The Great Migration and Those Left Behind

Gabrielle Grafton and Michael Neubauer*

March 20, 2026

Click here for most recent version.

Abstract

Despite the extensive literature on the Great Migration, very little is known about its economic and social consequences for Southern non-migrants. We study the impact out-migration imposes on those who remain behind, focusing on Black Americans who remained in the South during the first decade of the Great Migration. We build a shift-share instrument for the net Black out-migration rate using shocks to manufacturing employment in Northern cities. We find that both Black and White non-migrants experienced increases in wages due to out-migration. Black out-migration also led to a decrease in the number of farms as well as less racial hostility in Southern communities. We identify labor supply shocks, occupation switching, and changes in the production process as the mechanisms driving our results. At least in the short term, out-migration can be beneficial for non-migrants.

1 Introduction

During the Great Migration (1940-1970), more than 4 million Black Americans, over 40% of the Southern Black population in 1940, migrated from the South to the North and West. While the Great Migration has been studied extensively, most studies on the Great Migration focus on the effects on migrants themselves, or the places to which they moved (Collins 2021). The migration led to both economic gains and losses for Black migrants, and changed the political and social landscape of the places to which they migrated (Boustan 2010; Collins and Wanamaker 2014; Derenoncourt 2022; Calderon, Fouka, and Tabellini 2023).

However, very little is known about the consequences of the Great Migration on Southern non-migrants. The South during the mid-20th century was underdeveloped, with the agricultural labor market still relying heavily on Black labor for production (Holley 2000). Losses in the labor force, by way of Black out-migration, could have decreased non-migrant wages through sorting and agglomeration effects. Alternatively, Southern employers may have

*Both of Brown University. Neubauer also of Reed College. Neubauer is grateful to the Population Studies and Training Center at Brown University, which receives funding from the NIH, for training support (T32 HD007338) and for general support (P2C HD041020). The authors declare that they have no relevant or material financial interests that relate to the research described in this paper.

increased wages for non-migrants to attract labor to fill vacancies left by Black migrants. Additionally, in this time of Jim Crow laws and heightened racial tensions, the economic ramifications of the Black out-migration could have affected how White Southerners treated the Black residents who remained behind. In this context, the effect of Black out-migration on non-migrants is unclear, as multiple mechanisms push in opposite directions.

In this paper, we investigate how the Great Migration affected the socioeconomic outcomes of those who did not migrate. Motivated by racial differences in agricultural work, we begin by adapting Dustmann, Schönberg, and Stuhler (2017)'s model of heterogeneous labor supply shocks to the agriculturally intensive Southern labor market during the Great Migration. Assuming Black and White workers may be imperfect substitutes due to workplace segregation and discrimination, and that employers cannot alter their input decisions in the short run, our model predicts that a negative shock to the Black labor supply will increase wages of Black non-migrants more than White non-migrants. As a result, the number of firms, which we measure through the number of operational farms, will decline as wages increase variable costs. The model also predicts that Black non-migrant wages will increase overall, under fairly plausible sizes of the labor share and agglomeration parameter. We test these predictions in our empirical analysis.

To empirically estimate the effects of the Great Migration on those left behind, we link data on Southern Black net out-migration rates (1940 to 1950) to data on Southern Black and White non-migrants. For our sample of non-migrants, which are individuals who lived in the same Southern county in 1940 and 1950, we link 445,410 Black and 1,629,999 White individuals with positive wage income in 1940 to their 1950 census records using census record links from the Multigenerational Longitudinal Panel (MLP) project. Historical census records include extensive demographic, geographic, and employment information, including wage income and detailed occupation categories.

We use this data to estimate the effect of 1940-1950 Black out-migration on the change between 1940 and 1950 in non-migrant income, county farming operations, and racial attitudes. Given our model predictions, we expect the Great Migration, a negative shock to the Black labor supply in the South, to increase Black non-migrants wages and decrease the number of farms. However, the relationship between Black out-migration and racial attitudes towards Black non-migrants is theoretically ambiguous. White Southerners may become more repressive as economic conditions decline (R. W. Anderson, Johnson, and Koyama 2017; D. M. Anderson, Crost, and Rees 2020), or they may become less hostile towards Black residents as Black out-migration worsens the labor shortage in the South (Feigenbaum, Soumyajit Mazumder, and Smith 2020). To test whether Black out-migration during the Great Migration impacts racial animosity, we use data on the county-level vote share for George Wallace in the 1968 election and the number of KKK chapters in the 1960s.

Since the economic conditions that push people to leave a given Southern county may also directly affect the wages and occupational choices of those who did not migrate, we instrument for Black out-migration with a shift-share instrument. Our instrument is similar in spirit to Boustan (2010) and Derenoncourt (2022), but it predicts migrants leaving the South instead of migrants entering the North. Our shift-share instrument uses two forms of variation, from the "shares" and "shocks" (shifts), respectively. Our shares represent migration connections between Southern counties and Northern cities established before 1940. For our shocks, we use predicted growth in non-Black manufacturing employment in

Northern cities during the 1940s. We predict this employment growth using a set of LASSO-selected 1940 predictors. By combining our shares and shocks, we form our shift-share instrument for Black out-migration.

Our identifying assumption is that while shocks to manufacturing in Northern destination cities during the 1940s encouraged Black migration, they are uncorrelated with the economic conditions in Southern sending counties. For example, if one were concerned that Black families chose to move out of Southern counties that were already experiencing economic decline, our instrument addresses this issue so long as shocks to Northern non-Black manufacturing employment are not related to pre-existing trends in Southern counties. We check the validity of our exclusion restriction by testing for pre-trends in 1930s Black out-migration and changes in outcomes between 1930 and 1940. We find that our instrument is not correlated with 1930-1940 Black net out-migration rates from Southern counties, and we do not find evidence that our results will be systematically biased based on our estimates using outcomes prior to 1940.

Using our pull-factor shift-share instrument, we find that Black out-migration during the 1940s led to increases in wage income of non-migrants. This increase was larger for Black non-migrants than White non-migrants. Specifically, we find that among Black non-migrants one additional Black departure per 100 Black residents results in a 4% increase in wages. Putting our result into context, the estimated increase in wages for non-migrants in a county with the average amount of Black out-migration is comparable, though smaller, than that reported for Black migrants who moved north. For White non-migrants, we find that one additional Black departure per 100 Black residents results in a 2% increase in wages. We find that lower-income individuals, as well as Black and White non-migrants employed as farm employees and service workers in 1940, received the largest gains in wage income. Men and women of both races experienced gains in wage income.

We also find that Black out-migration led to a decline in the number of farmers and farms, along with less racial animosity. Places with high Black out-migration experienced increases in average wage income, farm value, and farm size, but the number of farmers and farms declined by 1.8% and 0.88%, respectively. The increase in farm value and farm size, together with the reduction in farms, suggests that Black out-migration may have caused smaller farms to shut down or consolidate. With respect to racial attitudes, we find that Black out-migration during the 1940s had a negative effect on the number of KKK chapters in the 1960s. Trying to prevent the loss of more Black labor, White Southerners may have become more tolerant of the Black Southerners who remained behind.

Since the sorting and agglomeration effects of positively-selected out-migration would decrease wages of non-migrants, our finding that out-migration increased wages are driven by occupation switching, changes in the production process, and the negative shock to the labor supply. Taking into account declines in average education due to sorting and reduced agglomeration effects due to population declines, our estimates for the effect of out-migration on wage income would be slightly larger in the absence of these effects. However, we find that Black out-migration pushed Black and White non-migrants into higher-status occupations as well as occupations that saw the largest wage gains from the negative labor supply shock. Additionally, we find that places with high out-migration saw decreases in the share of income accruing to labor and increases in the share of income accruing to capital. Since the marginal product of labor increases with higher capital investment, an increase in capital investment

may explain increases in wages for non-migrant workers who remained employed.

In our robustness checks, we address concerns regarding selection, measurement of farmers' wage income, and effect linearity. If, as we find evidence for in our sample, migration is positively selected, then non-migrants in high-migration counties are drawn from a lower-ability distribution than non-migrants in low-migration counties, biasing our estimates of the effect of migration on non-migrant wages downward. Similarly to Mishra (2007), we show that our results are robust to restricting our sample to a population for which the estimates are likely to be less biased: individuals with a low predicted likelihood of migrating. Furthermore, since our non-migrant sample relies on census linking, our results may be biased if there is non-random selection into the likelihood of being linked across the 1940 and 1950 censuses. We find that our results are robust to re-weighting our sample by the inverse probability of being linked in the census. Additionally, since the 1940 census did not report income for self-employed workers, we show that our main results are robust to excluding the 1940 farmers who have positive reported wage income. Finally, while in our main specification we assume that the effects of Black out-migration on wages are linear, it is possible that only places experiencing either low or high levels of out-migration drive our wage income results. After estimating the effect of Black out-migration on non-migrant wage income separately for predicted high- and low-migration counties, we find limited evidence of nonlinear effects.

This paper adds to the extensive literature studying the Great Migration by examining the effect of the migration on those left behind. The existing literature mainly focuses on the experiences of the migrants and the Northern cities that received a large influx of Black migrants (Collins 2021). With regard to wages, Boustan (2016) finds that the wages of migrants are higher than non-migrants, though Akbar et al. (2022) notes that much of those increased wages may have been spent on soaring rental costs for housing in segregated areas of Northern cities. Derenoncourt (2022) examines inter-generational effects and finds that the Great Migration had a negative impact on upward mobility for Black families in the North. Relative to this literature, our contribution is to focus on Black residents of Southern communities who themselves did not migrate. We find that the migration was beneficial for them in the short-run, at least in terms of income and racial attitudes.

We also contribute to the literature studying the effects of migration on sending communities and non-migrants more generally by showing that domestic migration, at least in the United States, can benefit those who do not migrate. Jia et al. (2023)'s review of the literature on internal migration within the United States highlights the lack of literature on non-migrants. One exception is Deryugina, Kawano, and Levitt (2018), which found increases in long-run income for those who remained in New Orleans after Hurricane Katrina, likely due to reductions in labor supply from widespread out-migration.

In contrast with the relatively limited literature on domestic non-migrants, numerous studies in the international context have found that migration out of a community increases human capital accumulation, income, and entrepreneurship in that community due to remittances and increased opportunities for those who remained behind (Yang 2008; Docquier and Rapoport 2012; Dinkelman and Mariotti 2016; Shrestha 2017; Khanna et al. 2022). Similarly, Mishra (2007), who studies the effect of emigration on the wages of Mexican workers who do not migrate, finds that immigration to the U.S. has positive effects on Mexican wages. Further, Akram, Chowdhury, and Mobarak (2017) studies seasonal migration in Bangladesh

and finds that out-migration increases male agricultural wages in sending villages. Similar to these findings, we find that migration can have positive effects on non-migrants and sending communities, though our focus is on domestic migration within the United States where remittances and short-term migration are not as common.

Lastly, our findings are relevant to the literature studying policies to help left-behind populations by demonstrating that, at least in the context of the Great Migration, out-migration can have positive impacts on those who do not migrate. Motivated by contemporary trends in declining cities and regions of the U.S., this literature has largely focused on examining "place-based" policies. Busso, Gregory, and Kline (2013) examine areas designated by the federal government as Empowerment Zones and find large increases in local employment and wages. Large-scale infrastructure programs, such as those undertaken by the Tennessee Valley Authority, can benefit local populations (Kline and Moretti 2014). Austin, Glaeser, and Summers (2018) argue that subsidizing employment and investing in education may be effective tools to help those living in areas with chronically high rates of non-employment. In contrast with these policy prescriptions, migration is sometimes seen as helping those who migrate but hurting those who remain behind (Economist 2017). For example, Diamond (2016) shows how sorting, induced by migration from changes in local labor demand, can increase inequality, specifically hurting low-skilled workers who live in expensive cities with amenities that high-skilled workers value. However, our findings run contrary to this view of migration, since we show that out-migration can be beneficial to non-migrants in the short term, especially when those who leave and those who remain were previously competing in the same sector of the labor market.

The remainder of the paper is organized as follows. Section II gives a brief overview of the labor market and racial attitudes in the South during the 1940s. Section III presents our conceptual framework. Section IV describes our data sources and sample construction, and provides descriptive statistics on this sample. In Section V, we present our empirical strategy for addressing endogeneity concerns and identifying the causal effect of 1940s Black out-migration. Section VI reports our empirical results. In Section VII, we discuss potential mechanisms driving our results. Section VIII tests the robustness of our results. Lastly, Section IX concludes.

2 Background

Out-migration could have affected life for non-migrants in a myriad of complex ways. In this paper, we focus on two aspects of life that are arguably among the most salient: the ability to earn wage income from employment, and relations between races. In this section, we provide some background on these two issues.

2.1 1940s Southern Labor Market

Agriculture had the most workers of any industry in the South during the 1940s, employing over 4 million Southerners in 1940.[1] Of employed Southern men and women in 1940, 32%

[1]Data comes from the 1940 full count census (Ruggles, Fitch, et al. 2024, Ruggles, Flood, et al. 2025). Over 4.3 million employed people living in the South in 1940 were employed in the agriculture industry,

were employed in agriculture, with 51% employed in the rural South.[2] Within Southern agriculture, Black workers played an essential role in the supply of farm wage labor. While White workers tended to be in higher-up positions on farms, such as farmer owners and tenant farmers, Black agricultural workers made up almost half of all farm wage labor in 1940, one of the lowest positions in the hierarchy (Holley 2000).

Despite technological advancements in agriculture production during the mid 20th century, much of Southern agricultural operations were still labor intensive. The labor-intensive nature of Southern agriculture stems from the reliance on Black slave labor for agricultural production before the Civil War. Even in the decades following Black emancipation, White farmers continued to rely on Black labor in the form of sharecropping and wage farm labor (Fite 1979; Holley 2000). For White farm owners, Black labor, which was often low-cost and plentiful, proved to be less expensive than investing in technological advancement, and thus discouraged full mechanization. By the early 1940s, cotton production still required 99 man-hours per acre on average (Holley 2000).

During the 1940s, the South experienced major labor shortages due to mass Black outmigration. Many Southerners, especially Black Southerners, moved north as increased economic opportunities opened up from WWII production and halted European immigration (Carrington, Detragiache, and Vishwanath 1996; Foote, Whatley, and Wright 2003; Collins and Wanamaker 2015). Studies have also pointed to the role of environmental and natural disasters, such as floods and pest infestations, in driving workers out of the South (Lange, Olmstead, and Rhode 2009; Boustan, Kahn, and Rhode 2012; Hornbeck and Naidu 2014; Ager, Brueckner, and Herz 2017; Feigenbaum, Soumyajit Mazumder, and Smith 2020). Racial violence, such as lynchings, also pushed Black Southerners to leave the South during the Great Migration (Tolnay and Beck 1992; Christian 2017). With many Black Southerners leaving the region, farmers lost a large share of their farm labor, with some lamenting that the labor supply in 1944 was up to 25% below that of the previous year (Holley 2000).

Therefore, while the agriculture industry employed the most Southern workers during the 1940s, it was also the industry most impacted by the Great Migration. The percentage of Southerners employed in the agriculture industry shrank from 32% to 21% by 1950. Among those employed in Southern agriculture, the share of those employed as wage farm laborers dropped from 22% in 1940 to 18% in 1950. We do not find the same trend in other major industries in the South. Manufacturing, the industry employing the second-highest number of workers in the South, increased its employment share from 15% in 1940 to 17% in 1950, while the wholesale and retail trade industry increased its share of Southern employment from 12% in 1940 to 15% by 1950 (Appendix Table A.1).

Despite these large-scale changes occurring due to the Great Migration, how the migration affected those who did not migrate has not been extensively studied. Farmers who relied heavily on farm labor in their production processes could not quickly implement mech-

which was 32% of all employed Southerners in 1940. Manufacturing was the second largest industry with 15% of Southerners working in this industry in 1940. Wholesale and Retail Trade followed closely behind with 12% of Southerners working in this industry in 1940 (Table A.1).

[2]Of those employed in the agricultural industry in 1940, 58% were farm owners or tenant farmers and 22% were wage farm laborers. In Appendix Table A.1, we report shares in agricultural employment in the South in 1940 and 1950. All shares are calculated using the employed Southern population in 1940 and 1950, including both men and women.

anization, a costly investment. Struggling to shift from labor-intensive to capital-intensive production, farmers may have increased wages to attract workers and ameliorate labor shortages. Additionally, depending on how substitutable White unskilled labor was for Black labor, White laborers may have also benefited from the decrease in the Black labor supply. We examine these conjectures both theoretically and empirically in the later sections of this paper.

2.2 Racial Attitudes in Southern Communities

In addition to affecting the labor market, a migration-induced shock to the Southern Black labor supply could also impact race relations. While changing racial attitudes in the post-migration South could have a substantial impact on the daily lives of Southern Black residents, the relationship between out-migration and racial attitudes is theoretically ambiguous. Empirical examinations of the effects of economic conditions on attitudes towards minorities have found mixed results (R. W. Anderson, Johnson, and Koyama 2017).

On the one hand, some studies have found increasingly hostile attitudes towards minorities as economic conditions worsen (R. W. Anderson, Johnson, and Koyama 2017; D. M. Anderson, Crost, and Rees 2020). These studies span both historical and contemporary time periods, in Europe and the United States. While increased competition for economic resources may strain relations between individuals in different racial or ethnic groups, the mechanisms linking economic stress to conflict are not fully understood (Burke, Hsiang, and Miguel 2015).

On the other hand, research focusing on migration of racial minorities has documented declining hostility in response to out-migration. Studying racial relations earlier in the 20th century, Feigenbaum, Soumyajit Mazumder, and Smith (2020) found that Southern society became less repressive and less violent in response to Black out-migration that was caused by boll weevil infestations. While out-migration may alter existing economic patterns and cause economic stress, the desire to prevent further out-migration may create incentives for the majority group to treat minority groups better.

Given the conflicting theoretical and empirical predictions of how out-migration will affect racial attitudes, we will examine this question empirically in the context of the Great Migration. Specifically, we determine how racial attitudes in Southern counties during the 1960s, measured using votes for George Wallace and the number of KKK chapters, varied based on the amount of Black out-migration those counties experienced in the 1940s.

3 A Model of Race-Specific Labor Supply Shocks

As described above, the South during the 1940s provides a unique context in which one racial group, the one that predominantly supplies unskilled labor, experiences a large negative shock to labor supply. This setting demands a model that links race-specific labor supply shocks to the economic well-being of all those in the community affected by the shock.

We provide such a model by proposing a theoretical framework that explains why race-specific labor supply shocks may affect non-migrant wages. This framework captures the distinctive landscape of the Southern labor market during the early 20th century. We find

that, because workers of one race may be imperfect substitutes for workers of another race, a large race-specific labor supply shock induced by out-migration increases the wages of non-migrants in the racial group causing the shock more than the wages of non-migrants in groups not causing the shock. Additionally, this labor supply shock and resultant increase in non-migrant wages raises variable costs for farms, which may result in farms for which variable costs now exceed total revenue to shut down. While this model assumes workers are homogenous (except for race), a model in which workers could change occupations would provide another avenue for non-migrant wages to increase.

3.1 Model Set-up

We adapt and apply the model of Dustmann, Schönberg, and Stuhler (2017) that links immigration-induced labor supply shocks to native wages and employment decisions to the Southern labor market during the Great Migration. Specifically, we establish how race-specific migration-induced labor supply shocks are theoretically linked to non-migrant wages and the number of operational farms. We allow Black and White laborers to be imperfect substitutes in the Southern labor market at this time, and derive the effect of Black out-migration on the wages of both Black and White non-migrants. Importantly, we focus on short-run effects, for which we assume that employers cannot alter their input decisions, such as their level of capital utilization. These model predictions have testable implications, and we pair our model predictions with empirical analysis of the effects of out-migration on non-migrant wages and the number of operational farms later in the paper.

3.1.1 Production

We assume that agricultural output Q in a given county is produced by combining labor and capital in a Cobb-Douglas production function with agglomeration effect σ, where $E = L^{\sigma}$:

$$Q = AK^{\alpha}(LE)^{1-\alpha} \tag{1}$$

Further, we assume that labor L used to produce this agricultural output is a CES aggregate of Black (B) and White (W) labor. We treat labor of Black and White workers separately, similarly to Hornbeck and Naidu (2014), since laborers of different racial groups likely could not perfectly substitute for each due to workplace segregation and discrimination:

$$L = [\theta_B L_B^{\beta} + \theta_W L_W^{\beta}]^{\frac{1}{\beta}} \tag{2}$$

where $\theta_W + \theta_B = 1$ and $\beta \leq 1$.

We next show how decreases in the Black labor force affect wages of Black and White workers in this model differentially. We also examine the number of operational farms, as agriculture employed the most workers of any sector at this time.

3.1.2 Linking Race-Specific Labor Supply Shocks to Non-migrant Wages

We model the effects of out-migration on non-migrant wages. Holding all else equal, reduced agglomeration would decrease wages when the Black labor force shrinks, while the labor supply shock would increase wages. Though these effects work in opposite directions, we

show in the model that the net effect on Black wages is positive for usual values of the labor share and agglomeration effect, and under the assumption that the levels of capital K and total factor productivity A do not change due to out-migration. Additionally, we show that the effect on Black non-migrant wages is more positive than the effect on White non-migrant wages, for which the sign of the effect is ambiguous.

