above- or below- median segregation levels. In Appendix Figure A5, split the counties in the study based on their level of segregation, and compute coefficients and confidence intervals for 2SLS regressions. With the exception of a decline in literacy-

[22]Once again, a lower $\Delta Black$ value signifies a higher level of out-migration out of the county between the years 1940 and 1970.

conditioned segregation in high-segregation counties that experienced most out-migration, the results are insignificant.

Curiously, the Great Migration does not appear to have a statistically significant relationship with the "residual" variation in occupational segregation that is not explained by literacy. This could be an effect of the small sample used: with a panel of 138 county group observations, it is difficult to draw robust causal inferences. It is also possible that splitting the group into skilled and unskilled samples on the basis of completing secondary education does not accurately capture the magnitude of different education wage premiums. If, for example, an increasing number of black workers graduated from college without changing the high school graduation rate, yet graduating from college was what allowed black workers to enter higher skill, lower segregation workplaces, then our measure of skill acquisition may not adequately capture the effect.

This observation also raises questions about the political economy of discrimination. Research by Boustan and Tabellini (2019) indicates that the Great Migration increased support for anti-segregation candidates. If the Great Migration changed electorate preferences toward segregation, it is reasonable that this shift should have trickled down to employers' taste for segregation. Why might this political shift not led to a decrease in "residual" segregation, as measured by this index? This question suggests a fruitful avenue for future political economy research about the effectiveness and spillover effects of desegregation measures.

Table 6: Effect of Black Out-migration on Segregation: All Counties

	Observed		Variation in Literacy		Residual	
Dependent variable	(1) D_{obs}	(2) $Gini_{obs}$	(3) D_{cr}	(4) $Gini_{cr}$	(5) D_*	(6) $Gini_*$
Panel A. OLS						
$\Delta Black$	-0.0250	-0.0330	0.0434	0.0729	-0.0718*	0.111**
	(0.0302)	(0.0342)	(0.0319)	(0.0520)	(0.0414)	(0.0542)
Panel B. 2SLS (Shift-Share)						
$\Delta Black$	0.4097	0.6275	1.247**	2.064**	-0.708	-1.114
	(0.658)	(0.758)	(.05577)	(1.066)	(0.560)	(0.678)
Panel C. First Stage						
$\Delta Black$	0.414**	0.414**	0.414**	0.414**	0.414**	0.414**
	(0.154)	(0.154)	(0.154)	(0.154)	(0.154)	(0.154)
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	414	414	414	414	414	414

The sample includes a panel of 134 Southern US county groups that had more than one black resident in 1950. The sample consists workers aged 25 to 65. D_{obs} is the observed occupational segregation index and D_c is the conditionally random index after randomly reassigning race within literacy groups. D indicates the extent of occupational segregation after filtering out literacy. Panel A reports OLS estimates. Panel B instruments the cha-nge in the black share with a "shift-share" instrument. Panel C reports the first stage relationship between the actual and instrumented change in the black share.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

6 Conclusion

My findings are consistent with the idea that a period of labor scarcity-induced, skill-biased technical change spurred on by mechanization and manufacturing growth incentivized occupational upgrading for African Americans in higher out-migration counties. Using a modified version of the canonical "shift-share" instrument, I show that the Great Migration is associated with the modernization of Southern agriculture and manufacturing and the improvement of African Americans' occupational outcomes as measured by an increase in the return to education, an increase in the share of African Americans holding high-skill jobs, and a decrease in occupational segregation as explained by variation in literacy. All my findings jointly imply that this episode created incentives and rewards for human capital acquisition among African Americans in the South.

As is the case with many historical events, the Great Migration occurred at the same time as other transformative events that were bound to have a permanent effect on the Southern economy. Although I have tried to control for the two main ones – World War II and the New Deal – I recognize that there is no way to entirely disentangle the effects of one from the other. That is not, as far as I am concerned, the goal of economic history. My research merely sheds light one of many channels that might have facilitated education and more equitable work conditions for black Southerners.

I conclude with some thoughts about the potential direction of future work. First, I only focused on the contemporaneous affect of the Great Migration. A natural question concerns what the effect of the Great Migration might be on cohorts that were born during this episode of Southern out-migration. This would provide a natural counterpart to the Derenoncourt (2019) paper that estimates changes in socioeconomic mobility for cohorts of children born during the Great Migration. It would also complement this paper by further clarifying the persistence of the higher return to skill for black labor, and propose links between short-term, medium-term, and long-term effects of a large-scale out-migration episode. Second, this paper does not concern itself with the political economy response to black out-migration, as historical evidence seems to indicate that the Great Migration played a central role in politically empowering the black population. Finally, more information is needed about

the mechanisms through which Southerners that left for the North might have affected the human capital acquisition of their relatives in the South. Remittances and immigrant trade networks play a role in the growth of developing countries, and it is not unimaginable that they might have encouraged economic transformation in this instance.