Assuming that firms (farms) are price takers, they choose the amount of labor they employ to equate the marginal cost of labor and the marginal product of labor, giving:

$$\log \omega_B = \log[(1-\alpha)(1+\sigma)A] + \alpha[\log K - \log L] + \sigma(1-\alpha)[\log L] + \log \theta_B + (\beta - 1)[\log L_B - \log L]$$
$$\log \omega_W = \log[(1-\alpha)(1-\sigma)A] + \alpha[\log K - \log L] + \sigma(1-\alpha)[\log L] + \log \theta_W + (\beta - 1)[\log L_W - \log L]$$

We take the derivatives of the above expression for wages with respect to the size of the Black labor force, L_B, to determine how out-migration affects wages for Black (ω_B) and White (ω_W) workers holding all else constant:

$$\frac{\partial \log \omega_B}{\partial L_B} = \frac{\Big((1-\alpha)(1+\sigma) - \beta\Big)\theta_B L_B^{\beta-1}}{\theta_B L_B^{\beta} + \theta_W L_W^{\beta}} - \frac{(1-\beta)}{L_B}$$
$$\frac{\partial \log \omega_W}{\partial L_B} = \frac{\Big((1-\alpha)(1+\sigma) - \beta\Big)\theta_B L_B^{\beta-1}}{\theta_B L_B^{\beta} + \theta_W L_W^{\beta}}$$

As we show in Appendix A.3.1, $\frac{\partial \log \omega_B}{\partial L_B} < 0$ for values of α and σ within the usual range, so a decrease in L_B, the size of the Black labor force, increases Black wages ω_B. Additionally:

$$\frac{\partial \log \omega_B}{\partial L_B} - \frac{\partial \log \omega_W}{\partial L_B} = -\frac{(1-\beta)}{L_B} \implies \frac{\partial \log \omega_B}{\partial L_B} < \frac{\partial \log \omega_W}{\partial L_B}$$

Therefore, a decrease in L_B does not increase White wages as much as it increases Black wages. In fact, the sign of the effect on White wages is ambiguous, and depends on the degree of substitutability, β, between White and Black workers. As the degree of substitutability between Black and White labor increases, the effects of out-migration on Black and White wages become more similar, as laborer race matters less when the two groups of workers approach perfect substitutability.

3.1.3 Linking Race-Specific Labor Supply Shocks to Farm Operation

Shifting from wage changes to farming operations, our model describes the effect of a negative labor supply shock in the short run, in which we assume that input decisions are fixed.[3]

Profits for individual farm a are given by:

$$\pi_a = pQ_a - \omega_B L_{B_a} - \omega_W L_{W_a} - rK_a - \epsilon_a \tag{3}$$

[3]In the longer-run, farms would adjust their optimal inputs of labor and capital in response to the supply shock (Hornbeck and Naidu 2014).

where we assume that p is the common price faced by all farms for selling their agricultural output, ω are wages paid, and ϵ_a is the farm's fixed cost of production.

In the short-run, farms shut down if variable costs exceed total revenue:

$$pQ_a < \omega_B L_{B_a} + \omega_W L_{W_a} + rK_a \tag{4}$$

The increase in ω_B due to the negative supply shock increases variable costs in the short run, given that L_{B_a}, L_{W_a}, and rK_a are fixed. Assuming there is heterogeneity in variable costs across farms (for example, due to pre-existing differences in the racial composition of workers, L_{B_a} and L_{W_a}), there may now be some farms for which variable costs exceed total revenues, leading these farms to shut down.

Summarizing the results of our theoretical model, we have shown how out-migration of Black workers can lead to increases in the wages of remaining Black workers, smaller increases in the wages of non-migrating White workers, and a decline in the number of operational farms. Additionally, our expression for log wages shows that wages should increase if the production process becomes more capital-intensive and the capital stock (K) grows. On the other hand, log wages should fall if average productivity (A) decreases because higher-ability individuals leave, or if agglomeration effects (σ) shrink as the population decreases. Below, we describe the data we use to test these predictions empirically.

4 Data and Descriptive Statistics

To estimate the effect of Black out-migration on those who do not migrate, we use data on 1940 county characteristics, including the amount of Black out-migration between 1940 and 1950, farming operations, and racial attitudes, and individual-level data on non-migrant wages and employment in 1940 and 1950. This individual-level data allows us to analyze an understudied population and estimate the causal effect of Black out-migration on individual earnings and granular occupation choices. We find that places that saw more migrants leaving tend to be more rural, have lower-valued, smaller farms, and have higher Black population shares in 1940. Additionally, we use this data to compare our samples of migrants and non-migrants, and we find that migrants are positively selected on 1940 characteristics.

4.1 Data

We use data from the 1940 and 1950 full-count censuses, links between the 1940 and 1950 censuses, and data on 1940s Black out-migration to measure out-migration and identify non-migrants in the South. With this data we can also measure wage income and occupational decisions in 1940 and 1950, which provide the basis for our main outcomes of interest. Additionally, we use county-level data on income, farming, and racial attitudes to understand changes occurring in local communities. Finally, we construct our instrument for Black out-migration using data on shocks to Northern employment in 1940.

To measure Black out-migration, we use the net out-migration rates of the Black population from Southern counties between 1940 and 1950 from Boustan (2016). Boustan (2016) uses the census survival method to estimate the 1940-1950 Black net migration rate, which is the number of net Black out-migrants from a county between 1940 and 1950 per 100

Black residents. Using net out-migration, instead of total out-migration, differences out in-migration into the area. As the model in the previous section shows, what matters is the overall change in the Black labor supply, which is reflected in the net change in the Black population instead of the total number of Black residents that moved out of the area. To improve the power of our instrument, we restrict our analysis to Southern counties that experienced net Black out-migration between 1940 and 1950. During the 1940s, 75% of Southern counties experienced net Black out-migration.[4]

In conjunction with the 1940 and 1950 full-count decennial censuses (Ruggles, Fitch, et al. 2024; Ruggles, Flood, et al. 2025), we use census record links from the Multigenerational Longitudinal Panel (MLP) project (Ruggles, Ozder, et al. 2025) to construct our sample of Black and White non-migrants. The MLP data allows us to track both men and women over time by linking their 1940 and 1950 census records. The full-count census data provides detailed information on demographics and geographic location, including county of residence. For our main analysis, we classify non-migrants as Black and White Southerners who lived in the same Southern county in 1940 and 1950. In our sample of non-migrants, we are able to link 37.6% of Black Southerners and 55% of White Southerners in 1940 to their 1950 census records.[5]

Using our sample of non-migrants, we construct our individual-level outcomes, the change in wage income, occupational income score (OCCSCORE), and likelihood of being in a specific occupation or employment status between 1940 and 1950.[6] For income, we use the census variable for wage and salary income (INCWAGE), which measures all earnings received as an employee over the previous year in dollars.[7] OCCSCORE corresponds to the median total income (in hundreds of 1950 dollars) of all individuals for a given occupation (Ruggles, Fitch, et al. 2024; Ruggles, Flood, et al. 2025). We use this continuous measure of occupations as a proxy for occupational income and prestige. Lastly, we measure the likelihood of being in a specific occupation using the census variable OCC1950, which categorizes individuals' occupations using the 1950 classifications. In our investigation, we focus on the following occupations: farmers, farm employees, and service workers, which were the most

[4]Among Southern counties with net Black out-migration, variation in the share of Black migrants leaving between 1940 and 1950 is visible in Figure 1, Panel A.

[5]See Appendix Table A.2 for a comparison of the full sample of Black and White Southerners living in the counties in our sample in 1940 to our linked non-migrant samples. In Section 8 we address concerns regarding selection in linking.

[6]The change in occupational income score and wage income between 1940 and 1950 are calculated as the relative (or decimal) change, while the change in occupation between 1940 and 1950 is the difference between indicator variables for the likelihood of being in the occupation in 1940 and 1950. The change in occupation status is also restricted to those who are in the labor force in 1940 and/or 1950. Our main sample of non-migrants is all Black and White Southerners who lived in the same Southern county in 1940 and 1950 that have a non-missing value for at least one of our outcomes. In other words, the sample of non-migrants must have at least one of the following: (1) non-zero wage income in 1940, (2) non-zero occscore in 1940, (3) in the labor force in 1940, or (4) in the labor force in 1950. For our main specification, we focus only on those with non-zero wage income in 1940.

[7]To deal with top-coded values, we follow Goldin and Margo (1992). We multiply the top-coded values for wage income in 1940 and 1950 by 1.4. Additionally, to deal with extremely low annual earnings, we replace wage income values with zero if they are less than half of the minimum weekly wage ($6 in 1940 and $8 in 1950). We also adjust for inflation before we calculate our change in wage income variables using the CPI inflation calculator from Labor Statistics (2025).

common occupation categories among Black Southerners in 1940.[8]

Since the 1940 census does not measure income for the self-employed, we follow Carruthers and Wanamaker (2017) and exclude individuals with missing wage income in either 1940 or 1950 from our main wage income results. Our main outcome for income, the relative change in wage income between 1940 and 1950, already excludes individuals with zero wage income in 1940. Dropping those with missing wage income in 1940 or 1950 excludes an additional 1,427 Black and 7,393 White non-migrants from our main outcome.

Due to errors in the 1950 full-count census personal income variables, we follow Collins and Zimran (2025) and provide estimates using non-migrants with reported income greater than zero in 1950.[9] Ruggles, Flood, et al. (2025) have disclosed problems with misreporting of personal income variables in the 1950 census, estimating that the variable for personal wage income is accurate 70-75% of the time. Wages are more likely to be underestimated for erroneous observations (Ruggles, Flood, et al. 2025). Provided that there are no systematic differences in misreporting by our treatment, county-level Black out-migration between 1940 and 1950, these income errors should not bias our estimates. Following Collins and Zimran (2025), we provide estimates using non-migrants with reported income greater than zero in 1950 by analyzing the change in log wage income between 1940 and 1950. Using the change in log wage income alleviates bias by excluding individuals whose wage income was misreported as zero or missing in the 1950 census.

To understand aggregate changes in the local community following Black out-migration, we use data on county-level changes in income, farming employment and operations, and racial attitudes. For income, we calculate the change in average wage income in a county between 1940 and 1950 using the 1940 and 1950 full count censuses (Ruggles, Fitch, et al. 2024; Ruggles, Flood, et al. 2025). For changes in farmer employment, we use the 1940 and 1950 full count censuses to measure the change in the number of farmers between 1940 and 1950. For information on farming operations, we use data from the 1940 and 1950 county-level decennial censuses and 1947, 1952, and 1956 County Data Books from Haines (2010). Specifically, we calculate county-level changes in the number of farms, average farm value, average farm size, and total farming area between 1940 and 1950.[10] Given the issues with measuring farmers' income using individual-level data, county-level measures of farm operations provides proxies for farmers' economic status and allows us to infer how Black out-migration affects their business. We also use measures for county-level vote share for George Wallace in the 1968 President election from Clubb, Flanigan, and Zingale (2006) and the number of KKK chapters in the 1960s from Shom Mazumder (2018) to assess how changing social and economic conditions affect racial attitudes.

Lastly, to construct our shift-share instrument for Black out-migration, we again use data from 1940 and 1950 censuses, this time with the addition of the 1947 County Data

[8] In our main sample of Black stayers, 18% were farmers or tenant farmers, 27% were farm employees, and 18% were in the service industry (Appendix A.4). We classify service workers as those with the top service jobs in our sample. These jobs include: truck and tractor drivers; operative and kindred workers; private household workers; porters; janitors and sextons; and other laborers.

[9] At the time of writing this paper, only the preliminary release of the 1950 full-count census is available.

[10] Farm value, measured in dollars, is the value of farm property. This includes land and buildings, implements and machinery, and livestock (Census 1947). Average farm value is calculated as the average value of farm property per farm in 1940 dollars.

Book (Haines 2010). We construct the shares of our shift-share instrument, which represent migrant network links between Southern counties and Northern cities, using the record from the 1940 census of the location where an individual lived five years prior. We construct the shifts (or shocks) using the 1940 and 1950 censuses to measure growth in non-Black manufacturing employment in Northern cities. To create predicted shocks, we use 1940 city-level pull factors from the full-count censuses to create all of our variables, with the exception of a proxy for city-level WWII spending per capita, which comes from data on county-level WWII contracts from the 1947 County Data Book (Haines 2010).

4.2 Descriptive Statistics

To better understand the type of places migrants left behind, we explore characteristics of the Southern counties in our sample and discuss how these characteristics relate to Black out-migration during the 1940s. We then describe how income and employment changed between 1940 and 1950 for non-migrants. Lastly, we compare non-migrants' 1940 characteristics with the 1940 characteristics of migrants who had moved away from the South by 1950.

4.2.1 County-level Characteristics

We begin by using the data described above to explore Southern counties that Black Americans left during the 1940s. Below we summarize the amount of Black out-migration and other county characteristics for both high- and low-migration counties.

Our sample counties are characterized by having large numbers of Black out-migrants. In our sample, which is restricted to counties that experienced Black out-migration on net, the average Black net out-migration rate is 23.19 per 100 Black residents (Appendix Table A.3). Changes in incomes of Black Southerners are correlated with county-level net out-migration rates.[11] In order to identify a causal, as opposed to correlational, relationship, our empirical strategy will need to address endogeneity issues caused by selection and shocks that precipitated both out-migration and economic decline in Southern counties.

In Appendix Table A.3, we provide county-level characteristics for the counties in our main sample, and compare counties with high and low net Black out-migration rates between 1940 and 1950. We define a high out-migration county as a county with a net Black out-migration rate at or above the median, while a low out-migration county's net Black out-migration rate is below the median. When comparing counties with low net Black out-migration during the 1940s to those with high amounts of out-migration, we find that places that experienced higher Black out-migration were more rural, had lower average income, had smaller and less valuable farms, and had more agricultural employment. Additionally, places with high levels of Black out-migration during the 1940s had more Black lynchings between 1882 and 1930 and higher Black population shares in 1940. The notable differences between the counties in our sample provides further evidence that we need to address endogeneity issues associated with the relationship between out-migration and economic conditions in

[11] We can see this in Figure 1 and Appendix Figure A.1. When we plot the binscatter of the county-level 1940-1950 net Black out-migration rate and county-level mean in the change in wage income between 1940 and 1950, we find that Black non-migrant wages are positively correlated with Black out-migration.

Southern counties to be able to identify the causal effect of Black out-migration on non-migrant outcomes.

4.2.2 Non-migrant Sample

We next describe our sample of Black and White non-migrants. We use the linked 1940 and 1950 census data to analyze this individual-level data.

For our sample of Black and White non-migrants with positive wage income in 1940, we report descriptive statistics of net out-migration and our individual-level outcomes in Table 1. The change in wage income and occupational income score between 1940 and 1950 are calculated as the relative change. The change in log wage income between 1940 and 1950 is calculated as the difference between log wage income in 1950 and log wage income in 1940. The change in farm employment between 1940 and 1950 is the difference between the indicator variables for farm employment in 1940 and 1950.[12]

For Black non-migrants, real wages decreased by 4% on average during the 1940s. Analyzing log wage income, which requires that an individual have positive income in both 1940 and 1950, reveals an increase in wage income of 94%. Similarly, White non-migrants in our sample saw real wages decrease by 49% on average, while White non-migrants with positive wages in 1940 and 1950 saw their wage income increase by 49%. The declines in real income are driven by those who shift from positive reported income in 1940 to zero reported income in 1950, as well as high inflation during this time period.[13] White non-migrants experienced a larger reduction in real wage income, on average, relative to Black non-migrants, because they were more likely to have zero wage income in 1950 and had higher levels of income in 1940. White non-migrants were more likely to report zero wage income in 1950 because White non-migrants were older and more likely to be self-employed in 1950, indicating that they may be leaving the labor force or no longer receiving income from wages.[14] Additionally, real wage increases during this time period are concentrated among those who had lower levels of income in 1940. Since average White non-migrant income was more than $500 more in 1940 than average Black non-migrant income (Table 1), White non-migrants received less of the wage gains enjoyed by lower income individuals.

4.2.3 Selection of Migrants

Turning to a comparison of migrants and non-migrants in Appendix Tables A.4 and A.5 , we find that Black and White migrants tend to be positively selected on 1940 characteristics.

We compare the average characteristics of Black non-migrants and migrants in Appendix Table A.4. In column 1, we report Black non-migrants' average demographics and occupational choices in 1940, and in column 2 we report averages in the same characteristics for

[12]We calculate the change in employment and occupation categories using the difference between the indicator variables in 1940 and 1950 since the relative change would only be identified for those where the indicator equals one in 1940. For example, if we used the relative change in farm employment, we would only be estimating occupation changes for those who were employed in farming in 1940. Since we want to capture switching into and out of specific occupations, we use the difference instead of the relative change.

[13]Between 1940 and 1950, the consumer price index increased by about 72% (Labor Statistics 2025).

[14]For those who may also have non-wage income, in Section 8 we show that our results are robustness to excluding these workers.

Black migrants.[15] Column 3 reports differences in characteristics between non-migrant and migrants, and column 4 reports the adjusted difference. Similar to Collins and Wanamaker (2015), we define the adjusted difference as the difference remaining after controlling for age, gender, and county-of-residence fixed effects in 1940. We perform the same analysis for White Southerners (Appendix Table A.5).

We find that both Black and White migrants are positively selected on education, wage income, and occupation relative to non-migrants. Both Black and White migrants are over four years younger than non-migrants. For Black Southerners, migrants have over a quarter of a year more education than non-migrants, and this gap widens to over half of a year after controlling for age, gender, and county of origin. White migrants have over a half of a year more education than non-migrants after adjusting for age, gender, and origin location. For income, both Black and White migrants have higher wage income and occscore prior to migrating, compared to similar non-migrants living in the same county in 1940. Additionally, migrants are more likely to be tenant farmers and laborers, while non-migrants are more likely to be farm owners. Black non-migrants are 2 percentage points more likely to be a farmer owner, on average, and White non-migrants are 5 percentage points more likely to be a farmer owner, than migrants are after adjusting for age, gender, and county of residence in 1940.

Our results are consistent with Collins and Wanamaker (2014) and Black et al. (2015), who find evidence consistent with positive selection, in terms of education and earnings potential, of Black migrants during the Great Migration. Consistent with our findings for 1940 after controlling for income and county of residence, Collins and Wanamaker (2014) notes that in 1930 the majority of non-migrants were employed in farming, while many migrants were laborers and fewer were involved in farming.[16] Taken together, our evidence indicates that there was likely negative selection into our sample of non-migrant Black Southern residents.

5 Empirical Strategy

Given our model predictions, we use the data described above to estimate the effect of 1940-1950 Black out-migration on wage income for Black and White non-migrants. Since economic conditions that drive out-migration may also have a direct effect on non-migrants, we instrument for Black out-migration using a shift-share instrument that relies on migrant connections and shocks to employment in Northern cities.

Specifically, to determine the extent to which out-migration affects those who do not migrate, we want to estimate β in the causal model given by equation 5. In this equation, y_i is the outcome of individual i who stays in county j and x_j is the net Black out-migration rate of county j between 1940 and 1950.

$$y_i = \alpha + \beta\, x_j + \varepsilon_i \tag{5}$$

[15] Migrants are defined as individuals who lived in a Southern county in 1940 and were located in a different county in 1950.

[16] Boustan (2010) presents a more heterogeneous story of selection in which Black migrants drew from both the top and bottom (but not the middle) of the socio-economic distribution.

We face two main endogeneity concerns in estimating β. First, economic conditions can encourage people to leave an area and directly affect the outcomes of those who do not leave. Second, if migrants are positively selected on ability, we would expect those who remain in places that experience more out-migration to be lower in the ability distribution than those who remain in places with less out-migration. We formalize and discuss our approach to addressing these issues below.

5.1 Endogeneity Issues

Two issues can lead to selection into treatment, being in a high-migration county, among stayers. We formalize these issues, which may prevent us from obtaining unbiased estimates of the theoretical relationships derived in our model, using potential outcomes notation in Appendix A.3.2.

First, there could be a correlation between treatment and the economic conditions that lead to out-migration. That is, if areas with high and low out-migration are different, then we would expect counterfactual outcomes among stayers in these areas to be different. To address this issue, we construct a shift-share instrument to exploit quasi-random variation in out-migration. By providing quasi-random variation in out-migration, the shift-share instrument helps us address concerns that pre-existing differential trends in wages caused counties to experience different levels of out-migration, as changes in these counties' economic conditions could affect both wages and out-migration.

Second, stayers in areas of high and low out-migration will be different if migration is selected (i.e., if the distribution of counterfactual outcomes is different for stayers versus migrants). In Appendix A.3.3, we augment our model from Section 3 by adding worker heterogeneity by skill. At the county level, we show that average wages of non-migrants could change if selected migration causes the skill composition of the remaining workers to change. To address this issue, our analysis primarily focuses on changes in wages at the individual level, which is not directly affected by these compositional shifts.

At the individual level, though, selected migration leads to bias if people in low-migration counties who do not migrate are different from people in high-migration counties who do not migrate.[17] Specifically, if stayers have higher ability in counties with low out-migration than in counties with high out-migration due to selection, this selection will likely bias downward our estimates of the effect of out-migration on the income of those who do not migrate. We provide a formal statement of the assumptions required to produce this downward bias in Appendix A.3.4, and demonstrate this downward bias by simulating a simple Roy model in Appendix A.3.5.

To address selection into our sample of stayers, we focus on outcomes that are less likely to shift with stayer status, such as changes in wages instead of the level of wages. We show in Appendix A.3.4 that in a simple Roy model of selected migration, selection would lead to biased estimates of the effect of out-migration on the level of wages, but it would not lead to biased estimates of the effect of out-migration on changes in wages over time. Though we explore augmenting the Roy model to allow for the possibility of selected migration biasing

[17] The effect of this migrant selection has not received widespread attention in the literature on migration, even though it may impact the interpretation of results even if shocks to migration are randomly assigned.

changes in wages over time in the Appendix, we still believe that estimating the effects of out-migration on changes in economic outcomes instead of the level of these outcomes will produce estimates that are less likely to be biased.

Additionally, we use the approach taken by authors including Mishra (2007) to focus on populations with low probability of migration. That is, we examine population groups with less selection into staying, those for whom the decision to stay is (less) related to wage changes and the county's amount of out-migration. In other words, in a population (such as the elderly) that is unlikely to migrate, we do not expect the treatment to shift the wage distribution among stayers, as we assume that migrants within this population have chosen to migrate for idiosyncratic reasons unrelated to ability.

To restrict our non-migrant sample to populations with low probability of migration, we examine the sample of individuals who are predicted, using 1940 covariates, to be stayers. Specifically, we run a probit regression to predict whether an individual stayed in the same county in 1940 and 1950 using variables from the 1940 census such as age, gender, education, occupation, income, and farm ownership status. We then restrict our sample to individuals with greater than a given threshold probability of being a stayer.

Restricting our sample to predicted stayers alleviates issues of selection into the non-migrant sample by constructing comparable samples of stayers across our treatment. That is, we expect that in this population the difference in counterfactual outcomes between stayers in high and low migration counties will be smaller. In Section 8, we show our estimates using the predicted sample of stayers and find that the estimated coefficients are stable as we change the threshold probability for being a predicted stayer.

5.2 Shift-Share Instrument for Out-Migration Rate

We construct an instrument to isolate net out-migration that is unrelated to changes in county j's underlying economic conditions that could both change wages and induce out-migration. The shift-share instrument z_j that we construct for each Southern county creates quasi-random variation in our treatment, the net Black out-migration rate x_j, using two forms of variation. It makes use of variation in connections between Southern counties and Northern cities in 1940 as well as variation between Northern cities in the growth of manufacturing employment among non-Black individuals during the 1940s.

5.2.1 Construction of Shift-Share Instrument

We construct the shift-share instrument z_j for net Black out-migration from Southern county j between 1940 and 1950 as follows:

$$z_j = \sum_c s_{jc} g_c \tag{6}$$

where, borrowing notation from Borusyak, Hull, and Jaravel (2022), s_{jc} (share) is the share of the Black migrant population of Southern county j that moved to Northern city c between 1935 and 1940, and g_c (shift, or shock) is predicted growth in non-Black manufacturing employment in city c between 1940 and 1950.

We use the 1940 full count census to calculate the shares s_{jc}. These shares vary at the county level, so an individual i residing in county j will inherit the county-level shares:

$s_{ijc} = s_{jc}$. The 1940 full count census records where Black individuals who moved to Northern cities lived in both 1935 and 1940. We define

$$s_{jc} = \frac{\text{Number of Black migrants from } j \text{ to } c \text{ between 1935 and 1940}}{\text{Total number of Black migrants from } j \text{ between 1935 and 1940}} \tag{7}$$

The shocks g_c measure predicted growth in the number of non-Black manufacturing employees in city c between 1940 and 1950. We focus on non-Black manufacturing to reduce endogeneity concerns stemming from Black migrants who left the South affecting employment growth in Northern cities. To construct g_c, we estimate the fitted values of the following equation using predictors selected using LASSO:

$$\tilde{g}_{c,1950} = \beta_0 + P'_{c,1940}\,\boldsymbol{\beta} + \varepsilon_{c,1950}$$

where $\tilde{g}_c$ is the percent growth in the number of working age (18-55 years old) non-Black manufacturing employees in city c between 1940 and 1950 and $P'_{c,1940}$ are predictor variables, measured in 1940, as selected by LASSO.[18] Identification comes from our assumption that growth in non-Black manufacturing employment attracted Black individuals to move from Southern counties to Northern cities, but is otherwise unrelated to economic conditions in Southern sending counties.

5.2.2 Instrument Validity

The recent econometric literature on shift-share instruments provides three recommendations for conducting correct inference. These recommendations address issues of instrument exclusion, inference, and bias, and share elements with a more general literature on instrument "recentering".

First, identification for shift-share instruments relies on either exogenous shocks or exogenous shares. Our instrument for the net Black out-migration rate relies on exogenous shocks, and satisfies the exclusion restriction if our shocks are conditionally exogenous. Specifically, the shocks g_c, which measure predicted growth in non-Black manufacturing employment in Northern cities, must be unrelated to weighted, unobserved determinants of changes in economic conditions in Southern counties (Borusyak, Hull, and Jaravel 2022). We will write and test our exclusion restriction at the shock level, as explained below.

Second, Borusyak, Hull, and Jaravel (2022) demonstrate that shift-share regressions need to be run at the level of the shocks to obtain correct standard errors.[19] We follow their

[18] The predictors selected by LASSO are labor force participation share, farm share, homeowner share (restricted to working age population), occupational income score per capita (restricted to working age population), 1940 wage income per capita (restricted to working age population), 1940 Black population share, and WWII spending per capita. Our data for $\tilde{g}_c$ comes from the 1940 full count census, 1950 full count census, and 1947 County Data Book. In Appendix Table A.6, we show that our main results are robust to using actual growth in non-Black manufacturing employment in city c between 1940 and 1950 as our shocks.

[19] Both Borusyak, Hull, and Jaravel (2022) and Adao, Kolesár, and Morales (2019) show that conventional standard errors may be invalid because observations with similar values of the shares s_{jc} may have correlated residuals.

suggestion with our sample of 493 Northern cities, as described below.[20] Therefore, the unit of observation in our regressions are Northern cities that received black migrants from Southern counties. There are 1,037 Southern counties represented in this sample.

To estimate the shock-level regression, we convert our individual-level measures of changes in economic outcomes and county-level measures of changes in the net out-migration rate from equation 5 to the level of the shocks.[21] The estimating equation that generates our main results is therefore the following shock-level regression, which is estimated using regression weights $s_c = \sum_i s_{ijc}$ and instrument g_c (the shocks):

$$\bar{y}_c^{\perp} = \alpha + \beta \bar{x}_c^{\perp} + \bar{\varepsilon}_c^{\perp} \tag{8}$$

Though we obtain our results from estimating the shock-level equation 8, the point estimate of β is equivalent to that from using a standard shift-share instrument in the individual-level regression equation 5 (Borusyak, Hull, and Jaravel 2022). Therefore, our estimates of β can be interpreted at the individual level, and reflect the magnitude of the changes in economic outcomes resulting from a one percentage point increase in the county Black net out-migration rate.

We can further use the shock-level notation to write the exclusion restriction. Our exclusion restriction, equation 9, means that predicted city-level shocks to manufacturing employment g_c must be orthogonal to county-level unobservables.[22] For example, there should not be systematic differences between the unobserved determinants of economic conditions in Southern counties that have strong migrant-network connections to Northern cities with low versus high predicted growth in non-Black manufacturing employment.

$$\mathbb{E}\left[\sum_c g_c \, s_c \tilde{\varepsilon}_c\right] = 0 \tag{9}$$

Though we cannot explicitly test the exclusion restriction, we can test whether potential confounding variables are correlated with the instrument, as recommended by Borusyak, Hull, and Jaravel (2022) in a type of "pretrend test". A natural choice for a confounding variable is the outcome for the period prior to our analysis. Unfortunately, the 1930 Census did not record income. Instead, we analyze our measures of occupational prestige and two measures of employment in the 1930s as proxies for income in these pretrend tests. For the individual-level variables, we estimate equation 10 using data on individual i in 1930 and 1940, a vector of individual-level controls $\mathbf{c_i}$, and instrument z_j.[23] A statistically significant coefficient δ would indicate that our instrument is correlated with an unobserved confounding variable.

[20] As suggested by Borusyak, Hull, and Jaravel (2022), we compute the inverse of the Herfindahl index of the exposure shares to measure the effective sample size of this regression. Though our sample includes 493 cities, our effective sample size using city-level shocks is 16.

[21] In practice, we use the Stata program *ssaggregate* to do so, converting x_j to $\bar{x}_c^{\perp}$ and y_i to $\bar{y}_c^{\perp}$. See Appendix A.3.6 or Borusyak, Hull, and Jaravel (2022) for more details.

[22] One may think of the county-level unobserved characteristics as the average ε_i for individuals in a given county. Individual-level unobservable characteristics ε_i are transformed to the shock level and weighted by exposure shares to become the $\tilde{\varepsilon}_c$ used in the exclusion restriction. See the Appendix for details on the construction of s_c and $\tilde{\varepsilon}_c$, or refer to Borusyak, Hull, and Jaravel (2022).

[23] Though we wrote this equation at the individual level, we estimate this regression at the shock level, as explained above.

$$\Delta_{1930,1940}\, y_i = \alpha + \delta\, x_{j\,1940,1950} + \gamma \mathbf{c_i} + \varepsilon_i \tag{10}$$

The results of the pretrend tests are displayed in Appendix Tables A.7, A.8 and A.9. These results indicate that county-level out-migration before 1940 and our occupational stature variables are not strongly correlated with the instrument in the pre-period for either Black or White stayers. In Appendix Table A.7, we find a statistically significant relationship in column one for Black non-migrants, but it is no longer significant after adding controls. When we use change in log occscore between 1930 and 1940 as an outcome, we have a significant result for Black non-migrants; however, the effects are small in magnitude and are not robust to adding controls.[24] For White non-migrants, we do not find statistically significant effects for change in occscore. The results for change in log occscore are small and not robust to adding individual-level controls.

In Appendix Table A.8, we do not find compelling evidence of pretrends in employment status. For both Black and White non-migrants, we have a significant result when we use change in employment status as an outcome (column 1), but it is no longer significant after adding controls (column 2). We find no effect when we use change in labor force participation as an outcome for Black non-migrants (column 4). For White non-migrant, we find a statistically significant relationship in columns three and four. However, these relationships are of opposite sign, and not robust to the inclusion or exclusion of our individual-level control variables. Therefore, we do not believe these estimates indicate that our IV results for income will be systematically biased.

Additionally, in Appendix Table A.9, we analyze whether our instrument is predictive of migration flows before 1940 by regressing the treatment variable for the period prior to our analysis (net out-migration from 1930 to 1940) on our instrument for out-migration in the 1940s. We do not find evidence of a statistically significant relationship, buttressing our exclusion restriction and indicating that our instrument can detect a change in migration flows that occurred in the 1940s, instead of simply predicting long-standing trends in out-migration from Southern counties.

Following the third recommendation in the recent econometric literature, we control for the individual-level sum of exposure shares in our regressions. Borusyak, Hull, and Jaravel (2022) and Borusyak and Hull (2023) demonstrate that one should control for the sum of weighted exposure shares $\sum_c s_{ijc}$ in order to remove bias and "recenter" the shift-share instrument. Therefore, we control for the individual-level sum of exposure shares in our regressions. For a given individual, the sum of these weighted exposure shares is the fraction of Black migrants from their county that migrated to the Northern cities in our data.

5.2.3 First Stage

Our instrument z_j predicts county-level net Black out-migration rates (Figure 2). The first-stage relationship is positive, as expected, indicating that counties with strong migrant-

[24] In Appendix Table A.7 column 3, we have that a one percentage point increase in the net Black out-migration rate decreases occscore by 0.5%. However, compared to an average increase of 13%, this effect appears fairly small. Additionally, this effect and the effect of Black out-migration on change in log occscore between 1940 and 1950 (Table A.10 column 2) are of opposite sign. Therefore, if our results were biased, it is likely that our results would be underestimated.

network connections to Northern cities that experienced more growth in non-Black manufacturing employment sent out more migrants than counties with strong connections to Northern cities that experienced less growth in non-Black manufacturing employment. F-statistics for each specification are included in our tables of results.

6 Results

Using our instrumental variables strategy to empirically analyze changes in non-migrant income, we find that out-migration increases incomes of those who do not migrate. The income gains from migration are higher in lower-skilled occupations and among lower-income individuals. Additionally, we find that counties with high out-migration saw declines in the number of farmers and farms, as well as lower racial animosity in the 1960s.

6.1 Income

We find that Black non-migrants living in places with high out-migration during the 1940s experienced increases in wage income. In Table 2, we report our main income results for Black Southerners, in which our outcome of interest is change in income ($y_i = \Delta\omega_B$). Regressing change in income on our treatment, the net Black out-migration rate, we find using OLS that a one percentage point increase in the Black net out-migration rate increases incomes among those who do not migrate by 1.4% (Table 2 Panel A, column 1). When we instrument for net Black out-migration using our shift-share instrument constructed in Section 5, we find that a one percentage point increase in the Black net out-migration rate causes Black non-migrant income to increase by 4.2% (Table 2 Panel B, column 1).[25]

All of these results are statistically significant, and we find similar effects when we use alternative specifications for the change in wage income. The effects on wage income are similar, through larger, when we control for age, martial status, gender, and occscore in 1940 and the county-level growth in occscore per capita between 1930 and 1940 (Table 2 Panel B, column 3). When we use change in log wage income as our outcome, we find that, among Black non-migrants with positive reported wage income in 1940 and 1950, a one percentage point increase in the Black net out-migration rate increases wage income by 2.5% (Appendix Table A.10 Panel A, column 3). We also find increases in wage income for Black non-migrants when we include those with zero wages in 1940 and control for individual characteristics (Appendix Table A.11 Panel A).

To put our results on wage income in context, Collins and Wanamaker (2014) finds that, after the first wave of the Great Migration (1910-1930), the incomes of migrants were 89-

[25]Our SSIV results are larger than our OLS results because the economic characteristics in Southern counties that are correlated with Black out-migration most likely would have reduced non-migrant wages. For instance, negative shocks to Southern agriculture were possible factors that pushed Black Southerners to the North (Lange, Olmstead, and Rhode 2009; Boustan, Kahn, and Rhode 2012; Hornbeck and Naidu 2014; Ager, Brueckner, and Herz 2017; Feigenbaum, Soumyajit Mazumder, and Smith 2020). Ager, Brueckner, and Herz (2017) finds that boll weevil infestations in the South between 1889 and 1929 decreased farm wages. These negative shocks to agriculture, such as pest infestations or natural disasters, likely increased out-migration and worsened economic outcomes, lowering incomes for those who remained in the South. Therefore, our OLS results should be biased downward.

90 log points higher than those of non-migrants in a simple comparison of the two groups. Analyzing changes in individual income over time, they find that between 1910 and 1930 the earnings of migrants increased, on average, by 63 log points more than the earnings of non-migrants. Boustan (2016) finds even larger income increases for migrants, suggesting that the annual earnings of Southern Black men could have increased by 82 log points through migration to the North in 1940. Our estimated coefficients imply that the income increases one could realize from the average net Black out-migration rate (17.4) are comparable, though smaller, than those reported above for migration to the North.[26] These estimates are on the order of 54 to 70 log points depending on the exact specification used.[27] Our own analysis of the data suggests that gains in wage income were larger for the average migrant than for the average non-migrant, for both Black and White workers (Column four of Appendix Tables A.12 and A.13, respectively).

We find similar, though smaller, effects on the income of White non-migrants. We repeat the above analyses with a sample of White individuals who remained in the South between 1940 and 1950. In Panel B of Table 3, we find that income increased by around 2% for every one percentage point increase in the Black net out-migration rate. We find similar income effects when we include county and individual controls in columns 2 and 3. When we use change in log wage income as our outcome, we find that a one percentage point increase in the Black net out-migration rate increases wage income by 2.4% for White non-migrants (Table A.10 Panel B, column 3). Our results, though, suggest that while incomes increased due to Black migration for White individuals who remained in the South, consistent with our theoretical model, the estimated increases in income are smaller than for the Black non-migrant population.

6.1.1 Heterogeneity

We estimate the effects of Black out-migration on wage income separately by occupation, gender, and income level. We find that the largest increases in income are in the farming and service sectors, sectors that lost many workers due to out-migration. We also find that wage gains are concentrated among those with low levels of income in 1940.

When we estimate the effects of Black out-migration on wage income by occupation in Table 4, we find that non-migrants employed in farming and services in 1940 saw the largest increases in wage income. When we restrict the sample to those employed in a farming occupation in 1940, we find that a one percentage point increase in the Black net out-migration rate increased wage income by 4.4%. For service workers, we find an increase of 5%.

As with the Black population, we find large increases in income for White residents who are employed in farming and service occupations. As shown in Table 5, we find that a one percentage point increase in the Black net out-migration rate increased wage income

[26] The average net Black out-migration rate for our sample of Black non-migrants is 17.44. See Table 1 for details.

[27] Our estimated coefficients also imply that the income increases one could realize from a one standard deviation (10) increase in the net Black out-migration rate are comparable, though smaller, than those reported above for migration to the North. These estimates are on the order of 35 to 47 log points depending on the exact specification used.

by 2.5% for White non-migrants employed in farming in 1940. We find similar effects for White service workers. Additionally, White non-migrants experience wage increases for all occupation categories, with wage increases for White workers being largest for farm employees and service workers. The difference between the effect on Black and White workers may imply a segregated labor market for farm workers, in which White workers were not a ready substitute for Black workers.

For both Black and White non-migrants, the effect of out-migration on wage income do not differ significantly by gender.[28] The main results are very similar to the results for Black men, but the income results for Black women are larger, yet insignificant (likely due to a smaller sample size). For White non-migrants, the income results are generally similar across genders, though income increases more for White men than for White women. Though there are small differences in the coefficients for men and women, these coefficients are not statistically different from each other for either Black or White non-migrants. However, given the large increases in income for women, the labor shortages caused by out-migration may have encouraged non-migrant women to work more hours or shift into higher paying occupations that were previously held by migrants.

Income increases also had distributional effects. We find that the wage gains due to out-migration are concentrated in the bottom half of the 1940 income distribution. When we focus on Black non-migrants with income in 1940 below the median level in the sample, we find that the results are comparable to our main results (Appendix Table A.16 columns 1 and 2). However, we find no wage effects for the sample in the top half of the 1940 income distribution (Appendix Table A.16 columns 3 and 4). We find similar results for White non-migrants (Appendix Table A.17). These results suggest that the negative shock to the Black labor supply benefited low-income non-migrants the most.

6.2 County-level Results

To understand how the reduction in the Black labor supply impacted local communities, we estimate the effect of Black out-migration on county-level income, farming, and racial attitudes. Using the same empirical strategy outlined in Section 5, we find that, while counties with high Black out-migration saw increases in county-level wage income for Black and White residents, they also experienced declines in the number of farms and farmers. Additionally, these counties have less racial animosity on average in the 1960s. Our evidence suggests that while there were large gains for Black non-migrants in regards to wages and social environment, some farmers had to shut down or consolidate their farms due to the higher costs for labor.

We find that places with high levels of Black out-migration experienced overall increases in wage income between 1940 and 1950. In Table 6, we report county-level changes in average wage income by race. We find that a one percentage point increase in the Black net out-migration rate increased overall average wage income by 1.9% in the South. For Black and White Southerners, we find a 2.5% and 3.1% increase, respectively. When using aggregate data, we find wage gains for both Black and White Southerners between 1940 and 1950 that are in line with our results using individual-level data. As described in Section 6.1, though,

[28] See Appendix Tables A.14 and A.15 for details.

these increases in income are smaller than those enjoyed by out-migrants.

Examining agricultural operations, we find declines in the number of farms and farmers, though farms also became larger and more valuable on average. In Table 7, we report county-level changes in farming, specifically the number of farmers, the number of farms, average farm value, average farm size, and total farming area. We find that a one percentage point increase in the Black net out-migration rate decreased the number of farmers by 1.8%. We also find that places with high levels of Black out-migration saw declines in the number of farms but increases in the average farm value and farm size, with no change in the total farming area. These findings suggest that the Great Migration led to farm consolidation, with some Southern non-migrants, particularly White owners of certain types of farms, experiencing economic consequences. The labor supply shock may have pushed small farmers to close their farms or to consolidate with other farms as previously-cheap farm labor became more expensive.

Next, to understand how Black out-migration impacted White Southerners' attitudes, we estimate the effect of out-migration on racial attitudes and find that counties with more Black out-migration had less racial animosity in the 1960s. In Table 8, we report the effects of Black out-migration on county-level vote share for George Wallace in the 1968 presidential election and the number of KKK chapters in the 1960s. We find that places with high levels of Black out-migration during the 1940s had lower vote share for George Wallace in 1968; however, the estimates are not statistically significant. Examining KKK chapters, we find that a one percentage point increase in the Black net out-migration rate decreases the number of KKK chapters by 0.1 percentage points. These results are robust to the inclusion of the change in county average Black wage income between 1940 and 1950 and the total Black population in 1940 as controls (Table 8 columns 2 and 4), indicating that out-migration had an effect on racial attitudes that was separate from the increases in income we have discussed above.