A Appendix – Formal Model

A.1 A Model of Skill Acquisition and Segregation

As in Autor (2011), assume that there exists a continuum of identical firms and each firm i formally maximizes:

$$U_i = pF(N_b + N_a) - w_a N_a - w_b N_b - h_a(N_a) - h_a(N_b) - d_i(N_b)N_b \tag{11}$$

where p is the price level, F is the production function, w_x and h_x are the wage level and hiring cost respectively for group group $x = \{a, b\}$, and N_x is the number of workers in each group working for the firm. I assume that the hiring cost h_x is exponentially decreasing in N_x ($h'_x(N_x) < 0$ and $h''_x(N_x) > 0$), as the hiring process is assisted by family and acquaintance networks. Finally, $d(N_a)$ represents an employer's "discrimination coefficient" – a measure of the disutility associated with having an additional worker from the group they discriminate against. We can think of $d(N_a)$ as a function increasing in N_a ($\frac{\partial d(N_a)}{\partial N_a} > 0$ because of the costs of racial conflict in the workplace. Assume that there is a continuous distribution of the discrimination coefficient that varies across the population of employers with a CDF $G(d)$.

Since black and white workers are perfect substitutes in production and utility is linear in discriminatory distaste, each employer simply hires the type of labor that is "cheaper" to him. The employer will only hire workers from group b if $w_a - w_b \geq d(N_a) + [h'_b(N_b) - h'_a(N_a)]$, and vice-verse if the inequality is reversed.

Now, assume the Great Migration induces a reduction in the number of black workers, such that $N_{b2} < N_{b1}$. By extension, $h(N_{b2}) > h(N_{b1})$ and $h'(N_{b2}) > h'(N_{b1})$. Make the additional assumption that there is no change in the number of white workers, and therefore no change in $d(N_a)$ for a given employer. Since black and white workers are still perfect substitutes, the new threshold for employer i to hire workers from group b is $w_a - w_{b2} \geq d(N_a) + [h'_{b2}(N_{b2}) - h'_a(N_a)]$. Because, ceteris paribus, $[h'_{b2}(N_{b2}) - h'_a(N_a)] > [h'_{b1}(N_{b1}) - h'_a(N_a)]$. Therefore, after the Great Migration, an employer will only hire black workers for a lower $d(N_a)$ in the "discrimination coefficient" distribution – effectively implying that a smaller number of firms will hire black workers as opposed to below. Still, there is no change to the extent of discrimination.

When might an out-migration episode change levels of occupational segregation? We will assume, as in Section 5.3.1, that out-migration affects black workers' opportunities to invest in their skills or their return to education (thereby increasing incentives to invest in skills). Suppose that, as before, individuals are either black (b) or white (w), but they are now differentiated into two skill groups: those that possess either high (H) or low (L) skill. All firms have identical production functions with constant returns to scale in capital, high-skilled labor and low-skilled labor:

$$U_i^H = \pi_i - h_a(N_a^H + N_a^L) - h_b(N_b^H + N_b^L) - d_i(N_a^H + N_a^L) \times [N_b^H + N_b^L] \tag{12}$$

where

$$\pi_i = pF(K, N_b^H + N_a^H, N_b^L + N_a^L) - rK - \sum_{\{x,S\}} w_x N_x^S \tag{13}$$

As above, p is the price level, F is the production function, w_x^T and h_x^T are the wage level and hiring cost respectively for group group $x = \{a, b\}$ and skill level $S = \{H, L\}$ and N_x^T is the number of workers in each respectively indexed skill and race group working for the firm. Each firm requires a fixed and equal number of low and high-skilled workers, and compensates them according to the average marginal product of their skill type.

Assume that the number of high-skilled workers, $N^H < N^L$, and there is always a deficit of high-skilled workers relative to each firm's needs. Ordering firms by their d_i, high-skilled blacks are allocated first to the least discriminatory firms. The last firm to be matched with a high-skilled black worker is the marginal discriminator. In equilibrium, the black-white wage gap must be just high enough to compel the marginal discriminator to match with high-skilled blacks but not high enough to induce the next most discriminatory firm to do the same.

Some randomly chosen portion of firms with d_i greater than that of the marginal discriminator pair with white high-skilled workers, until those run out. However, because $N^H < N^L$, some firms with d_i higher than the marginal discriminator's will pair with white low-skilled workers, instead of either black or white low-skilled workers. However, firms face a long-term penalty for paying less productive workers (low-skilled whites) the average of the marginal product of high-skilled workers. If we endogenize d_i, we can claim that there is an incentive

to lower their d_i and compete for the opportunity to hire low-skilled blacks. We continue to assume that no firm prefers to replace low-skilled workers with those of the other race.