Though we believe out-migration affected racial animosity independently of wage changes, racial animosity may have affected wages. We provide evidence of this when we estimate our main results on wage income for counties with and without a history of lynching. In Appendix Table A.18, we find that places with any Black lynchings between 1882 and 1930 saw larger increases in wage income for both Black and White non-migrants, relative to counties with no history of lynchings. Since places with more lynchings between 1882 and 1990 had higher out-migration (Appendix Table A.3), these results suggest that wage increases were used to compensate for racial violence and discourage Black laborers from leaving for the North. We find similar results when splitting counties by low and high George Wallace vote share (Appendix Table A.19). Taken together with our results using 1960s measures of racial attitudes, White Southerners may have used higher wages and less animosity as ways to retain Black labor. Additionally, we have shown that the reduction in the Black labor supply provided higher-paying employment opportunities for both Black and White non-migrants, so the wage gains for White non-migrants may have also allowed a less hostile social environment to develop. These findings align with those of Feigenbaum, Soumyajit Mazumder, and Smith (2020), who studied Black out-migration due to the boll weevil, and show that Black out-migration at that time (before the Great Migration) reduced racial animosity.

7 Mechanisms

Our main results indicate that out-migration increases income among both Black and White non-migrants, with the effect being larger for Black non-migrants. That the empirical findings align with the theoretical predictions from Section 3 suggests that the migration-induced labor supply shock is a key mechanism driving our wage results. In this section, we investigate other mechanisms that could explain non-migrant wage gains following mass Black out-migration. Specifically, we consider sorting, agglomeration effects, occupation switching, and changes in the production process. We find that, in addition to the negative labor supply shock, switching occupations and changes in the production process drive our main results.

7.1 Sorting

Sorting, induced by selected migration, may change the composition and total factor productivity of the workforce. This effect is embodied in the A term our production function, equation 1. Since migrants have more education than non-migrants, A may decline as those better-educated migrants leave the Southern labor force.

Previous empirical studies have demonstrated that the overall education level of the workforce can affect the productivity of each worker. These social returns to education have been studied extensively. In the literature cited by Ciccone and Peri (2006), 10% is the highest estimate of the social returns to one year of education that they cite (though their own estimates were much lower).

Using a 10% social return to an additional year of education to provide an upper bound on our back-of-the-envelope calculation of the size of the education externality, we expect that sorting could change incomes by less than one percent. In Appendix Table A.4, we find that Black migrants have an extra 0.42 years of average education (4.83 years) than Black non-migrants (4.41 years). For a one percentage-point increase in the Black out-migration rate, and given that the average county in our sample has a Black share of 0.27, the average education level of the population would therefore decline by $0.42 * 0.01 * 0.27 = 0.001$. For a 10% social return on each year of additional education, a decrease in average education of 0.001 years leads to an estimated decrease in wages, due to the smaller positive externality, of 0.01%. This estimate is more than two orders of magnitude smaller than our main estimates of the effect of out-migration on wages in Table 2, and has the opposite sign. That is, we find that out-migration has a positive effect on the wages of non-migrants despite the loss of better-educated migrants, which, by itself, may slightly decrease wages.

7.2 Agglomeration

Migration patterns affect the population density of the areas in which the Southern residents who do not migrate live. The production function in our model, equation 1, allows for the resulting changes in agglomeration economies to affect wages. Agglomeration effects are embodied in the σ term of our model, and our expression for log wages demonstrates that wages are a function of the agglomeration parameter σ. To determine whether agglomeration effects could be driving our main results, we follow the approach used by De la Roca and Puga (2017).

The first step of this approach is to calculate the static advantages of living in a county with a larger population. To do this calculation, we regress log wages in 1940 of individual i living in county j on worker characteristics x_i' (education and dummies for occupational categories) and county fixed effects μ_j:

$$\log y_{ij} = \alpha + \boldsymbol{x_i'}\beta + \mu_j + \eta_{ij} \tag{11}$$

In the second step, we regress the estimated county fixed effects μ_j from the first step on the log county population as measured in 1940:

$$\mu_j = \alpha + \beta \log(\text{pop}_j) + \eta_j \tag{12}$$

Our estimate of β from equation 12, as displayed in column one of Table 9, represents an elasticity of 0.15 of the earnings premium with respect to county population. This estimate is much larger than the city-size premium of 0.0455 estimated by De la Roca and Puga (2017), though the binscatter in Figure 3 displays some non-linearity in our estimated agglomeration effects.

This estimated wage premium elasticity, as De la Roca and Puga (2017) argue, does not account for the potential of workers to sort according to unobserved ability based on county size. We investigate the possibility of worker sorting on unobserved ability based on county size by including individual fixed effects. To do so, we use the matched 1940 and 1950 census sample. To estimate both individual and county fixed effects, we only include individuals who moved between counties, and who have data on wage income and education (or occupation - the results are similar) in both time periods. Including individual fixed effects in the first step of our estimation leads us to estimate the following two regression equations:

$$\begin{aligned} \log y_{ijt} &= \alpha + \beta x_{it} + \mu_j + \phi_i + \xi_{ijt} \\ \mu_j &= \alpha + \beta \log(\text{pop}_j) + \xi_j \end{aligned} \tag{13}$$

Our estimate of β when including individual fixed effects and using workers' location in 1940, as displayed in column two of Table 9, represents an elasticity of approximately 0.13 of the earnings premium with respect to county population. This estimate is smaller than when we exclude the individual fixed effects, similarly to De la Roca and Puga (2017).[29]

Interestingly, when we include workers' 1950 location in the computation of the county fixed effects, the estimated agglomeration effect is much larger. In column three of Table 9, including individual fixed effects leads the estimated elasticity of wages with respect to 1940 population to more than double. This estimate provides suggestive evidence that, between 1940 and 1950, individuals with higher unobserved ability moved away from larger counties, so our estimate of the true agglomeration effect increases when we account for this sorting.

Therefore, though we find evidence that larger counties have higher wages, we find that agglomeration effects likely do not explain our results. Overall, given that the average county in our sample has a Black share of 0.27, a one percentage-point increase in the out-migration rate would decrease the population by 0.27%. This reduction in population would lead wages

[29]However, the decrease is not as substantial as in De la Roca and Puga (2017), where including the individual fixed effects reduces the estimated elasticity by about half.

to fall by between approximately 0.03 percent and 0.1 percent due to the agglomeration effects estimated in Table 9. Therefore, our main estimated effects of how out-migration increases wage income, as displayed in Table 2, would likely have been slightly larger if not for the reduced agglomeration effects that accompany out-migration.

7.3 Occupation Switching

Since we have shown that sorting and agglomeration should have led out-migration to decrease non-migrant wages slightly, we next examine whether switching occupations could explain our main finding that non-migrant wages increased with out-migration. In line with these predictions, Foged and Peri (2016) find that non-migrating Danish workers switched occupations when immigration changed the supply of workers in different occupations. In our context, we find evidence that Black and White non-migrants move into higher-paid occupations and occupations that saw wage increases following Black out-migration. Black non-migrant men in high out-migration counties are less likely to remain farmers, while Black non-migrant women are more likely to become farm laborers by 1950. White non-migrants are more likely to switch into non-farming service occupations.

For our sample of non-migrants with non-missing values for our main outcome, change in wage income between 1940 and 1950, we find some evidence of Black non-migrants switching to more prestigious occupations. In Table 10, we report our results using change in OCCSCORE between 1940 and 1950 as our outcome. In Panel A, we find that Black out-migration has no statistically significant effect on Black non-migrants' OCCSCORE (column 1). However, controlling for age, marital status, gender and years of education in 1940, we find that a one percentage point increase in the Black net out-migration rate increases occupational stature for Black non-migrants by 0.006 (column 2). The standard deviation of the net Black out-migration rate in our sample is approximately 10, so this coefficient implies that a one standard deviation increase in the net Black out-migration rate would increase the occupational score by about 0.06 (compared to a mean decrease in occupational score of around 0.02 in this sample).

When we stratify the sample by gender, we find that Black non-migrant women move into higher-paid occupations. For Black non-migrant men, the effect of Black out-migration on OCCSCORE is not robust to controlling for 1940 individual characteristics (Appendix Table A.20 Panel A). However, for Black non-migrant women, we find that a one percentage point increase in the Black net out-migration rate increases occupational stature by 0.014 (Appendix Table A.20 Panel A, column 3). When we control for age, marital status, and years of education in 1940, we find that occupational stature increases by 0.021 (Appendix Table A.20 Panel A, column 4). This coefficient implies that a one standard deviation increase in the net Black out-migration rate would increase the occupational score by about 0.21 (compared to a mean decrease in occupational score of around .25 in this sample).

Looking at changes in occupations, we find that Black non-migrants also switched into low-skilled occupations that were previously held by migrants. We find that a one percentage point increase in the Black net out-migration rate decreased the likelihood of being a farmer by 0.008 percentage points, and the likelihood of being a farm laborer increased by 0.008 percentage points (Table 11 Panel A). Stratifying the sample by gender, we find that Black non-migrant women were less likely to remain in occupations categorized as "other worker"

(primarily laundresses, teachers, and cooks, as shown in Appendix Table A.21 Panel B, column 5), while Black non-migrant men were less likely to remain as service workers (Appendix Table A.21 Panel A, column 4). These results suggest that Black out-migration caused some Black non-migrants to shift into lower-skilled occupations that were newly available and offering higher wages due to the negative labor supply shock.

White non-migrants also switched to occupations with higher occupational income scores. In Table 10 Panel B, we find that a one percentage point increase in the Black net out-migration rate increases occupational stature for White non-migrants by 0.007 (column 1). We find similar results when we include individual-level controls. This coefficient implies that a one standard derivation increase in the net Black out-migration rate would increase the occupational score by about 0.07 (compared to a mean decrease in occupational score of around 0.05 in this sample). However, we also find some evidence consistent with White non-migrants switching into occupations vacated by Black workers who out-migrated. For example, in Appendix Table A.22 we show that White non-migrants are more likely to earn less than the federal minimum wage in 1950 in counties that experienced more Black out-migration, which reduces the gap between the share of White non-migrants earning less than the minimum wage and the share of Black non-migrants earning less than the minimum wage.

Consistent with increases in occscore, we find that White non-migrants are more likely to switch into service occupations. We find that a one percentage point increase in the Black net out-migration rate increased the likelihood of being a service worker by 0.016 percentage points (Table 11 Panel B). These results are robust to the inclusion of individual-level control variables (Appendix Table A.23). When we stratify the sample by gender, we find similar changes for White non-migrant men and women, but the increases are more pronounced for women (Appendix Table A.24). Similar to Black women, White non-migrant women are also less likely to remain in other worker occupations. For White women, these occupations were primarily teachers, sales clerks, and stenographers, typists, and secretaries. These results suggest that Black out-migration pushed both male and female White non-migrants into non-farming service occupations.

Overall, out-migration promotes some who remain to move into higher-status occupations; it encourages others to move into the lower-status occupations most affected by the labor shortage. From the OCCSCORE results we find that non-migrants are switching into higher-status occupations, which can increase income for non-migrants. Additionally, from the occupation results, we find evidence of Black non-migrants, mainly Black women, switching into wage labor positions, which were occupations that migrants were more likely to be employed in in 1940.[30] We find that White non-migrants are switching to different occupations, such as non-farming service roles; these occupations were also occupations that Black migrants were more likely to be employed in in 1940.[31] Given our model of migration-induced labor supply shocks, we would expect occupations directly impacted by the labor shortage, such as wage farm laborers, to see the largest wage increases. Therefore, occupation switching, whether through moving into higher-status occupations or through switching into low-paying occupations that experienced wage increases, can explain our main results.

[30] See Appendix Table A.4.

[31] See Appendix Table A.4.

7.4 Changes in Production Process

While occupation switching may lead to wage increases, out-migration can also influence the wages of those who do not migrate by changing the production process. Faced with changing relative costs of labor and capital, farm owners may change to a more capital-intensive productive process in the long run. This adjustment has been documented by Hornbeck and Naidu (2014) for an earlier, more localized, migration of Southern Black residents. This shift towards capital investment may both increase the share of income flowing to the owners of capital and also increase the marginal product of labor, thereby changing wages. In our model, these effects are both embedded in equation 1. Our expression for log wages demonstrates how decreasing the labor share $1-\alpha$ would likely decrease wages, while increasing the level of capital K would increase wages.

To measure changes in the production process, we calculate the labor share of total income in each Southern county in 1950.[32] Specifically, we aggregate total wage income, variable INCWAGE in the full-count 1950 census, and total income, variable INCTOT, for each county. We then take the ratio of these two variables. We additionally perform this calculation when only including wage income of farm laborers (individuals whose occupation category is "farm laborers, wage workers") in the numerator. While researchers sometimes back out the capital share from the computation of the labor share, we also compute an explicit capital share using business and farm income, INCBUSFM, as the numerator in our ratio.

Our results indicate that counties with more Black out-migration have a lower labor share of income and a higher capital share of income. This evidence is consistent with the findings of Hornbeck and Naidu (2014). In column one of Table 12, we see that an increase of one additional Black net out-migrant per 100 Black residents reduces the labor share by about 0.01, which is a negative and statistically significant reduction. The increase in the capital share is smaller, at 0.005, but is also statistically significant. Consistent with our finding that the number of operational farms decreases, the share of all income accruing to farm laborers declines with out-migration as well, though it declines less than the overall labor share.[33]

Our main results, that wages for those who do not migrate increase with out-migration, can be consistent with an increase in the capital intensity of the production process. The marginal product of labor increases with higher capital intensity, raising wages for those who remain employed. However, these changes in factor prices may also reduce labor demand, lowering employment and the aggregate income of laborers. Our results are mostly consistent with this story, as we demonstrate in Appendix Table A.25 that employment and labor force participation drop more among individuals in counties with more out-migration (though the statistical significance of these coefficients is not robust to our inclusion of control variables).

[32] Total income is not available in the 1940 Census, preventing us from doing the same calculation for 1940.

[33] The share of all income accruing to farm laborers and the labor share are measured with the same denominator, so the decrease in the share of all income accruing just to farm laborers cannot exceed the decline in the share of all income accruing to labor (the labor share).

8 Robustness

In this section we test the robustness of our results by addressing concerns about selection into being a non-migrant, selection into census linking, measurement of farmers' wages, and nonlinear effects. We find that our results are robust to using comparable samples of non-migrants and to re-weighting the sample by the inverse probability of being linked to the census. Additionally, we find that, even when we exclude individuals classified as farm owners or tenant farmers in 1940, our main results for wage income still hold. We do not find strong evidence of nonlinear effects.

8.1 Selection into Being a Non-Migrant

As outlined in Section 5.1, our main estimates for Black non-migrants may be underestimated due to selection. If migrants are positively selected, the quality of non-migrants will vary across our treatment. Specifically, those who do not migrate will be drawn from lower in the ability distribution in high-migration counties, likely biasing our results downward.

To address this selection issue, we restrict our sample to populations that are less likely to migrate and show that our results are robust to using this sample. We do this by constructing a sample of predicted non-migrants using 1940 covariates and estimating the effect of Black out-migration on wage income for this sample. We find that our main results on income are robust to using predicted non-migrants. In Table 13, we estimate the effect of Black out-migration using our sample of predicted stayers. The results are similar in magnitude to those in our main results table. They are also robust to changing the probability threshold at which we classify someone as having a high predicted probability of being a non-migrant (Figure A.2).[34]

8.2 Selection into Census Linking

Compared to the full sample of Black and White Southerners with non-zero income in 1940, our sample of non-migrants, those for whom we can link their 1940 census records to their 1950 census records, are older, more likely to be male, more likely to be married, and have higher wage income in 1940.[35] If there is non-random selection into the likelihood of being linked in the census, our results may be biased. For our results to be overestimated, our sample of non-migrants with non-zero income in 1940 would need to be overrepresented by non-migrants with wage gains. This bias would also need to be larger in higher out-migration counties and smaller in lower out-migration counties.

To address this selection issue, we follow Aizer, Grafton, and Pérez (2025) and weight our sample by the inverse probability of being linked to the census. Specifically, we estimate a probit model of the likelihood of being linked to the 1940 and 1950 census using 1940 characteristics.[36] Using this estimated probability, we find that our results are robust to

[34]The effects are noisy for high probability thresholds due to small sample sizes at thresholds greater than .85.

[35]See Appendix Table A.2 for details. These differences are from comparing columns 2 and 5.

[36]We use the following 1940 covariates: urban status, gender, martial status, occupation, county of residence, employment status, income, birth place, industry, and age.

re-weighting the sample to account for selection on observables between individuals linked and not linked to the 1940 and 1950 censuses (Appendix Table A.26).

8.3 Farmers' Wages

As discussed in Section 4.1, the 1940 census does not measure wage income for self-employed individuals, including farmers.[37] In all our results on wage income, we drop individuals with missing wage income in 1940 or 1950. Despite dropping a sizable number of farmers, there still remain some farmers in our sample of non-migrants.[38] These remaining farmers most likely have a small amount of positive income from employment that is not farming.

Since wage income in the 1940 census does not accurately represent their economic standing, we show that our main results are robust to dropping those classified as farmers in 1940. The results are shown in Appendix Table A.27. When dropping those classified as farmers in 1940, a one percentage point increase in the Black net out-migration rate increases wage income by 3.5%. While this estimate is smaller in magnitude than our main results, it is still large and statistically significant.[39] For White non-migrants, a one percentage point increase in the Black net out-migration rate increases wage income by 1.6%, which is comparable to our main results, when we drop those classified as farm owners or tenant farmers in 1940. The results are similar to the main results when we control for individual characteristics. Our main findings still hold when we exclude farmers who may have inaccurate measures of wage income.

Since the 1940 census does not include in-kind payments (i.e. living quarters, meals, supplies, etc.) as a part of wage income either, we show that our main results are also robust to dropping those large in-kind payments. In-kind payments may have been prominent among agricultural workers (Carruthers and Wanamaker 2017). Therefore, wage income for workers who also received in-kind payments will be underestimated, which can bias our measure of change in wage income. To alleviate this issue, we follow Carruthers and Wanamaker (2017) by including a measure of change in wage income that excludes those

[37] The census occupation code for farmers includes both farm owners and tenant farmers, who we collectively refer to as "farmers".

[38] There are 27,765 Black and 102,617 White non-migrant farmers with non-zero income in 1940 and reported wage income in 1950.

[39] The results for non-farmers may be smaller for two reasons: issues with measuring 1940 income for farmers may cause our main results to be overestimated and farmers actually experienced larger wage income increases than non-farmers.

First, if there is measurement error in the 1940 census measure for personal wage income that causes the change in wage income for farmers to be too large, our main results would be overestimated. Even if this is the case, we still find that non-farmers experience large and statistically significant increases in wage income, so our findings are not solely driven by farmers.

Second, it may be that farmers in our sample actually experienced larger increases in wage income than non-farmers. Farmers may have switched to jobs that earn wages in 1950, or those who remained farmers in both 1940 and 1950 may have earned more from a secondary job that is paid in wages, both of which would lead to farmers having larger wage increases. Since the census only records one occupation, most likely the primary job, it is unclear whether some of the wage income comes from secondary jobs. Since farmers in 1940 should not have wage income (because they are self-employed), it may be the case that those with reported wage income have another occupation that is not reported in the census. For example, the large negative labor supply shock may have caused farmers to seek additional income, especially since farm employees saw the largest increases in wage income.

with large in-kind payments. Using the 1940 census variable for non-wage income, we provide estimates of change in wage income between 1940 and 1950 that exclude those with large in-kind payments. Specifically, when we exclude individuals with more than $50 in non-wage earnings in 1940, we find that our results are robust to excluding these individuals (Appendix Table A.28). We also find similar results when we exclude both farmers and those with more than $50 in non-wage earnings in 1940 (Appendix Table A.27 columns 3 and 4).

8.4 Nonlinearity

In all our specifications so far, we assume that the net Black out-migration rate has a linear effect on our outcomes. However, it may be that net Black out-migration actually has a non-linear effect. If effects are non-linear, counties with low levels of out-migration will have different effects, potentially even in the opposite direction, than counties with high levels of out-migration. For example, places experiencing a small labor supply shock could have a large positive effect on wage income due to the change in supply of labor paired with small changes in consumption. However, places with a relatively large share of Black Southerners leaving may actually have smaller or negative wage income effects due to a sizable reduction in consumption. A substantial decline in consumption could ultimately reduce labor demand and decrease non-migrant wages, despite the reduction in the labor supply.

To check for nonlinearity, we conduct a heterogeneity analysis where we stratify counties by low and high predicted net Black out-migration. First, we predict our treatment, the 1940-1950 net Black out-migration rate, using possible predictors from Appendix Table A.3.[40] After constructing a measure of predicted Black out-migration, we classify counties with a predicted net Black out-migration rate at or above the median value as "predicted high counties". Counties below the median value are considered "predicted low" Black out-migration counties. Using this classification for counties, we estimate the effects of the 1940-1950 net Black out-migration rate on the change in wage income for Black and White non-migrants by predicted Black out-migration (Appendix Table A.29).

We find suggestive (but statistically insignificant) evidence that Black out-migration has a nonlinear effect, though both types of counties still experience positive effects of out-migration on non-migrant wage income. Comparing columns 1 and 2 in Appendix Table A.29, we find that the effect of Black out-migration on Black non-migrant wage income in predicted low counties is higher than the effect for those in predicted high counties. However, the effect for the predicted low sample is very noisy, and the predicted high sample still experiences a 3.4% increase in wage income following a one percentage point increase in the net Black out-migration rate. For White non-migrants, we find the opposite. While While out-migration has positive effects in both predicted low and high counties, the effect of Black out-migration on White non-migrant wage income is lower for the predicted low sample than for the predicted high sample. While there is some suggestive evidence of nonlinear effects, the coefficient estimates for the predicted low and high-migration counties are not

[40] We use LASSO to select our set of 1940 predictors. Using OLS, we estimate the fitted values of the regression of 1940-1950 net Black out-migration rate on LASSO selected predictors. The set of LASSO predictors are: Black share; % employed, % employed in agriculture; area; % urban; average farm value; average farm size; share of white farmers; share of full owned farms; share of tenant farms; % males in labor force; WWII spending per capita. All variables are measured in 1940, except for WWII spending per capita.

statistically different from each other, and we still see large, positive effects for counties with both high and low levels of Black out-migration.

9 Conclusion

While most studies of migration have focused on migrants themselves, or impacts on the destination locations, migration also affects those who remain behind. However, the effect of out-migration on those who do not migrate is theoretically ambiguous.

We address two econometric issues to empirically investigate the impact of migration on stayers. First, push factors such as changing economic conditions may both encourage migration and directly affect those who do not migrate. To address this issue, we focus on a time period during which migration flows sharply increased among Black Americans: the Great Migration. We construct a shift-share instrument based on pre-existing migrant networks between Southern counties and Northern cities, and growth in predicted non-Black manufacturing employment in those Northern cities. Our instrument satisfies the exclusion restriction if the shocks are exogenous, that is if changes in predicted non-Black manufacturing employment in Northern cities are unrelated to unobserved determinants of economic characteristics in Southern counties.

Second, if out-migrants are positively selected on ability, then the average ability of non-migrants will differ between areas with high and low out-migration due to selection. To address this issue, we check our results for robustness by analyzing a subpopulation that has a low probability of migrating, for whom selection into migration is less likely.

Addressing these econometric issues using the approaches described above, we find that Black net out-migration increases wage income for both Black and White non-migrants, with larger effects for the former, aligning with the predictions of our theoretical model. Investigating heterogeneity in effects by occupation, we find that the largest increases in income are for farm employees, an occupation in which county-level employment dropped sharply during the Great Migration. Despite large wage gains for non-migrants, counties with high out-migration saw substantial reductions in the number of farmers and farms, suggesting there were losses for self-employed, predominately White, farm owners.

Along with the negative supply shock for Black labor, we find that non-migrant wage gains from Black out-migration can be explained by occupation switching and changes in the production process. Though sorting and agglomeration should depress non-migrant wages in the presence of selected out-migration, we estimate that these effects are small, and therefore do not explain the overall positive effect of out-migration on wages. We do find that Black out-migration increases occupation stature for both Black and White non-migrants. At the same time, Black non-migrants are also more likely to shift into occupations that were directly impacted by the negative labor supply shock. These results suggest that changing occupations explains our wage results as some non-migrants switch into higher-status occupations and others switch into low-skilled occupations that enjoyed large wage increases following Black out-migration. We also find that counties with high out-migration saw increases in the capital share, which can increase wages by raising the marginal product of labor. Moreover, our results cannot be explained by non-migrant selection, census linking selection, measurement of farmers' wage income, or non-linearity of effects.

If the economic fortunes of those who are left-behind continue to diverge from those of everyone else, additional policies will be needed to help those who do not or cannot migrate. This issue has become especially relevant recently, as many American cities are experiencing population decline. While many proposed policies focus on helping left-behind individuals directly, our results suggest that encouraging those who are able to leave less-productive areas to do so may have a salutary effect on those who remain, especially if those who leave and those who remain were previously competing in the same sector of the labor market. In the context of the Great Migration, we show that out-migration increases non-migrant wages.

During the Great Migration, migrants left their jobs in rural farm communities to seek better employment opportunities in the North. Southern farmers still needed laborers, even as economic conditions changed, since land for farming operations could not be moved and agricultural sales do not depend solely on local demand. However, for labor markets that depend primarily on providing services for the local community, future researchers may find that out-migration decreases non-migrant wages because out-migration leads to declines in local consumption, reducing the demand for labor. More research is needed to understand how out-migration impacts non-migrants in labor markets where production relies on local demand. Additionally, since our study occurs in a time period in which there were lower levels of direct government assistance and infrastructure spending, more research is needed into how out-migration affects the provision and quality of government services enjoyed by those who remain.

References

Adao, Rodrigo, Michal Kolesár, and Eduardo Morales (2019). "Shift-share designs: Theory and inference". In: The Quarterly Journal of Economics 134.4, pp. 1949–2010.

Ager, Philipp, Markus Brueckner, and Benedikt Herz (2017). "The boll weevil plague and its effect on the southern agricultural sector, 1889–1929". In: Explorations in Economic History 65, pp. 94–105.

Aizer, Anna, Gabrielle Grafton, and Santiago Pérez (2025). "Daughters as Safety Net? Family Responses to Parental Employment Shocks: Evidence from Alcohol Prohibition". In: National Bureau of Economic Research.

Akbar, Prottoy A, Sijie Li Hickly, Allison Shertzer, and Randall P Walsh (2022). "Racial segregation in housing markets and the erosion of black wealth". In: Review of Economics and Statistics, pp. 1–45.

Akram, Agha Ali, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak (2017). "Effects of emigration on rural labor markets". In: National Bureau of Economic Research.

Anderson, D Mark, Benjamin Crost, and Daniel I Rees (2020). "Do economic downturns fuel racial animus?" In: Journal of Economic Behavior & Organization 175, pp. 9–18.

Anderson, Robert Warren, Noel D Johnson, and Mark Koyama (2017). "Jewish persecutions and weather shocks: 1100–1800". In: The Economic Journal 127.602, pp. 924–958.

Austin, Benjamin A, Edward L Glaeser, and Lawrence H Summers (2018). Jobs for the Heartland: Place-based policies in 21st century America. Tech. rep. National Bureau of Economic Research.

Black, Dan A, Seth G Sanders, Evan J Taylor, and Lowell J Taylor (2015). "The impact of the Great Migration on mortality of African Americans: Evidence from the Deep South". In: American Economic Review 105.2, pp. 477–503.

Borjas, George J (1987). "Self-Selection and the Earnings of Immigrants." In: American Economic Review 77.4.

— (2003). "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market". In: The Quarterly Journal of Economics 118.4, pp. 1335–1374.

Borusyak, Kirill and Peter Hull (2023). "Nonrandom exposure to exogenous shocks". In: Econometrica 91.6, pp. 2155–2185.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel (2022). "Quasi-experimental shift-share research designs". In: The Review of Economic Studies 89.1, pp. 181–213.

Boustan, Leah Platt (2010). "Was postwar suburbanization "white flight"? Evidence from the black migration". In: The Quarterly Journal of Economics 125.1, pp. 417–443.

— (2016). Competition in the promised land: Black migrants in northern cities and labor markets. Vol. 39. Princeton University Press.

Boustan, Leah Platt, Matthew E Kahn, and Paul W Rhode (2012). "Moving to higher ground: Migration response to natural disasters in the early twentieth century". In: American Economic Review 102.3, pp. 238–244.

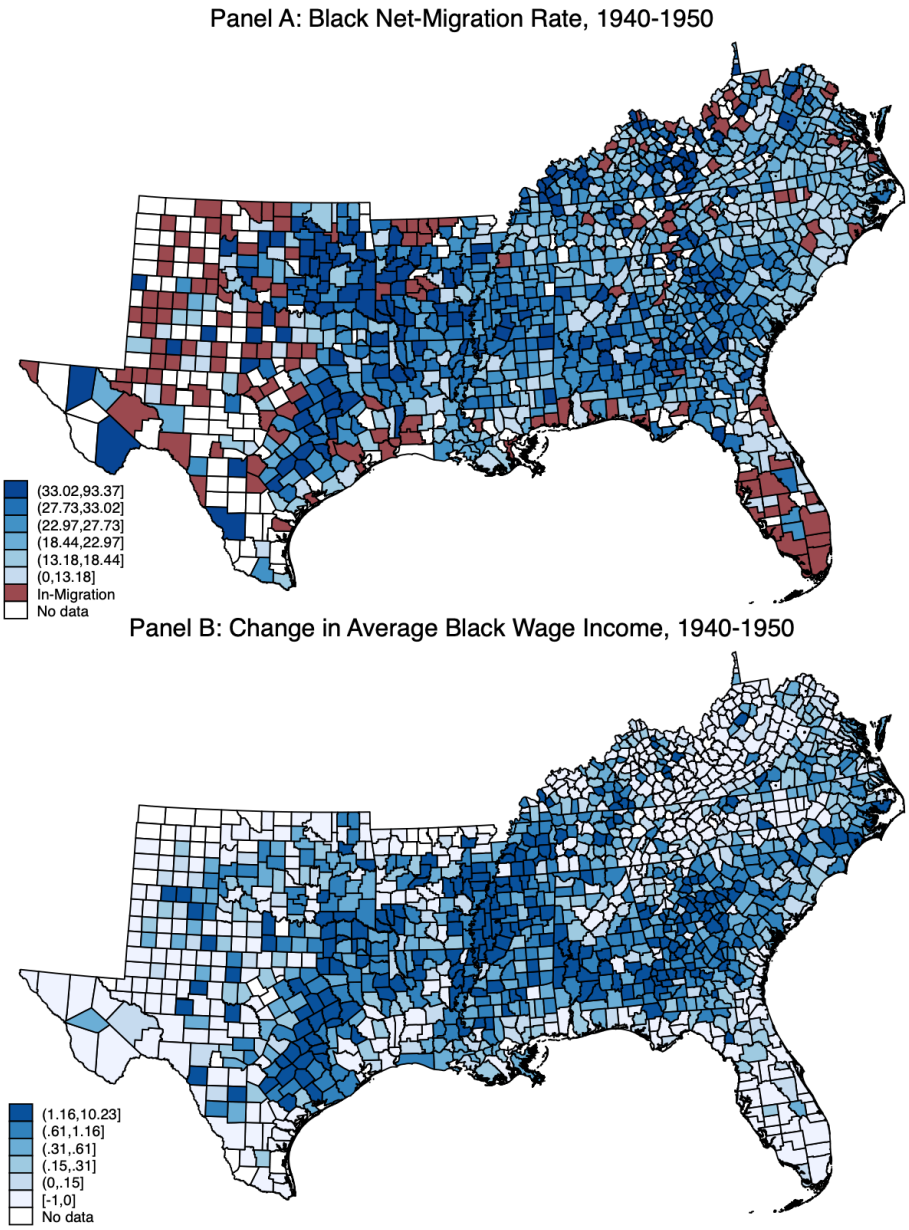
Burke, Marshall, Solomon M Hsiang, and Edward Miguel (2015). "Climate and conflict". In: Annu. Rev. Econ. 7.1, pp. 577–617.

Busso, Matias, Jesse Gregory, and Patrick Kline (2013). "Assessing the incidence and efficiency of a prominent place based policy". In: American Economic Review 103.2, pp. 897–947.
Calderon, Alvaro, Vasiliki Fouka, and Marco Tabellini (2023). "Racial diversity and racial policy preferences: the great migration and civil rights". In: The Review of Economic Studies 90.1, pp. 165–200.
Carrington, William J, Enrica Detragiache, and Tara Vishwanath (1996). "Migration with endogenous moving costs". In: The American Economic Review, pp. 909–930.
Carruthers, Celeste K and Marianne H Wanamaker (2017). "Separate and unequal in the labor market: human capital and the jim crow wage gap". In: Journal of Labor Economics 35.3, pp. 655–696.
Census, United States. Bureau of the (1947). Statistical Abstract of the United States, 1947. 68. US Bureau of the Census.
Christian, Cornelius (2017). "Lynchings, labour, and cotton in the US south: A reappraisal of Tolnay and Beck". In: Explorations in Economic History 66, pp. 106–116.
Ciccone, Antonio and Giovanni Peri (2006). "Identifying human-capital externalities: Theory with applications". In: The Review of Economic Studies 73.2, pp. 381–412.
Clubb, Jerome M., William H. Flanigan, and Nancy H Zingale (2006). Electoral Data for Counties in the United States: Presidential and Congressional Races, 1840-1972. https://doi.org/10.3886/ICPSR08611.v1. Inter-university Consortium for Political and Social Research [distributor].
Collins, William J (2021). "The Great Migration of Black Americans from the US South: A guide and interpretation". In: Explorations in Economic History 80, p. 101382.
Collins, William J and Marianne H Wanamaker (2014). "Selection and economic gains in the great migration of African Americans: new evidence from linked census data". In: American Economic Journal: Applied Economics 6.1, pp. 220–252.
— (2015). "The great migration in black and white: New evidence on the selection and sorting of southern migrants". In: The journal of economic history 75.4, pp. 947–992.
Collins, William J and Ariell Zimran (2025). "World War II service and the GI Bill: New evidence on selection and veterans' outcomes from linked census records". In: Explorations in Economic History 96, p. 101649.
De la Roca, Jorge and Diego Puga (2017). "Learning by working in big cities". In: The Review of Economic Studies 84.1, pp. 106–142.
Derenoncourt, Ellora (2022). "Can you move to opportunity? Evidence from the Great Migration". In: American Economic Review 112.2, pp. 369–408.
Deryugina, Tatyana, Laura Kawano, and Steven Levitt (2018). "The economic impact of Hurricane Katrina on its victims: Evidence from individual tax returns". In: American Economic Journal: Applied Economics 10.2, pp. 202–233.
Diamond, Rebecca (2016). "The determinants and welfare implications of US workers' diverging location choices by skill: 1980–2000". In: American economic review 106.3, pp. 479–524.
Dinkelman, Taryn and Martine Mariotti (2016). "The long-run effects of labor migration on human capital formation in communities of origin". In: American Economic Journal: Applied Economics 8.4, pp. 1–35.

Docquier, Frédéric and Hillel Rapoport (2012). "Globalization, brain drain, and development". In: Journal of economic literature 50.3, pp. 681–730.
Dustmann, Christian, Uta Schönberg, and Jan Stuhler (2017). "Labor supply shocks, native wages, and the adjustment of local employment". In: The Quarterly Journal of Economics 132.1, pp. 435–483.
Economist, The (2017). "The right way to help declining places". In: The Economist.
Feigenbaum, James J, Soumyajit Mazumder, and Cory B Smith (2020). When Coercive Economies Fail: The Political Economy of the US South After the Boll Weevil. Tech. rep. National Bureau of Economic Research.
Fite, Gilbert C (1979). "Southern agriculture since the Civil War: an overview". In: Agricultural History 53.1, pp. 3–21.
Foged, Mette and Giovanni Peri (2016). "Immigrants' effect on native workers: New analysis on longitudinal data". In: American Economic Journal: Applied Economics 8.2, pp. 1–34.
Foote, Christopher L, Warren C Whatley, and Gavin Wright (2003). "Arbitraging a discriminatory labor market: black workers at the Ford Motor Company, 1918–1947". In: Journal of Labor Economics 21.3, pp. 493–532.
Gardner, John and William Cohen (1992). Demographic Characteristics of the Population of the United States, 1930-1950: County-Level. https://doi.org/10.3886/ICPSR00020.v1. Inter-university Consortium for Political and Social Research [distributor].
Goldin, Claudia and Robert A Margo (1992). "The great compression: The wage structure in the United States at mid-century". In: The Quarterly Journal of Economics 107.1, pp. 1–34.
Haines, Michael R.and ICPSR (2010). Historical, Demographic, Economic, and Social Data: The United States, 1790-2002. https://doi.org/10.3886/ICPSR02896.v3. Inter-university Consortium for Political and Social Research [distributor].
Hines, Elizabeth and Eliza Steelwater (1882–1930). Historical American Lynching Data Collection Project. Project HAL [dataset].
Holley, Donald (2000). Second Great Emancipation: Mech. cottonpicker, Black Migration & Modern South (c). University of Arkansas Press.
Hornbeck, Richard and Suresh Naidu (2014). "When the levee breaks: black migration and economic development in the American South". In: American Economic Review 104.3, pp. 963–990.
Jia, Ning, Raven Molloy, Christopher Smith, and Abigail Wozniak (2023). "The economics of internal migration: Advances and policy questions". In: Journal of economic literature 61.1, pp. 144–180.
Khanna, Gaurav, Emir Murathanoglu, Caroline B Theoharides, and Dean Yang (2022). Abundance from abroad: Migrant income and long-run economic development. Tech. rep. National Bureau of Economic Research.
Kline, Patrick and Enrico Moretti (2014). "Local economic development, agglomeration economies, and the big push: 100 years of evidence from the Tennessee Valley Authority". In: The Quarterly journal of economics 129.1, pp. 275–331.

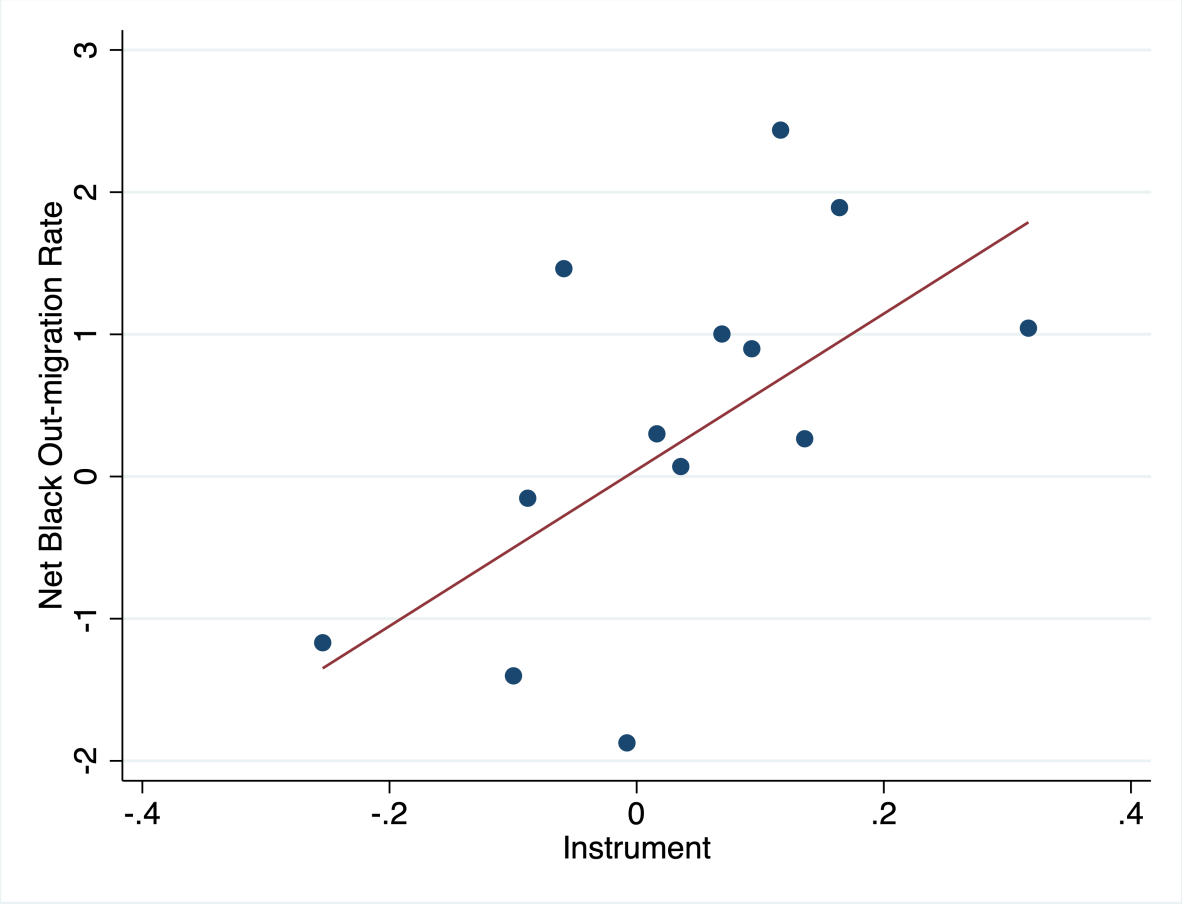
Labor Statistics, U.S. Bureau of (2025). CPI Inflation Calculator. URL: https://www.bls.gov/data/inflation_calculator.htm. (accessed: July 15, 2025).
Lange, Fabian, Alan L Olmstead, and Paul W Rhode (2009). "The impact of the boll weevil, 1892–1932". In: The Journal of Economic History 69.3, pp. 685–718.
Manson, Steven, Jonathan Schroeder, David Van Riper, Katherine Knowles, Tracy Kugler, Finn Roberts, and Steven Ruggles (2024). IPUMS National Historical Geographic Information System: Version 19.0 [dataset]. http://doi.org/10.18128/D050.V19.0. Minneapolis, MN: IPUMS: Minneapolis: University of Minnesota.
Mazumder, Shom (2018). Replication Data for: The Persistent Effect of U.S. Civil Rights Protests on Political Attitudes. https://doi.org/10.7910/DVN/WKJJ3Z. Harvard Dataverse.
Mishra, Prachi (2007). "Emigration and wages in source countries: Evidence from Mexico". In: Journal of development economics 82.1, pp. 180–199.
Roy, Andrew Donald (1951). "Some thoughts on the distribution of earnings". In: Oxford economic papers 3.2, pp. 135–146.
Ruggles, Steven, Catherine A. Fitch, Ronald Goeken, J. Robert Warren, J. David Hacker, Matt Nelson, Evan Roberts, and Matthew Sobek (2024). IPUMS Ancestry Full Count Data: Version 4.0 [dataset]. https://doi.org/10.18128/D014.V4.0. Minneapolis, MN: IPUMS: Minneapolis: University of Minnesota.
Ruggles, Steven, Sarah Flood, Matthew Sobek, Daniel Backman, Grace Cooper, Julia A. Rivera Drew, Stephanie Richards, Renae Rodgers, Jonathan Schroeder, and Kari C.W. Williams (2025). IPUMS USA: Version 16.0 [dataset]. https://doi.org/10.18128/D010.V16.0. Minneapolis, MN: IPUMS: Minneapolis: University of Minnesota.
Ruggles, Steven, Nesile Ozder, Catherine A. Fitch, Matthew Sobek, Julia A. Rivera Drew, J. David Hacker, Jonas Helgertz, Cheyenne Lonobile, Matt A. Nelson, Evan Roberts, and J. Robert Warren (2025). IPUMS Multigenerational Longitudinal Panel: Version 2.0 [dataset]. https://doi.org/10.18128/D016.V2.0. Minneapolis, MN: IPUMS: Minneapolis: University of Minnesota.
Shrestha, Slesh A (2017). "No man left behind: Effects of emigration prospects on educational and labour outcomes of non-migrants". In: The Economic Journal 127.600, pp. 495–521.
Tolnay, Stewart E and Elwood M Beck (1992). "Racial violence and black migration in the American South, 1910 to 1930". In: American Sociological Review, pp. 103–116.
Yang, Dean (2008). "International migration, remittances and household investment: Evidence from Philippine migrants' exchange rate shocks". In: The Economic Journal 118.528, pp. 591–630.