Therefore, the model predicts that change in the skill mix can introduce heterogeneity in the racial mix of some firms in two ways: first, by facilitating the interaction of high-skilled blacks and low-skilled whites by less-discriminatory firms (low d_i firms) and second, by incentivizing firms to lower their d_i for an opportunity to get higher productivity workers. This leads to a decrease in segregation.

B Appendix – Additional Figures and Tables

B.1 Descriptive Data

Table A1: List of Southern States

Alabama	North Carolina
Arkansas	Oklahoma
Florida	South Carolina
Georgia	Tennessee
Kentucky	Texas
Louisiana	Virginia
Mississippi	West Virginia

Note: The table presents the list of Southern states considered in my analysis. As Boustan (2010) and Calderon et al. (2019), I assign Maryland and Delaware to the North, as they were net recipient of black migrants during this period.

B.2 Instrument Exogeneity: Additional Pre-Trends

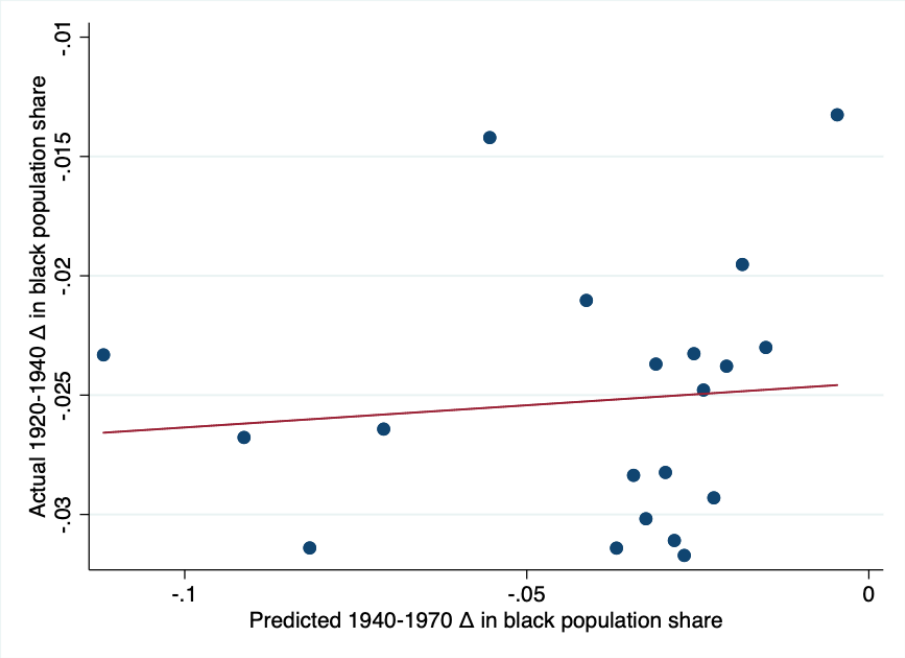
Table A2: Pre-Trends: Relationship between Net Migration in 1920-1940 and Instrumented Black Population Change in 1940-1970

Change in Black Share in:	(1) 1920-1940	(2) 1920-1930	(3) 1930-1940
$\Delta Black_\tau$	-0.0062	-0.00298	-0.0033
	(0.0173)	(0.0130)	(0.0097)
State by year FEs	Yes	Yes	Yes
1940 Black Share × year	Yes	Yes	Yes
Specification	Stacked FD	Stacked FD	Stacked FD
N	3,309	3,309	3,309

Regression of the instrumented change in the black share (1940-1970) on the pre-period change in the black share (1920-1940). Standard errors clustered at county level.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Figure A1: Pre-Trends: Relationship between Black Population Change in 1920-1940 and Instrumented Black Population Change in 1940-1970



Source: Data from ICPSR and IPUMS. Calculations are author's own.