Figure 1: Southern Black Net-Migration and Changes in Average Black Wages, 1940-1950



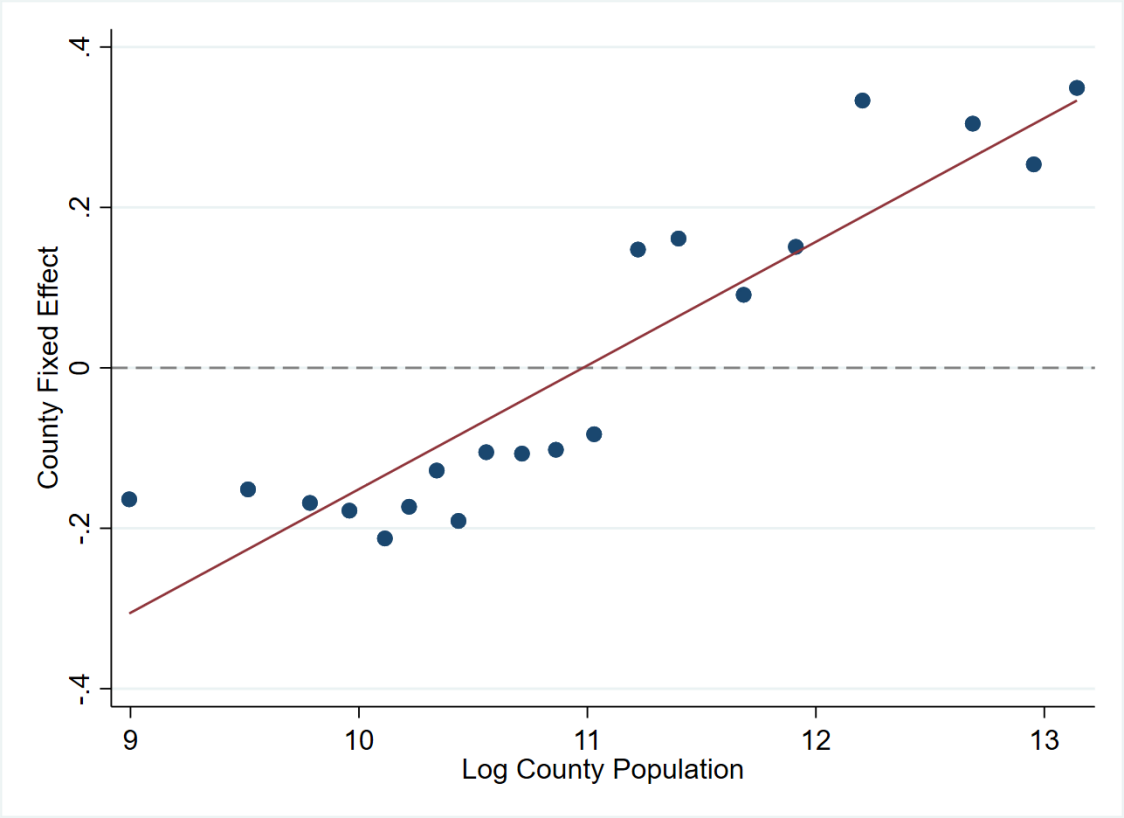
Note: Panel A reports the Black net-migration rate for counties with net out-migration between 1940-1950. The data on Black net-migration rates for Southern counties comes from Boustan (2016). Net-migration rate is Black net-migration between 1940-1950 divided by the county Black population in 1940, multiplied by 100. Panel B reports county-level change in average Black wage income (in 1940 dollars) between 1940 and 1950. Average county-level wage income is calculated as the percentage change in average wage income between 1940 and 1950 among all Black residents with non-missing wage income.

Figure 2: First stage on Net Black Out-Migration



Note: Binscatter plot the relationship between our shift-share instrument and the net Black Out-Migration rate at the shock (city) level. The coefficient equals 5.49 with a F-stat of 16.1. This figure includes 493 Northern cities.

Figure 3: Effect of County Population on Income



Note: Binscatter at the individual level. The dependent variable is the estimated fixed effect for each county in a regression of wage income on education, occupation category, and county fixed effects. The independent variable is the logarithm of the county's total population in 1940.

Table 1: Summary Statistics: Southern Non-Migrants, 1940-1950

	(1) Black Non-Migrants	(2) White Non-Migrants
Age (1940)	35.73	36.17
	(12.52)	(12.22)
Male	0.69	0.80
	(0.46)	(0.40)
Married (1940)	0.81	0.79
	(0.39)	(0.40)
Years of Education (1940)	5.18	8.46
	(3.35)	(3.69)
OCCSCORE (1940)	14.59	23.05
	(8.10)	(9.32)
Wage Income (1940)	319.22	823.37
	(318.77)	(782.20)
Net Black Out-migration Rate (1940-1950)	17.44	18.31
	(10.20)	(11.27)
Change in Wage Income (1940-1950)	-0.04	-0.49
	(8.17)	(4.38)
Change in Log Wage Income (1940-1950)	0.94	0.49
	(1.23)	(1.18)
Change in Occscore (1940-1950)	-0.02	-0.05
	(0.95)	(0.81)
N	445410	1629999

Note: Data from the 1940 full count census, 1950 full count census and Boustan (2016). The change in occupational income score and wage income between 1940 and 1950 are calculated as the relative change, while the change in farm employment status between 1940 and 1950 is the difference between indicator variables for each variable in 1940 and 1950.

Table 2: Main Results: Black Non-Migrants

Panel A: OLS

	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)
Net Black out-migration rate	0.0143***	0.0144***	0.0134***
	(0.0018)	(0.0018)	(0.0018)
Individual Observations	445,410	445,410	445,410
County Controls	No	Yes	Yes
Individual Controls	No	No	Yes
Dependent Mean	-0.04	-0.04	-0.04

Panel B: IV

	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)
Net Black out-migration rate	0.0417***	0.0436***	0.0584***
	(0.0148)	(0.0162)	(0.0176)
Individual Observations	445,410	445,410	445,410
Number of Shocks	493	493	493
County Controls	No	Yes	Yes
Individual Controls	No	No	Yes
Dependent Mean	-0.04	-0.04	-0.04
Effective F-statistic for IV	16.1	14.7	14.7

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. For county controls, we control for growth in per capita occscore between 1930 and 1940. Individual controls include age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table 3: Main Results: White Non-Migrants

Panel A: OLS

	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)
Net Black out-migration rate	0.0024***	0.0024***	0.0021***
	(0.0004)	(0.0004)	(0.0004)
Individual Observations	1,629,999	1,629,999	1,629,999
County Controls	No	Yes	Yes
Individual Controls	No	No	Yes
Dependent Mean	-0.49	-0.49	-0.49

Panel B: IV

	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)
Net Black out-migration rate	0.0193***	0.0204***	0.0206***
	(0.0032)	(0.0038)	(0.0042)
Individual Observations	1,629,999	1,629,999	1,629,999
Number of Shocks	493	493	493
County Controls	No	Yes	Yes
Individual Controls	No	No	Yes
Dependent Mean	-0.49	-0.49	-0.49
Effective F-statistic for IV	13.9	12.6	12.4

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. For county controls, we control for growth in per capita occscore between 1930 and 1940. Individual controls include age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table 4: Results by Occupation for Black Non-Migrants

	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0417***	0.0442***	0.0503***	0.0215
	(0.0148)	(0.0140)	(0.0123)	(0.0135)
Individual Observations	445,410	114,244	195,436	119,087
Number of Shocks	493	493	493	493
Individual Controls	No	No	No	No
Dependent Mean	-0.04	0.14	-0.22	-0.18
Effective F-statistic for IV	16.1	25.9	12.4	9.2
Sample	Full Sample	Farm Employee	Service Worker	Other Worker

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table 5: Results by Occupation for White Non-Migrants

	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0193***	0.0246**	0.0247***	0.0085***
	(0.0032)	(0.0099)	(0.0039)	(0.0020)
Individual Observations	1,629,999	204,784	499,961	883,723
Number of Shocks	493	493	493	493
Individual Controls	No	No	No	No
Dependent Mean	-0.49	-0.17	-0.48	-0.61
Effective F-statistic for IV	13.9	8.2	17.6	12.8
Sample	Full Sample	Farm Employee	Service Worker	Other Worker

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table 6: Change in Average Wage Income by Race - County-level

	Change in Average Wage Income	Change in Average Black Wage Income	Change in Average White Wage Income
	(1)	(2)	(3)
Net Black out-migration rate	0.0190***	0.0246**	0.0313***
	(0.0053)	(0.0103)	(0.0068)
County Observations	1,037	1,031	1,037
Number of Shocks	493	493	493
County Controls	No	No	No
Dependent Mean	-0.75	-0.30	-0.93
Effective F-statistic for IV	13.4	14.2	13.4

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are log changes measured at the county level between 1940 and 1950. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table 7: Changes in Farming - County-level

	Change in # of Farmers	Change in # of Farms	Change in Average Farm Value	Change in Average Farm Size	Change in # Total Farm Size
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	-0.0179***	-0.0088***	0.0092***	0.0110***	0.0021
	(0.0051)	(0.0029)	(0.0027)	(0.0032)	(0.0038)
County Observations	1,037	1,026	1,026	1,026	1,026
Number of Shocks	493	491	491	491	491
Individual Controls	No	No	No	No	No
Dependent Mean	-0.28	-0.13	0.19	0.18	0.05
Effective F-statistic for IV	13.4	13.1	13.1	13.1	13.1

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are log changes measured at the county level between 1940 and 1950. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table 8: Racial Attitudes - County-level

	George Wallace Vote Share (1968)	George Wallace Vote Share (1968)	# of KKK Chapters (1960s)	# of KKK Chapters (1960s)
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.4219	-0.2384	-0.0993***	-0.0759***
	(0.4433)	(0.4286)	(0.0306)	(0.0255)
Observations	493	493	493	493
County Observations	1,036	1,036	1,037	1,037
Number of Shocks	493	493	493	493
County Controls	No	Yes	No	Yes
Dependent Mean	39.63	39.63	0.56	0.56
Effective F-statistic for IV	13.2	12.5	13.4	12.7

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are measured in levels. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. For county-level controls, we control for change in average Black wage income between 1940 and 1950 and total Black population in 1940. *** indicates significance at the 1% level.

Table 9: Agglomeration Effects

	Pooled OLS	Individual FE: 1940 Location	Individual FE
	(1)	(2)	(3)
Log of 1940 County Population	0.1502***	0.1390***	0.3857***
	(0.0001)	(0.0024)	(0.0028)
Constant	-1.6498***	-2.3365***	-4.5783***
	(0.0013)	(0.0258)	(0.0309)
Observations	2630577	22746	35193

Note: Regression at the individual level. The dependent variable is the estimated fixed effect for each county in a regression of wage income on individual characteristics and county fixed effects. The independent variable is the logarithm of the county's total population in 1940. *** indicates significance at the 1% level.

Table 10: Changes in Occscore: Black and White Non-Migrants

Panel A: Black Non-Migrants

	Change in Occscore	Change in Occscore
	(1)	(2)
Net Black out-migration rate	-0.0022	0.0061**
	(0.0024)	(0.0025)
Individual Observations	428,767	428,767
Number of Shocks	493	493
Individual Controls	No	Yes
Dependent Mean	-0.02	-0.02
Effective F-statistic for IV	15.9	16.3

Panel B: White Non-Migrants

	Change in Occscore	Change in Occscore
	(1)	(2)
Net Black out-migration rate	0.0070***	0.0062***
	(0.0018)	(0.0018)
Individual Observations	1,588,468	1,588,468
Number of Shocks	493	493
Individual Controls	No	Yes
Dependent Mean	-0.05	-0.05
Effective F-statistic for IV	13.9	13.6

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1940 and 1950 using full count census data. Outcomes are restricted to those with a non-missing value for our main outcome, change in wage income between 1940 and 1950. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Column 2 controls for age, gender, martial status, and years of education in 1940. *** indicates significance at the 1% level.

Table 11: Changes in Occupation: Black and White Non-Migrants

Panel A: Black Non-Migrants

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	-0.0077***	0.0083***	0.0031	-0.0025	-0.0070***
	(0.0017)	(0.0026)	(0.0020)	(0.0015)	(0.0021)
Observations	493	493	493	493	493
Individual Observations	430,597	430,597	430,597	430,597	430,597
Number of Shocks	493	493	493	493	493
Individual Controls	No	No	No	No	No
Dependent Mean	0.06	-0.13	-0.11	-0.04	0.04
Effective F-statistic for IV	16.0	16.0	16.0	16.0	16.0

Panel B: White Non-Migrants

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	0.0015*	-0.0032***	-0.0048***	0.0157***	-0.0105***
	(0.0009)	(0.0006)	(0.0011)	(0.0032)	(0.0031)
Observations	493	493	493	493	493
Individual Observations	1,597,535	1,597,535	1,597,535	1,597,535	1,597,535
Number of Shocks	493	493	493	493	493
Individual Controls	No	No	No	No	No
Dependent Mean	0.03	-0.05	-0.09	-0.08	0.04
Effective F-statistic for IV	13.9	13.9	13.9	13.9	13.9

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1940 and 1950 using full count census data. Outcomes are restricted to those with a non-missing value for our main outcome, change in wage income between 1940 and 1950. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table 12: County-level Labor Share

	Labor Share	Capital Share	Farm Laborer Share
	(1)	(2)	(3)
Net Black out-migration rate	-0.0108***	0.0053***	-0.0007**
	(0.0019)	(0.0015)	(0.0003)
County Observations	1,048	1,048	1,048
Number of Shocks	482	482	482
Individual Controls	No	No	No
Effective F-statistic for IV	17.2	17.2	17.2

Note: Exposure-robust standard errors are clustered at the shock (city) level. The dependent variables are the labor and capital shares for each Southern county in 1950. The independent variable is the net Black out-migration rate in 1940. *** indicates significance at the 1% level.

Table 13: Main Results: Predicted Non-Migrants

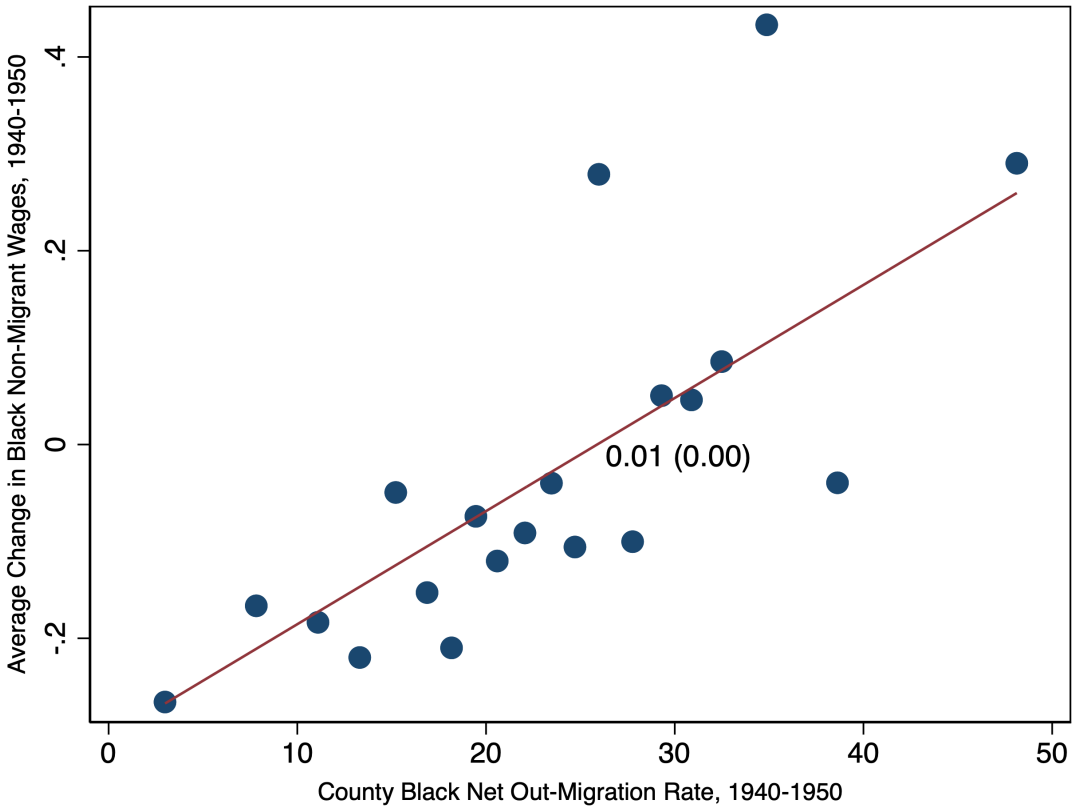
	Black Non-Migrants		**White Non-Migrants**	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0417***	0.0440***	0.0193***	0.0300***
	(0.0148)	(0.0139)	(0.0032)	(0.0075)
Individual Observations	445,410	121,315	1,629,999	258,321
Number of Shocks	493	493	493	493
Individual Controls	No	No	No	No
Dependent Mean	-0.04	-0.21	-0.49	-0.21
Effective F-statistic for IV	16.1	11.1	13.9	15.5
Sample	Actual	Predicted	Actual	Predicted

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

A Appendix

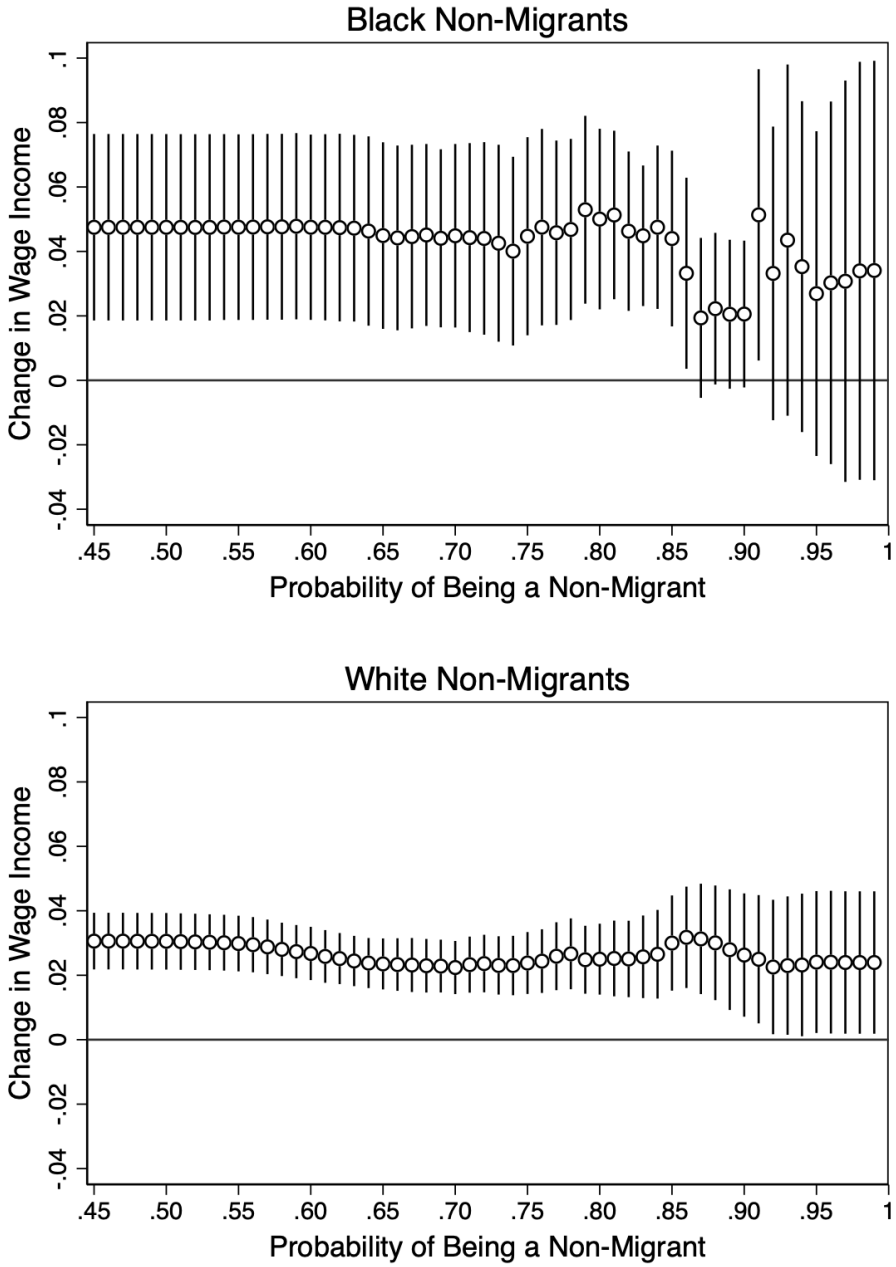
A.1 Appendix Figures

Figure A.1: Black Non-Migrant Wages and Black Out-Migration, 1940-1950



Note: Binscatter of the county-level 1940-1950 Black Net Out-Migration rate and county-level mean in the change in wage income between 1940 and 1950 for Southern Black Non-Migrants.

Figure A.2: Effect of Black Out-Migration on Change in Wage Income by Probability of Being a Non-Migrant



Note: This figure plots coefficients of regressions where change in income between 1940 and 1950 is the dependent variable and Black net out-migration rate is the independent variable, with 95% confidence interval. The x-axis denotes the threshold using to classify the predicted non-migrant sample. Black net out-migration is instrumented with our shift-share instrument. IV regressions are at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940.

A.2 Appendix Tables

Table A.1: 1940 Southern Agricultural Employment by Race

	(1) All	(2) Black	(3) White
Panel A: 1940 Southerners			
Share in Agriculture	0.32	0.11	0.21
Share in Agriculture (conditional on living in a rural area)	0.51	0.17	0.34
Share in Manufacturing	0.15	0.03	0.13
Share in Wholesale and Retail Trade	0.12	0.02	0.11
Share in Personal Services	0.09	0.06	0.03
Share in Construction	0.08	0.02	0.07
Non-Agricultural Laborers Share	0.13	0.05	0.07
Farmers Share (conditional on being in agriculture)	0.58	0.15	0.43
Wage Farm Laborers Share (conditional on being in agriculture)	0.22	0.10	0.11
Panel B: 1950 Southerners			
Share in Agriculture	0.21	0.07	0.14
Share in Agriculture (conditional on living in a rural area)	0.39	0.12	0.27
Share in Manufacturing	0.17	0.03	0.14
Share in Wholesale and Retail Trade	0.15	0.02	0.13
Share in Personal Services	0.07	0.04	0.03
Share in Construction	0.06	0.01	0.05
Non-Agricultural Laborers Share	0.06	0.03	0.03
Farmers Share (conditional on being in agriculture)	0.58	0.14	0.43
Wage Farm Laborers Share (conditional on being in agriculture)	0.18	0.08	0.10

Note: Shares calculated using the 1940 and 1950 full count censuses. The sample is restricted to those employed and living in a southern state. Panel A reports shares for 1940, while panel B reports shares for 1950. Column (1) reports the share of those employed in a specific occupation or industry, as a share of all working age(18-55 years old) employed southerners. Column (2) reports the share of Black Southerners employed in a specific occupation or industry, as a share of all working age employed southerners. Column (3) reports the same shares, but for White Southerners. The variable *Share in Agriculture (conditional on living in a rural area)* is the share of rural southerners employed in the agriculture industry, as a share of all working age employed southerners living in a rural area. While *Farmers Share* and *Wage Farm Laborers Share* are conditional on being employed in agriculture industry.

Table A.2: Comparison of Linked and Full Sample, Black and White Southerners, 1940

	(1) Full Sample	(2) Full Sample w/ Non-Zero 1940 Income	(3) Linked Sample	(4) Linked Non-Migrant Sample	(5) Linked Non-migrant Sample w/ Non-Zero 1940 Income	(6) Main Linked Non-Migrant Sample
Panel A: Black Southerners						
Age	25.95	34.45	22.70	23.60	35.73	29.75
Male	0.49	0.65	0.51	0.51	0.69	0.67
Married	0.47	0.70	0.47	0.48	0.81	0.61
Employed	0.53	0.87	0.54	0.54	0.88	0.74
In Labor Force	0.57	0.94	0.57	0.57	0.94	0.78
Years of Education	3.90	5.13	3.54	3.49	5.18	4.41
Occscore	5.32	13.76	4.79	4.87	14.59	8.89
Wage Income	79.26	287.96	69.04	70.34	319.34	129.04
N	7423760	1866577	2829122	2186635	446837	1198960
Panel B: White Southerners						
Age	27.96	34.98	26.38	27.81	36.18	32.01
Male	0.50	0.77	0.52	0.51	0.80	0.76
Married	0.49	0.69	0.54	0.56	0.79	0.63
Employed	0.46	0.90	0.48	0.48	0.91	0.75
In Labor Force	0.49	0.95	0.50	0.50	0.96	0.79
Years of Education	6.40	8.65	6.16	6.12	8.46	7.44
Occscore	7.34	22.26	7.35	7.50	23.05	14.45
Wage Income	183.69	761.15	189.26	187.32	823.06	360.71
N	19833912	4355085	10940779	7871288	1644942	4064162

Note: All columns are restricted to individuals living in Southern counties in 1940 that had net Black out-migration between 1940 and 1950. Panel A reports summary statistics for Black Southerners, while panel B reports summary statistics for White Southerners. Column (1) is the full sample from the 1940 full-count census. Column (2) is the full sample in 1940 with non-zero wage income in 1940. Column (3) is the sample from column (1) that we can link their records to the 1950 census. Column (4) is the sample from column (3) that are living in the same county in 1940 and 1950. Column (5) is the sample from column (4) but with non-zero wage income in 1940. Column (6) is our main sample of non-migrants. This sample include all individuals from column (5), as well as non-migrants from column (4) that are at least in the labor force in 1940 or 1950 and/or had a non-zero occscore in 1940.

Table A.3: County Summary Statistics - Low vs High Net Black Out-Migration

	(1) All	(2) Low Black Out-Migration	(3) High Black Out-Migration
Black Population Share (1940)	0.27	0.25	0.29
	(0.21)	(0.19)	(0.24)
% Males In Labor Force (1940)	80.16	79.69	80.64
	(4.25)	(4.48)	(3.96)
% Employed (1940)	44.22	44.66	43.79
	(5.86)	(5.14)	(6.47)
% Employed in Agriculture (1940)	49.63	42.48	56.76
	(19.61)	(19.71)	(16.71)
Density (1940)	94.78	149.89	39.79
	(455.15)	(639.10)	(24.30)
Area (sq. mile) (1940)	579.29	563.64	594.91
	(352.92)	(320.89)	(381.89)
% Urban (1940)	16.26	22.63	9.91
	(19.66)	(22.82)	(13.11)
Average value per farm ($'s) (1940)	4047.96	4572.04	3524.89
	(6022.89)	(6656.54)	(5270.46)
Average farm size (acres) (1940)	163.13	169.90	156.52
	(645.24)	(706.30)	(580.02)
Share of White Owned Farms (1940)	0.77	0.80	0.74
	(0.23)	(0.20)	(0.25)
Share of Farms with Full Owners (1940)	0.46	0.50	0.41
	(0.18)	(0.17)	(0.17)
Share of Farms with Part Owners (1940)	0.07	0.08	0.07
	(0.04)	(0.04)	(0.04)
Share of Manager-operated Farms (1940)	0.01	0.01	0.00
	(0.01)	(0.01)	(0.01)
Share of Tenant Farms (1940)	0.46	0.42	0.51
	(0.19)	(0.18)	(0.18)
Average Occscore (1940)	10.54	11.10	9.98
	(1.68)	(1.86)	(1.24)
Average Black Occscore (1940)	8.70	9.01	8.40
	(1.38)	(1.42)	(1.28)
Average Wage Income (1940)	230.63	272.98	188.37
	(113.80)	(121.47)	(87.13)
Average Black Wage Income (1940)	150.48	175.17	125.83
	(76.98)	(74.07)	(71.82)
WWII Spending (per capita) (1940-1945)	154.54	232.60	76.64
	(576.36)	(754.78)	(289.16)
Slave Population Share (1860)	0.35	0.34	0.36
	(0.21)	(0.20)	(0.22)
# of Black Lynchings (1882-1930)	2.12	1.84	2.40
	(3.56)	(3.37)	(3.72)
Net Black Out-Migration Rate	23.19	14.72	31.64
	(10.81)	(5.79)	(7.54)
N	1037	518	519

Note: Column 1 reports average characteristics for all the Southern counties in our main sample. Column 2 reports average characteristics for counties in our sample with low levels of Black out-migration during the 1940s. Column 3 reports the same characteristics for counties in our sample with high levels of Black out-migration. We define a high out-migration county as a county with a net Black out-migration rate at or above the median, and a low out-migration county is any county below the median. Data on county and farm characteristics comes from the 1947, 1952, and 1956 County Data Books from Haines (2010). Slave population data comes from Manson et al. (2024). County-level lynching data comes from Hines and Steelwater (1882–1930).

Table A.4: Comparison of Black Migrants and Non-Migrants, 1940

	(1) Non-Migrants	(2) Migrants	(3) Difference	(4) Adjusted Difference
Male	0.67	0.67	0.00***	-
	(0.47)	(0.47)	(0.00)	
Age	29.75	25.47	4.28***	-
	(16.40)	(13.42)	(0.03)	
Married	0.61	0.57	0.04***	-0.05***
	(0.49)	(0.49)	(0.00)	(0.00)
Employed	0.74	0.69	0.05***	0.01***
	(0.44)	(0.46)	(0.00)	(0.00)
Years of Education	4.41	4.83	-0.42***	-0.61***
	(3.21)	(3.28)	(0.01)	(0.01)
Occscore	8.89	8.17	0.72***	-0.73***
	(8.92)	(8.68)	(0.02)	(0.03)
Wage Income	129.04	118.09	10.95***	-20.28***
	(256.25)	(236.54)	(0.50)	(1.00)
Farm Owners	0.05	0.02	0.03***	0.02***
	(0.22)	(0.13)	(0.00)	(0.00)
Tenant Farmer	0.13	0.14	-0.00***	-0.01***
	(0.34)	(0.35)	(0.00)	(0.00)
Wage Farm Laborer	0.09	0.10	-0.01***	-0.02***
	(0.28)	(0.30)	(0.00)	(0.00)
Laborer	0.09	0.09	-0.00**	-0.02***
	(0.29)	(0.29)	(0.00)	(0.00)
Service Worker	0.18	0.17	0.01***	-0.03***
	(0.39)	(0.38)	(0.00)	(0.00)
Other Worker	0.18	0.17	0.01***	-0.01***
	(0.39)	(0.38)	(0.00)	(0.00)
Observations	1198960	355568	1554528	1554528

Note: This table uses our full sample of Black Southerners, those with at least one of the following: (1) non-zero wage income in 1940, (2) non-zero occscore in 1940, (3) in the labor force in 1940, or (4) in the labor force in 1950. Column 1 reports our sample of Black non-migrants' average demographics and occupational choices in 1940. Column 2 reports averages in the same characteristics for Black migrants. Column 3 reports the difference in characteristics between column 2 and column 3. Column 4 reports the adjusted difference in mean characteristics between column 2 and column 3. We define the adjusted difference as the difference remaining after controlling for age, gender, and county of residence fixed effects in 1940.

Table A.5: Comparison of White Migrants and Non-Migrants, 1940

	(1) Non-Migrants	(2) Migrants	(3) Difference	(4) Adjusted Difference
Male	0.76	0.76	-0.00*	-
	(0.43)	(0.43)	(0.00)	
Age	32.01	27.06	4.94***	-
	(15.99)	(13.67)	(0.01)	
Married	0.63	0.55	0.08***	-0.02***
	(0.48)	(0.50)	(0.00)	(0.00)
Employed	0.75	0.68	0.07***	0.03***
	(0.43)	(0.47)	(0.00)	(0.00)
Years of Education	7.44	7.83	-0.39***	-0.66***
	(3.85)	(3.98)	(0.00)	(0.02)
Occscore	14.45	12.89	1.57***	-0.78***
	(13.14)	(13.07)	(0.01)	(0.03)
Wage Income	360.71	361.06	-0.36	-79.72***
	(660.15)	(657.34)	(0.63)	(2.55)
Farm Owner	0.11	0.04	0.07***	0.05***
	(0.31)	(0.19)	(0.00)	(0.00)
Tenant Farmer	0.08	0.08	-0.00	-0.00***
	(0.27)	(0.27)	(0.00)	(0.00)
Wage Farm Laborer	0.03	0.05	-0.01***	-0.01***
	(0.18)	(0.21)	(0.00)	(0.00)
Laborer	0.05	0.06	-0.01***	-0.01***
	(0.22)	(0.24)	(0.00)	(0.00)
Service Worker	0.14	0.13	0.01***	-0.01***
	(0.35)	(0.34)	(0.00)	(0.00)
Other Worker	0.32	0.32	0.00***	-0.04***
	(0.47)	(0.47)	(0.00)	(0.00)
Observations	4064162	1646636	5710798	5710798

Note: This table uses our full sample of White Southerners, those with at least one of the following: (1) non-zero wage income in 1940, (2) non-zero occscore in 1940, (3) in the labor force in 1940, or (4) in the labor force in 1950. Column 1 reports our sample of White non-migrants' average demographics and occupational choices in 1940. Column 2 reports averages in the same characteristics for White migrants. Column 3 reports the difference in characteristics between column 2 and column 3. Column 4 reports the adjusted difference in mean characteristics between column 2 and column 3. We define the adjusted difference as the difference remaining after controlling for age, gender, and county of residence fixed effects in 1940.

Table A.6: Main Results for Black Stayers - Actual Growth vs Predicted Growth

	Change in Income	Change in Income
	(1)	(2)
Net Black out-migration rate	0.0587***	0.0417***
	(0.0160)	(0.0148)
Individual Observations	445,410	445,410
Number of Shocks	493	493
Individual Controls	No	No
Effective F-statistic for IV	9.0	16.1
Shock	Actual	Predicted

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.7: Pretrend Test: Occupation Score for Black and White Non-Migrants, 1930-1940

Panel A: Black Non-Migrants

	Change in Occscore		Change in ln(Occscore)	
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0074***	-0.0013	-0.0049***	-0.0022
	(0.0026)	(0.0017)	(0.0019)	(0.0015)
Individual Observations	774,495	774,495	625,705	625,705
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.10	0.10	0.13	0.13
Effective F-statistic for IV	16.9	16.9	17.1	17.1

Panel B: White Non-Migrants

	Change in Occscore		Change in ln(Occscore)	
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0018	0.0010	-0.0012**	0.0000
	(0.0012)	(0.0009)	(0.0006)	(0.0005)
Individual Observations	2,434,569	2,434,569	2,093,940	2,093,940
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.17	0.17	0.14	0.14
Effective F-statistic for IV	14.5	14.4	14.4	14.3

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1930 and 1940 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Individual-level control variables are measured in 1940, and are age, marital status, employment status, and gender. *** indicates significance at the 1% level.

Table A.8: Pretrend Test: Employment for Black and White Non-Migrants, 1930-1940

Panel A: Black Non-Migrants

	Change in Employment Status		Change in Labor Force Participation	
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0035***	-0.0008	-0.0024*	0.0025
	(0.0005)	(0.0007)	(0.0013)	(0.0016)
Individual Observations	623,143	623,143	851,179	851,179
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	-0.01	-0.01	-0.04	-0.04
Effective F-statistic for IV	17.1	17.1	17.1	17.1

Panel B: White Non-Migrants

	Change in Employment Status		Change in Labor Force Participation	
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0030***	-0.0003	-0.0030***	0.0027***
	(0.0006)	(0.0004)	(0.0005)	(0.0006)
Individual Observations	2,109,024	2,109,024	2,792,413	2,792,413
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.00	0.00	-0.00	-0.00
Effective F-statistic for IV	14.4	14.4	14.7	14.7

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1930 and 1940 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Individual-level control variables are measured in 1940, and are age, marital status, employment status, and gender. *** indicates significance at the 1% level.

Table A.9: Pretrend Test: County Migration Flows

	1930-1940 Black Net-Migration Rate
	(1)
Shock ($\tilde{g}_c$)	10.0574
	(14.1506)
Observations	493

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variable is the 1930-1940 net-migration rate. The county-level net out-migration rate is from Gardner and Cohen (1992). *** indicates significance at the 1% level.

Table A.10: Log Results: Black and White Non-Migrants

Panel A: Black Non-Migrants

	Change in Log Occscore	Change in Log Occscore	Change in Log Income	Change in Log Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0040***	0.0021*	0.0254***	0.0377***
	(0.0014)	(0.0012)	(0.0094)	(0.0101)
Individual Observations	563,345	563,345	62,935	62,935
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.17	0.17	0.94	0.94
Effective F-statistic for IV	16.5	16.5	14.3	14.4

Panel B: White Non-Migrants

	Change in Log Occscore	Change in Log Occscore	Change in Log Income	Change in Log Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0018***	0.0035***	0.0240***	0.0255***
	(0.0005)	(0.0007)	(0.0032)	(0.0039)
Individual Observations	2,160,569	2,160,569	224,691	224,691
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.14	0.14	0.49	0.49
Effective F-statistic for IV	14.2	14.1	13.2	13.0

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Column 2 controls for age, gender, and martial status in 1940. Columns 4 controls for age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table A.11: Linear Change Results: Black and White Non-Migrants

Panel A: Black Non-Migrants

	Linear Change in Income	Linear Change in Income
	(1)	(2)
Net Black out-migration rate	1.8555	4.3798**
	(2.1280)	(1.9411)
Individual Observations	1,102,410	1,102,410
Number of Shocks	493	493
Individual Controls	No	Yes
Dependent Mean	-15.33	-15.33
Effective F-statistic for IV	17.7	17.7

Panel B: White Non-Migrants

	Linear Change in Income	Linear Change in Income
	(1)	(2)
Net Black out-migration rate	11.4618***	10.3559***
	(1.7711)	(1.4593)
Individual Observations	3,726,635	3,726,635
Number of Shocks	493	493
Individual Controls	No	Yes
Dependent Mean	-223.37	-223.37
Effective F-statistic for IV	13.7	13.5

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are linear changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Column 2 controls for age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table A.12: Comparison of Income Gains: Black Migrants and Non-Migrants, 1940

	(1) Non-Migrants	(2) Migrants	(3) Difference	(4) Adjusted Difference
Male	0.69	0.74	-0.05***	-
	(0.46)	(0.44)	(0.00)	
Age (1940)	35.73	30.79	4.94***	-
	(12.52)	(10.58)	(0.04)	
Married (1940)	0.81	0.79	0.02***	-0.04***
	(0.39)	(0.41)	(0.00)	(0.00)
Years of Education (1940)	5.18	5.64	-0.45***	-0.34***
	(3.35)	(3.44)	(0.01)	(0.02)
Occscore (1940)	14.59	14.49	0.10***	-0.13***
	(8.10)	(7.92)	(0.03)	(0.04)
Wage Income (1940)	319.22	302.29	16.93***	-9.36***
	(318.77)	(295.34)	(1.00)	(1.59)
Change in Wage Income (1940-1950)	-0.04	0.19	-0.23***	-0.07**
	(8.17)	(8.56)	(0.03)	(0.03)
Change in Log Wage Income (1940-1950)	0.94	1.23	-0.29***	-0.18***
	(1.23)	(1.22)	(0.01)	(0.01)
Change in Occscore (1940-1950)	-0.02	0.14	-0.16***	-0.06***
	(0.95)	(1.09)	(0.00)	(0.01)
Observations	445410	126874	572284	572284

Note: This table uses our sample of Black Southerners with non-zero wage income in 1940. Column 1 reports our sample of Black non-migrants' average demographics in 1940 and income gains between 1940 and 1950. Column 2 reports averages in the same characteristics for Black migrants. Column 3 reports the difference in characteristics between column 2 and column 3. Column 4 reports the adjusted difference in mean characteristics between column 2 and column 3. We define the adjusted difference as the difference remaining after controlling for age, gender, and county of residence fixed effects in 1940.