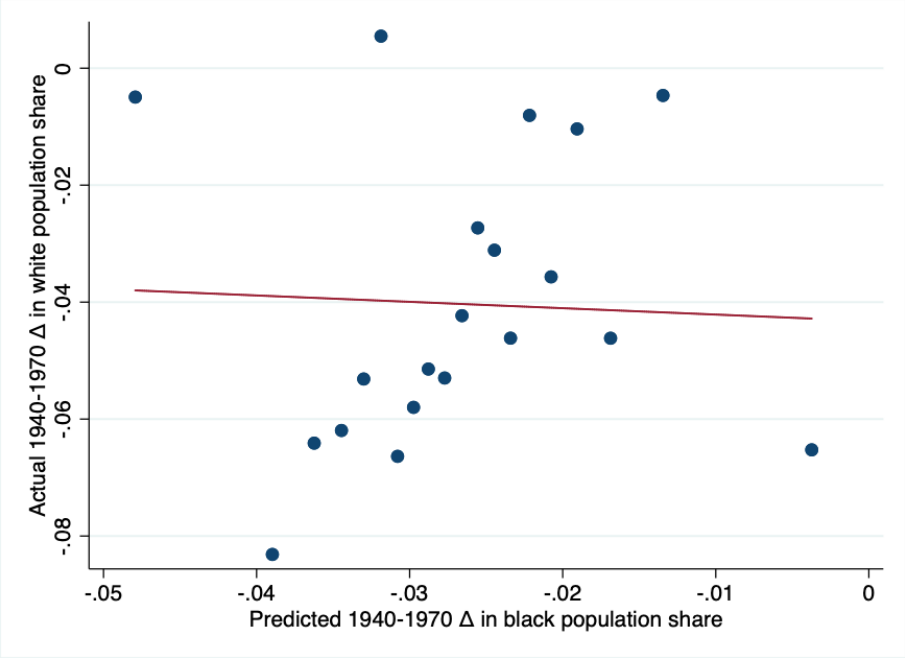
Table A3: Pre-Trends: Relationship between Instrument and WWII and New Deal Spending

Dependent Variable: *log*	(1) WW2 investment	(2) Agric. Adjustment	(3) Public Works	(4) Relief	(5) Loans
$\Delta Black_{\tau}$	1.67	0.736	1.085	0.197	0.832
	(1.84)	(0.725)	(0.882)	(0.146)	(0.645)
State by year FEs	Yes	Yes	Yes	Yes	Yes
1940 Black Share × year	Yes	Yes	Yes	Yes	Yes
Specification	Stacked FD	Stacked FD	Stacked FD	Stacked FD	Stacked FD
N	3,309	3,309	3,309	3,309	3,309
Outcome Mean	2.89	10.88	11.00	2.48	11.47

Regression of the instrumented change in the black share (1940-1970) on WW2 investment (1945) and New Deal spending.

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

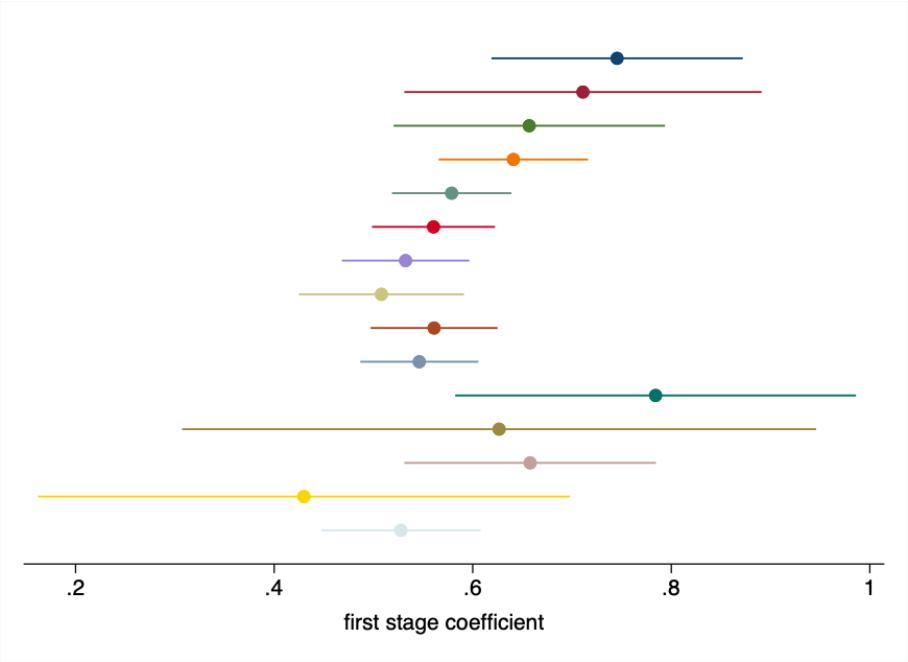
Figure A2: Pre-Trends: Relationship between White Population Change in 1940-1970 and Instrumented Black Population Change



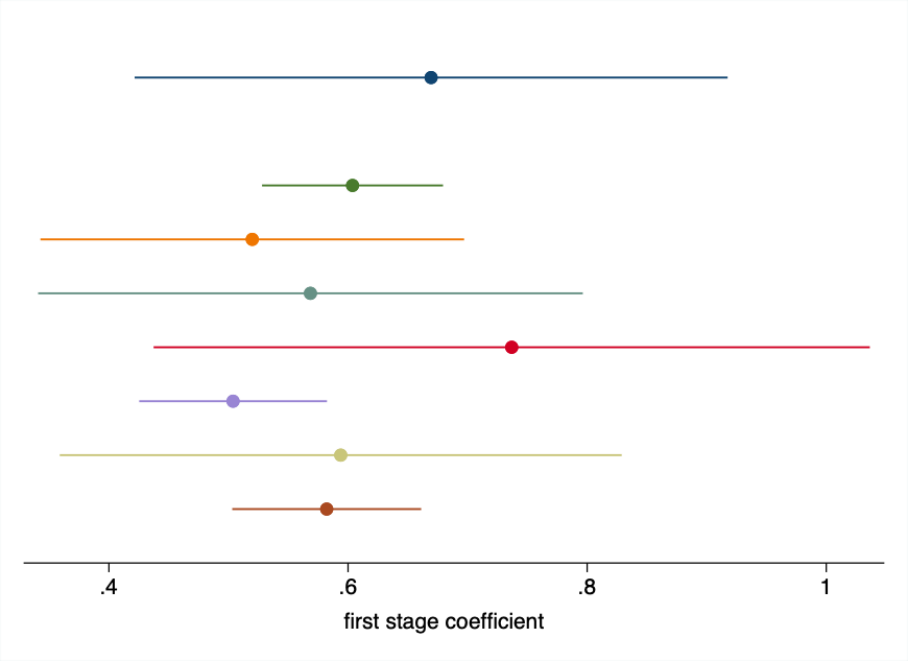
Source: Data from ICPSR and IPUMS. Calculations are author's own.

Figure A3: Leave-one-Out Cross-Validation

(a)



(b)



Note: The figure plots the first stage estimate and 95% confidence intervals. Each regression includes an interaction with the 1920 share of blacks going to different Northern states. I have excluded any Northern state that received fewer than 50 migrants.

B.3 Robustness Check: Re-estimating Regression Results Using WWII Instrument

Table A4: Effect of Black Migration on Tenancy (WW2 pull-factor instrument)

Change in:	Proportion tenancy		Farm Consolidation		Mechanization	
	(1)	(2)	(3) farms	(4) acres/farm	(5) machinery	(6) tractors
A. OLS						
ΔBl_τ	0.761***	0.296***	1.59***	-2.006***	-0.629**	-2.050**
	(0.0948)	(0.0850)	(2.33)	(0.234)	(0.301)	(-3.07)
B. 2SLS (WW2)						
ΔBl_τ	13.35***	18.84	-0.0718	-27.17***	-29.02	-0.00497
	(1.42)	(28.13)	(-0.42)	(8.146)	(51.39)	(-0.29)
C. First Stage						
ΔBl_τ	0.141***	0.0019	0.0674***	0.0674***	0.0135*	-0.00497
	(0.0163)	(0.0325)	(0.0240)	(0.0240)	(0.0039)	(-0.29)
State FEs	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
N	3,309	1,103	3,309	3,309	1,097	3,309

Robust standard errors in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

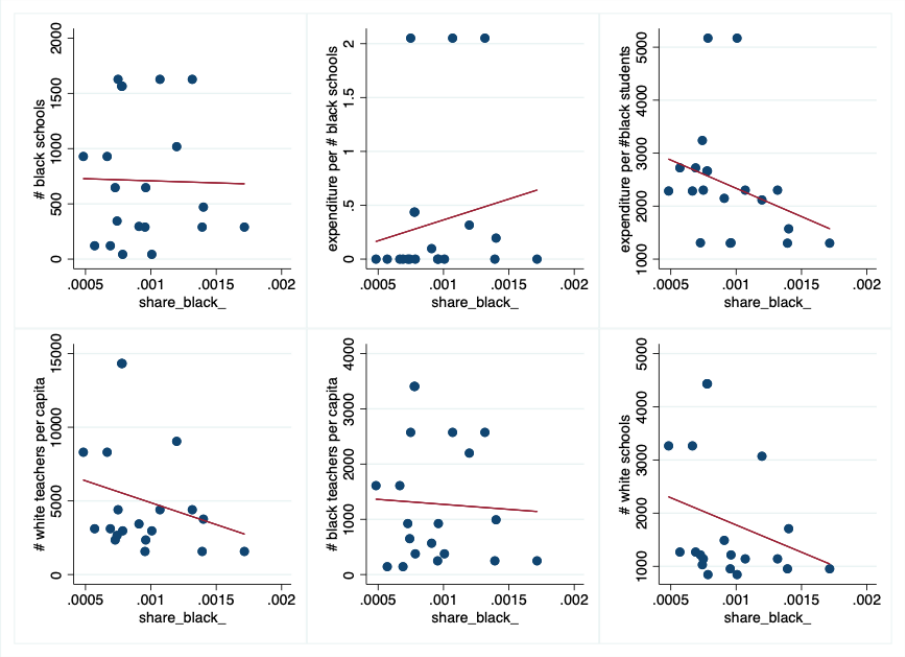
Table A5: Effect of Black Migration on Manufacturing (WW2 pull-factor instrument)

Change in:	(1) Mfg. Val. Added	(2) Share in Mfg.	(3) Real Mfg. Wages
Panel A. OLS			
ΔBl_τ	-2796.736***	-0.0481**	71.710
	(735.53)	(0.0268)	(2.33)
Panel B. 2SLS (WW2)			
ΔBl_τ	-1134.21	-2.619*	5233.96
	(6788.41)	(1.56)	(5518.1)
Panel C. First Stage			
ΔBl_τ	0.137***	0.0594*	0.0718**
	(0.0254)	(0.0325)	(0.0369)
State FEs	Yes	Yes	Yes
Controls	Yes	Yes	Yes
N	1,494	2,774	2,306

Robust standard errors clustered at the county level in parentheses. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

C Appendix – Black Out-Migration and Occupational Upgrading

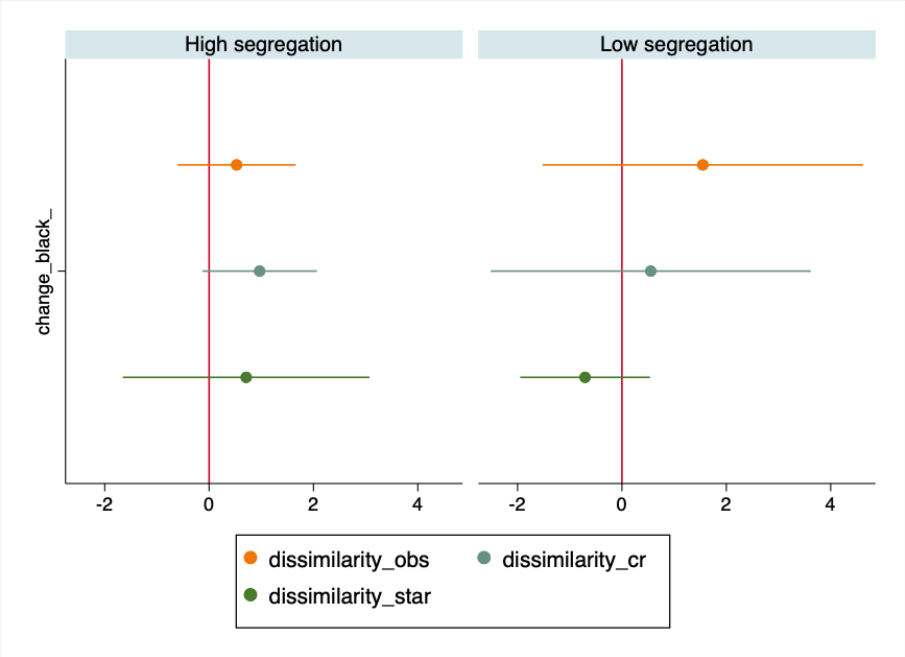
Figure A4: School Quality



Source: Data from IPUMS and Carruthers and Wanamaker (2019). Calculations are author's own.

D Appendix – Black Out-Migration and Segregation

Figure A5: Hetereogeneity: Effect on High and Low Segregation Counties



Source: Data from ICPSR and IPUMS. Calculations are author's own.

References

[1] Abramitzky, R., L.P. Boustan, and K. Eriksson (2012). "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review 102(5)*, 1832-1856.

[2] Abramitzky, R., L.P. Boustan, and K. Eriksson (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration" *Journal of Political Economy 122(3)*, 467-503.

[3] Abramitzky, R., R. Mill, and S. Perez (2018). "Linking Individuals Across Historical Sources: a Fully Automated Approach" Stanford Center for Global Poverty and Development Working Paper No. 1031.

[4] Acemoglu, D. (2010). "When does labor scarcity encourage innovation?" *Journal of Political Economy 118(6)*, 1037–1078.

[5] Acemoglu, D., and D. Autor (2012). "What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology". *Journal of Economic Literature 50(2)*, 426-463.

[6] Adão, R., M. Kolesár, and E. Morales (2019). "What Does Human Capital Do? A Review of Goldin and Katz's The Race between Education and Technology". *The Quarterly Journal of Economics 134(4)*, 1949–2010.

[7] Andersson, D., Karadja, M., and E. Prawitz. (2020). "Mass Migration and Technological Change". Working Paper.

[8] Altonji, J.G., and D. Card. "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives". In *Immigration, Trade, and the Labor Market*, 201 - 234. Chicago: University of Chicago Press.

[9] Becker, G. S. (1957). *The Economics of Discrimination*. Chicago: The University of Chicago Press.

[10] Bohren, J.A., K. Haggag, A. Imas, and D. G. Pope (2019). "Inaccurate Statistical Discrimination". NBER Working Paper 25935.

[11] Borusyak, K., P. Hull, and X. Jaravel (2018). "Quasi-Experimental Shift-share Research Designs". Working Paper.

[12] Boustan, L. P. (2010). "Is Post-War Suburbanization "White Flight"? Evidence from the Black Migration". *The Quarterly Journal of Economics 125*, 417–443.

[13] Boustan, L. P. (2016). *Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets.* Princeton University Press.

[14] Boustan, L. P., M. E. Kahn, and P. W. Rhode (2012). "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century." *American Economic Review 102(3)*, 238-244.

[15] Boustan, L. P., and M. Tabellini (2019). "Black Out-Migration and Southern Political Realignment." Mimeo.

[16] Brannen, R. (1924). "Relation of land tenure to plantation organization." Series: Department bulletin (United States. Dept. of Agriculture) ; no. 1269.

[17] Calderon, A., V. Fouka, and M. Tabellini (2019). "Legislators' Response to Changes in the Electorate: the Great Migration and Civil Rights." Working Paper.

[18] Card, D. (2001). "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics 19(1)*, 22-64.

[19] Carruthers, C. and M. Wanamaker (2019). "County-level school enrollment and resources in ten segregated Southern states." Ann Arbor: Inter-university Consortium for Political and Social Research [distributor]. doi:10.3886/E109625V1

[20] Clemens, M.A., E.G. Lewis, and H.M. Postel (2018). "Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion". *American Economic Review 108(6)*, 108(6), 1468-87.

[21] Collins, W.J. (1997). "When the Tide Turned: Immigration and the Delay of the Great Black Migration". *The Journal of Economic History 57(3)*, 607-632.

[22] Conrad, D.E. (1965). *The Forgotten Farmers: The Story of Sharecroppers in the New Deal*. Urbana: University of Illinois Press.

[23] Derenoncourt, E. (2019). "Can You Move to Opportunity? Evidence from the Great Migration". Draft.

[24] Duncan, B., and O.C. Duncan (1955). "A Methodological Analysis of Segregation Indexes" *American Sociological Review 20(2)*, 210-217.

[25] Dustmann, C. (2011). "Return migration, human capital accumulation and the brain drain." *Journal of Development Economics 95(1)*, 58-67.

[26] Edlund, L., C. Machado, and M. Sviatschi (2015). "Bright Minds, Big Rent: Gentrification and the Rising Returns to Skill" IZA Working Paper 9502.

[27] Fishback, P.V, W.C. Horrace, and S. Kantor. (2005). "Did New Deal Grant Programs Stimulate Local Economies? A Study of Federal Grants and Retail Sales during the Great Depression." *Journal of Economic History 65(1)*, 36–71.

[28] Fouka, V., S. Mazumder, and M. Tabellini (2018). "From Immigrants to Americans: Race, Status, and Assimilation During the Great Migration". Draft.

[29] Gardner, B.L. (2002). *American Agriculture in the Twentieth Century: How it Flourished and What It Cost* Cambridge: Harvard University Press.

[30] Gardner, J., and W. Cohen. (1992). "Demographic Characteristics of the Population of the United States, 1930-1950: County-Level." ICPSR00020-v1. Ann Arbor: Inter-university Consortium for Political and Social Research [distributor], 1992. doi:10.3886/ICPSR00020.v1.

[31] Goldin, C., and L.F. Katz (2010). *The Race between Education and Technology*. Cambridge: Harvard University Press.

[32] Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018). "Bartik Instruments: What, When, Why, and How". NBER Working Paper 24408.

[33] Gregory, J. N. (2005). *The Southern Diaspora: How the Great Migrations of Black and White Southerners Transformed America.* Chapel Hill: University of North Carolina Press.

[34] Haines, M. R. (2010). "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002". ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. doi:10.3886/ICPSR02896.v3.

[35] Heckman, J.J., H.J. Lockner, and P.E. Todd (2003). "Fifty Years Of Mincer Earnings Regressions". NBER Working Paper 9732.

[36] Hellerstein, J. K. and D. Neumark. (2008). "Workplace Segregation in the United States: Race, Ethnicity, and Skill." *The Review of Economics and Statistics 90(3)*, 459–477.

[37] Higgs, R. (2009). *Depression, War, and Cold War: Challenging the Myths of Conflict and Prosperity.* Oakland: Independent Institute.

[38] Hornbeck, R. and S. Naidu (2014). "When the Levee Breaks: Black Migration and Economic Development in the American South". *American Economic Review 104(3)*, 963-990.

[39] Jaeger, D. A., J. Ruist, and J. Stuhler (2018). "Shift-Share Instruments and the Impact of Immigration." NBER Working Paper 24285.

[40] Jaworski, T. (2017). "World War II and the Industrialization of the American South". *The Journal of Economic History 77(4)*, 1048-1082.

[41] Jung, Y. (2019). "Slavery and Evolution of the Local Economy: Labor Market Institutions and Return to Human Capital." Working Paper.

[42] Jung, Y. (2020). "The Long Reach of Cotton in the U.S. South: Tenant Farming, Mechanization and Low-Skill Manufacturing." *Journal of Development Economics 143, Forthcoming.*

[43] Katz, L. F. and R. A. Margo (2014). "Technical Change and the Relative Demand for Skilled Labor: The United States in Historical Perspective." In *Human Capital in History: The American Record*, 15-57. Chicago: University of Chicago Press.

[44] Kniesner, T., A. Padilla, and S. Polachek (1980). "The Rate of Return to Schooling and the Business Cycle: Additional Estimates". *Journal of Human Resources 15(2)*, 273-77.

[45] Koistinen, P.A.C. (1998). *Planning War, Pursuing Peace: The Political Economy of American Warfare, 1920–1939.* Lawrence: University of Kansas Press.

[46] Kuznets, S., A.R. Miller, and R.A Easterlin (1957). *Population Redistribution and Economic Growth, United States, 1870-1950: Vol. II: Analyses of Economic Growth.* Philadelphia: The American Philosophical Society.

[47] Lemieux, T. (2006). "The "Mincer Equation" Thirty Years After Schooling, Experience, and Earnings". In *Jacob Mincer: A Pioneer of Modern Labor Economics*, 127-145. Boston: Springer.

[48] Mandle, J. (1992). *Not Slave, Not Free: The African American Economic Experience Since the Civil War.* Durham: Duke University Press Books.

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

[50] McMillen, N.R. (ed) (1997). *Remaking Dixie: The Impact of World War II on the American South.* Jackson: University Press of Mississippi.

[51] Mincer, J. and S. Polachek (1978). "An Exchange: The Theory of Human Capital and the Earnings of Women: Women's Earnings Reexamined". *Journal of Human Resources 13(1)*, 118-134.

[52] Neal, D.A., and W. R. Johnson (1996). "The Role of Pre-market Factors in Black-White Wage Differences". *Journal of Political Economy 104(5)*, 869-895.

[53] Olmstead, A.L., and P.W. Rhode (2000). "The Diffusion of the Tractor in American Agriculture: 1910-60". NBER Working Paper 7947.

[54] Polachek, S.W. (2007). "Earnings Over the Lifecycle: The Mincer Earnings Function and Its Applications". IZA Working Paper 3181.

[55] Raper, A. (1946). "The Role of Agricultural Technology in Southern Social Change." *Social Forces 25(1)*, 21–30.

[56] Reich, S.A. (2014). *The Great Black Migration: A Historical Encyclopedia of the American Mosaic.* Santa Barbara: Greenwood Publishing Group.

[57] Ruggles, S., M. Sobek, T. Alexander, C. A. Fitch, R. Goeken, P. Kelly Hall, M. King, and C. Ronnander (2010). Integrated Public Use Microdata Series: Version 4.0 [Machine-readable database]. Minneapolis, MN: The University of Minnesota.

[58] Schulman, B.J. (1991). *From Cotton Belt to Sunbelt: Federal Policy, Economic Development, and the Transformation of the South, 1938–1980.* Oxford: Oxford University Press.

[59] Steckel, R.J. and W.J. White (2012). "Engines of Growth: Farm Tractors and Twentieth-Century U.S. Economic Welfare". NBER Working Paper 17879.

[60] Sundstrom, W.A. (2013). "African Americans Since Emancipation." In *Routledge Handbook of Modern Economic History*. London: Routledge.

[61] Tabellini, M. (2020). "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration". *Review of Economic Studies 87(1)*, 454–486.

[62] Tinbergen, J. (1974) "Substitution of Graduate by Other Labour." *Kyklos 27(2)*, 217-226.

[63] U.S. Civilian Production Administration. *War Industrial Facilities Authorized, July 1940 to August 1945.* Washington, DC: GPO, 1946.

[64] Whatley, W. C. (1983). "Labor for the Picking: the New Deal in the South." *The Journal of Economic History 43(3)*, Cambridge: Cambridge University Press.

[65] Winkler, R., Johnson, K. M., Cheng, C., Beaudoin, J., Voss, P. R., and Curtis, K. J. (2013). "County-Specific Net Migration by Five-Year Age Groups, Hispanic Origin, Race and Sex: 2000- 2010". ICPSR34638-v1. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor]. doi:10.3886/ICPSR34638.v1

[66] Wright, G. (1986). *Old South, New South: Revolutions in the Southern Economy since the Civil War.* New York: Basic Books.