Table A.13: Comparison of Income Gains: White Migrants and Non-Migrants, 1940

	(1) Non-Migrants	(2) Migrants	(3) Difference	(4) Adjusted Difference
Male	0.80	0.85	-0.05***	-
	(0.40)	(0.36)	(0.00)	
Age (1940)	36.17	32.21	3.97***	-
	(12.22)	(10.81)	(0.02)	
Married (1940)	0.79	0.76	0.04***	-0.02***
	(0.40)	(0.43)	(0.00)	(0.00)
Years of Education (1940)	8.46	9.15	-0.69***	-0.58***
	(3.69)	(3.80)	(0.01)	(0.02)
Occscore (1940)	23.05	22.65	0.41***	-0.38***
	(9.32)	(9.83)	(0.01)	(0.04)
Wage Income (1940)	823.37	798.60	24.76***	-53.06***
	(782.20)	(777.94)	(1.13)	(3.78)
Change in Wage Income (1940-1950)	-0.49	-0.22	-0.27***	-0.14***
	(4.38)	(5.85)	(0.01)	(0.01)
Change in Log Wage Income (1940-1950)	0.49	0.74	-0.25***	-0.08***
	(1.18)	(1.29)	(0.00)	(0.01)
Change in Occscore (1940-1950)	-0.05	0.06	-0.11***	-0.01
	(0.81)	(0.97)	(0.00)	(0.00)
Observations	1629999	682869	2312868	2312868

Note: This table uses our sample of White Southerners with non-zero wage income in 1940. Column 1 reports our sample of Black non-migrants' average demographics in 1940 and income gains between 1940 and 1950. Column 2 reports averages in the same characteristics for Black migrants. Column 3 reports the difference in characteristics between column 2 and column 3. Column 4 reports the adjusted difference in mean characteristics between column 2 and column 3. We define the adjusted difference as the difference remaining after controlling for age, gender, and county of residence fixed effects in 1940.

Table A.14: Heterogeneity by Gender: Black Non-Migrants

	Men		Women	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0378***	0.0500***	0.0664	0.0677
	(0.0074)	(0.0083)	(0.0429)	(0.0426)
Individual Observations	309,003	309,003	136,407	136,407
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	-0.23	-0.23	0.40	0.40
Effective F-statistic for IV	16.7	16.9	14.4	14.3

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. Columns 2 and 4 control for occscore, age, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.15: Heterogeneity by Gender: White Non-Migrants

	Men		Women	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0198***	0.0217***	0.0164***	0.0138***
	(0.0033)	(0.0038)	(0.0051)	(0.0051)
Individual Observations	1,299,867	1,299,867	330,132	330,132
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	-0.48	-0.48	-0.52	-0.52
Effective F-statistic for IV	13.4	13.2	15.1	14.9

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. Columns 2 and 4 control for occscore, age, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.16: Heterogeneity by 1940 Income: Black Non-Migrants

	Low Income		High Income	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0447***	0.0686***	-0.0005	0.0035**
	(0.0152)	(0.0169)	(0.0019)	(0.0014)
Individual Observations	220,760	220,760	224,650	224,650
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.59	0.59	-0.65	-0.65
Effective F-statistic for IV	22.7	22.6	8.9	8.7

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. Columns 2 and 4 control for gender, occscore, age, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.17: Heterogeneity by 1940 Income: White Non-Migrants

	Low Income		High Income	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0215***	0.0232***	-0.0020**	-0.0008
	(0.0046)	(0.0051)	(0.0008)	(0.0006)
Individual Observations	717,268	717,268	719,184	719,184
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	-0.25	-0.25	-0.80	-0.80
Effective F-statistic for IV	9.7	9.4	17.4	17.2

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. Columns 2 and 4 control for gender, occscore, age, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.18: Change in Income - Stratified by Non-Lynching and Lynching Counties

	Black Non-Migrants		White Non-Migrants	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0461***	0.0582***	0.0184***	0.0213***
	(0.0130)	(0.0217)	(0.0029)	(0.0070)
Individual Observations	150,361	295,049	799,453	830,546
Number of Shocks	405	442	405	442
Individual Controls	Yes	Yes	Yes	Yes
Effective F-statistic for IV	8.7	8.2	18.0	7.6
Lynching County	No	Yes	No	Yes

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Controls for age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table A.19: Change in Income - Stratified by Low and High George Wallace Vote Share

	Black Non-Migrants		White Non-Migrants	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0392***	0.0690***	0.0169***	0.0232***
	(0.0104)	(0.0215)	(0.0033)	(0.0045)
Individual Observations	126,998	318,412	808,836	821,163
Number of Shocks	403	440	403	440
Individual Controls	Yes	Yes	Yes	Yes
Effective F-statistic for IV	22.9	6.1	20.7	5.8
Sample	Low Vote Share	High Vote Share	Low Vote Share	High Vote Share

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Controls for age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table A.20: Occscore Heterogeneity by Gender

Panel A: Black Non-Migrants

	Men		Women	
	Change in Occscore	Change in Occscore	Change in Occscore	Change in Occscore
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0094***	0.0030	0.0144**	0.0205***
	(0.0030)	(0.0027)	(0.0061)	(0.0064)
Individual Observations	305,647	305,647	123,120	123,120
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.07	0.07	-0.25	-0.25
Effective F-statistic for IV	16.7	17.1	13.5	13.6

Panel B: White Non-Migrants

	Men		Women	
	Change in Occscore	Change in Occscore	Change in Occscore	Change in Occscore
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0019	0.0065***	0.0034***	0.0010
	(0.0018)	(0.0020)	(0.0013)	(0.0010)
Individual Observations	1,284,773	1,284,773	303,695	303,695
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.03	0.03	-0.37	-0.37
Effective F-statistic for IV	13.4	13.2	15.0	14.6

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. Outcomes are restricted to those with a non-missing value for our main outcome, change in wage income between 1940 and 1950. Columns 2 and 4 control for years of education, age, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.21: Occupation Heterogeneity by Gender - Black Non-Migrants

Panel A: Men

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	-0.0073***	0.0032***	0.0040	-0.0025***	-0.0008
	(0.0023)	(0.0012)	(0.0027)	(0.0006)	(0.0013)
Individual Observations	307,341	307,341	307,341	307,341	307,341
Number of Shocks	493	493	493	493	493
Individual Controls	Yes	Yes	Yes	Yes	Yes
Dependent Mean	0.08	-0.16	-0.14	0.03	0.08
Effective F-statistic for IV	17.4	17.4	17.4	17.4	17.4

Panel B: Women

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	-0.0004**	0.0106**	0.0013	0.0023	-0.0193***
	(0.0002)	(0.0044)	(0.0010)	(0.0047)	(0.0058)
Individual Observations	123,256	123,256	123,256	123,256	123,256
Number of Shocks	493	493	493	493	493
Individual Controls	Yes	Yes	Yes	Yes	Yes
Dependent Mean	0.00	-0.07	-0.02	-0.22	-0.06
Effective F-statistic for IV	13.6	13.6	13.6	13.6	13.6

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1940 and 1950 using full count census data. Outcomes are restricted to those with a non-missing value for our main outcome, change in wage income between 1940 and 1950. Controls for age, years of education, occscore, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.22: Federal Minimum Wage in 1950

	Black Non-Migrants		White Non-Migrants	
	Less Than Minimum Wage	Less Than Minimum Wage	Less Than Minimum Wage	Less Than Minimum Wage
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0036	0.0038	0.0090***	0.0077***
	(0.0025)	(0.0024)	(0.0016)	(0.0015)
Individual Observations	144,237	144,237	477,541	477,541
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.43	0.43	0.27	0.27
Effective F-statistic for IV	16.8	17.1	14.0	13.3

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. Columns 2 and 4 control for gender, occscore, age, income, and martial status in 1940. The main outcome is an indicator identifying whether an individual with positive wage income in the 1950 census received wage income less than the federal minimum wage in the year prior. The federal minimum wage in 1949 was $0.40 per hour, making the annual minimum wage $800 for someone working 40 hours per week for 50 weeks. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.23: Changes in Occupation with Controls : Black and White Non-Migrants

Panel A: Black Non-Migrants

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	-0.0070***	0.0085***	0.0046**	-0.0031**	-0.0051**
	(0.0017)	(0.0020)	(0.0021)	(0.0013)	(0.0020)
Individual Observations	430,597	430,597	430,597	430,597	430,597
Number of Shocks	493	493	493	493	493
Individual Controls	Yes	Yes	Yes	Yes	Yes
Dependent Mean	0.06	-0.13	-0.11	-0.04	0.04
Effective F-statistic for IV	15.9	15.9	15.9	15.9	15.9

Panel B: White Non-Migrants

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	0.0018**	-0.0013***	-0.0035***	0.0130***	-0.0123***
	(0.0009)	(0.0005)	(0.0011)	(0.0029)	(0.0030)
Individual Observations	1,597,535	1,597,535	1,597,535	1,597,535	1,597,535
Number of Shocks	493	493	493	493	493
Individual Controls	Yes	Yes	Yes	Yes	Yes
Dependent Mean	0.03	-0.05	-0.09	-0.08	0.04
Effective F-statistic for IV	13.7	13.7	13.7	13.7	13.7

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1940 and 1950 using full count census data. Outcomes are restricted to those with a non-missing value for our main outcome, change in wage income between 1940 and 1950. Controls for age, gender, occscore, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.24: Occupation Heterogeneity by Gender - White Non-Migrants

Panel A: Men

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	0.0020**	-0.0039***	-0.0051***	0.0104***	-0.0058***
	(0.0010)	(0.0008)	(0.0013)	(0.0020)	(0.0021)
Individual Observations	1,292,889	1,292,889	1,292,889	1,292,889	1,292,889
Number of Shocks	493	493	493	493	493
Individual Controls	Yes	Yes	Yes	Yes	Yes
Dependent Mean	0.04	-0.06	-0.10	-0.05	0.07
Effective F-statistic for IV	13.2	13.2	13.2	13.2	13.2

Panel B: Women

	Change in Farmer	Change in Wage Farm Laborer	Change in Laborer	Change in Service Worker	Change in Other Worker
	(1)	(2)	(3)	(4)	(5)
Net Black out-migration rate	-0.0002	-0.0003	-0.0004*	0.0189***	-0.0242***
	(0.0001)	(0.0003)	(0.0002)	(0.0047)	(0.0059)
Individual Observations	304,646	304,646	304,646	304,646	304,646
Number of Shocks	493	493	493	493	493
Individual Controls	Yes	Yes	Yes	Yes	Yes
Dependent Mean	-0.00	-0.01	-0.01	-0.24	-0.06
Effective F-statistic for IV	14.5	14.5	14.5	14.5	14.5

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1940 and 1950 using full count census data. Outcomes are restricted to those with a non-missing value for our main outcome, change in wage income between 1940 and 1950. Controls for age, occscore, years of education, and martial status in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.25: Change in Employment: Black and White Non-Migrants

Panel A: Black Non-Migrants

	Change in Labor Force Participation	Change in Labor Force Participation	Change in Employment	Change in Employment
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0086***	-0.0010	-0.0070***	0.0008
	(0.0017)	(0.0018)	(0.0025)	(0.0027)
Individual Observations	950,920	950,920	917,535	917,535
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.01	0.01	0.03	0.03
Effective F-statistic for IV	16.9	16.9	16.9	16.9

Panel B: White Non-Migrants

	Change in Labor Force Participation	Change in Labor Force Participation	Change in Employment	Change in Employment
	(1)	(2)	(3)	(4)
Net Black out-migration rate	-0.0055***	-0.0031***	-0.0036***	-0.0011
	(0.0010)	(0.0010)	(0.0007)	(0.0008)
Individual Observations	3,464,684	3,464,684	3,382,102	3,382,102
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	0.05	0.05	0.08	0.08
Effective F-statistic for IV	14.7	14.5	14.7	14.5

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Column 2 and 4 control for age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table A.26: Change in Income - With Inverse Probability Weights

	Black Non-Migrants		**White Non-Migrants**	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0417***	0.0445***	0.0193***	0.0230***
	(0.0148)	(0.0135)	(0.0032)	(0.0036)
Individual Observations	445,410	445,410	1,629,999	1,629,999
Number of Shocks	493	493	493	493
Individual Controls	No	No	No	No
Weights	No	Yes	No	Yes
Dependent Mean	-0.04	-0.04	-0.49	-0.49
Effective F-statistic for IV	16.1	15.3	13.9	13.6

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.28: Main Results: Excluding Workers with In-Kind Payments

	Black Non-Migrants		**White Non-Migrants**	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0369***	0.0562***	0.0183***	0.0189***
	(0.0129)	(0.0152)	(0.0028)	(0.0032)
Individual Observations	388,561	388,561	1,328,864	1,328,864
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	-0.08	-0.08	-0.50	-0.50
Effective F-statistic for IV	16.0	16.0	13.4	13.2

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. Columns 2 and 4 control for age, gender, martial status, and occscore in 1940. *** indicates significance at the 1% level.

Table A.27: Main Results: Non-Farmers

Panel A: Black Non-Migrants

	Change in Income	Change in Income	Change in Income (< $50 of Non-Wage Income)	Change in Income (<$50 of Non-Wage Income)
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0346**	0.0545***	0.0351***	0.0549***
	(0.0135)	(0.0161)	(0.0132)	(0.0156)
Individual Observations	417,645	417,645	375,345	375,345
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	-0.07	-0.07	-0.08	-0.08
Effective F-statistic for IV	15.1	15.1	15.3	15.3

Panel B: White Non-Migrants

	Change in Income	Change in Income	Change in Income (< $50 of Non-Wage Income)	Change in Income (<$50 of Non-Wage Income)
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0163***	0.0179***	0.0174***	0.0181***
	(0.0024)	(0.0030)	(0.0025)	(0.0030)
Individual Observations	1,527,382	1,527,382	1,283,461	1,283,461
Number of Shocks	493	493	493	493
Individual Controls	No	Yes	No	Yes
Dependent Mean	-0.51	-0.51	-0.50	-0.50
Effective F-statistic for IV	13.6	13.5	13.1	12.9

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. Columns 2 and 4 control for age, gender, martial status, and occscore in 1940. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.29: Change in Income - Stratified by Predicted Low and High Black Out-Migration

	Black Non-Migrants		**White Non-Migrants**	
	Change in Income	Change in Income	Change in Income	Change in Income
	(1)	(2)	(3)	(4)
Net Black out-migration rate	0.0430	0.0339**	0.0207***	0.0218*
	(0.0310)	(0.0145)	(0.0035)	(0.0115)
Individual Observations	320,731	124,679	1,170,644	459,355
Number of Shocks	465	382	465	382
Individual Controls	No	No	No	No
Effective F-statistic for IV	3.4	20.3	11.3	4.6
Sample	Predicted Low	Predicted High	Predicted Low	Predicted High

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variables are decimal changes measured at the individual level between 1940 and 1950 using full count census data. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

Table A.30: Selection: Model Simulation

	Stayer Skill
	(1)
Net Black out-migration rate	-0.0045***
	(0.0007)
Individual Observations	704,957
Number of Shocks	481
Individual Controls	No
Effective F-statistic for IV	16.0

Note: Regressions at the shock (city) level, for Northern cities that received Black out-migrants from the South in 1940. The dependent variable is the simulated values of $\Delta_t \epsilon_0$ from the Roy model described above. The county-level net out-migration rate is from Boustan (2016), instrumented for using a shift-share instrument. *** indicates significance at the 1% level.

A.3 Appendix Materials

A.3.1 Labor Model Derivation

$$
\begin{aligned}
\frac{\partial \log \omega_B}{\partial L_B} &= \frac{\Big((1-\alpha)(1+\sigma)-\beta\Big)\theta_B L_B^{\beta-1}}{\theta_B L_B^{\beta}+\theta_W L_W^{\beta}} - \frac{(1-\beta)}{L_B} \\
&= \frac{\Big((1-\alpha)(1+\sigma)-\beta\Big)\theta_B L_B^{\beta}}{L^{\beta} L_B} - \frac{(1-\beta)L^{\beta}}{L^{\beta} L_B} \\
&= \frac{\Big((1-\alpha)+\sigma(1-\alpha)-\beta\Big)\theta_B L_B^{\beta} - (1-\beta)[\theta_B L_B^{\beta}+\theta_W L_W^{\beta}]}{L^{\beta} L_B} \\
&= \frac{\Big(\sigma-\alpha(1+\sigma)\Big)\theta_B L_B^{\beta} - (1-\beta)(\theta_W L_W^{\beta})}{L^{\beta} L_B} \\
&< 0 \text{ iff } \alpha(1+\sigma) > \sigma, \text{ since } \alpha > 0, \beta \leq 1.
\end{aligned}
$$

For the usual range of these parameters, including the entire range $\alpha \in (\frac{1}{3}, 1)$, $\sigma \in (0, 0.2)$, $\alpha(1+\sigma) > \sigma$ so this condition holds and the derivative will be negative.

A.3.2 Potential Outcomes Formalization of Endogeneity Issues

Assuming a binary treatment D_i (high amounts of out-migration), an individual's realized outcome Y_i can be written in terms of potential outcomes:

$$Y_i = \begin{cases} Y_{1i} \text{ if } D_i = 1 \\ Y_{0i} \text{ if } D_i = 0 \end{cases}$$

In our setting, each individual also has a binary indicator S_i for whether they stay in their county of origin (i.e., whether they migrate or not).

The observed difference in outcomes among stayers, based on treatment status, is:

$$E[Y_i|D_i = 1, S_i = 1] - E[Y_i|D_i = 0, S_i = 1]$$

This observed difference, based on treatment status, among stayers can be decomposed:

$$\begin{aligned} &E[Y_i|D_i = 1, S_i = 1] \ - E[Y_i|D_i = 0, S_i = 1] = \\ &E[Y_{1i}|D_i = 1, S_i = 1] - E[Y_{0i}|D_i = 1, S_i = 1] + \qquad (14) \\ &E[Y_{0i}|D_i = 1, S_i = 1] - E[Y_{0i}|D_i = 0, S_i = 1] \qquad (15) \end{aligned}$$

Examining the decomposition, we find that the expression in equation 14 is the average effect of treatment (in our case, high levels of local out-migration) among the treated stayers, while the expression in equation 15 represents selection into treatment among stayers. Two issues can lead to the expression in equation 15 being non-zero.

First, there could be a correlation between treatment and the economic conditions that lead to out-migration. That is, if areas with $D_i = 1$ and $D_i = 0$ are different, then:

$$E[Y_{0i}|D_i = 1, S_i = 1] \neq E[Y_{0i}|D_i = 0, S_i = 1]$$

By providing quasi-random variation in D_i, our shift-share instrument helps us address concerns that $E[\Delta\omega_{B0i}|D_i = 1, S_i = 1] \neq E[\Delta\omega_{B0i}|D_i = 0, S_i = 1]$ because of changes in these counties' economic conditions that affected wages and caused out-migration.

Second, stayers in areas of high and low out-migration will be different if migration is selected (i.e., if the distribution of Y_{0i} is different for different values of S_i), again leading to $E[Y_{0i}|D_i = 1, S_i = 1] \neq E[Y_{0i}|D_i = 0, S_i = 1]$.[41] If stayers have higher ability in counties with low out-migration than in counties with high out-migration due to selection, and ability influences growth in wages, then $E[\Delta\omega_{B0i}|D_i = 1, S_i = 1] < E[\Delta\omega_{0i}|D_i = 0, S_i = 1]$, which would bias our estimates of the effect of out-migration on the income of those who do not migrate downward.

[41] Selection into treatment among the stayers can be further decomposed:

$$\begin{aligned} &E[Y_{0i}|D_i = 1, S_i = 1] - E[Y_{0i}|D_i = 0, S_i = 1] \ = \\ &E[Y_{0i}|D_i = 1, S_i = 1] - E[Y_{0i}|D_i = 1, S_i = 0] + \qquad (16) \\ &E[Y_{0i}|D_i = 1, S_i = 0] - E[Y_{0i}|D_i = 0, S_i = 0] + \qquad (17) \\ &E[Y_{0i}|D_i = 0, S_i = 0] - E[Y_{0i}|D_i = 0, S_i = 1] \qquad (18) \end{aligned}$$

where the expression in equation 16 is the difference in potential outcomes between movers and stayers among the treated, the expression in equation 17 represents selection into treatment among movers, and the expression in equation 18 represents the difference in potential outcomes between movers and stayers among the untreated.

A.3.3 Model with Worker Heterogeneity

Recall that in Section 3, overall labor supply is given by combining the labor of Black and White workers into a CES aggregate:

$$L = [\theta_B L_B^{\beta} + \theta_W L_W^{\beta}]^{\frac{1}{\beta}} \tag{19}$$

where $\theta_W + \theta_B = 1$ and $\beta \leq 1$.

Here, we introduce heterogeneity between workers of a given race by allowing the supply of Black and White labor to be divided by skill, inspired by Borjas (2003). We assume there are skilled workers, s, and unskilled workers, u, within each racial group. Workers of different skill levels are imperfect substitutes within each racial group, and their labor is combined using an additional CES aggregation:

$$L_B = [\alpha_{Bs} L_{Bs}^{\rho} + \alpha_{Bu} L_{Bu}^{\rho}]^{\frac{1}{\rho}}$$
$$L_W = [\alpha_{Ws} L_{Ws}^{\rho} + \alpha_{Wu} L_{Wu}^{\rho}]^{\frac{1}{\rho}}$$

where $\alpha_{Bs} + \alpha_{Bu} = 1$, $\alpha_{Ws} + \alpha_{Wu} = 1$, and $\rho \leq 1$.

At the county level, average wages for Black non-migrants now depend on the mixture of skill levels that remains in the county, with weights given by η:

$$\frac{1}{N} \sum_i \log \omega_B = \eta \log \omega_{Bs} + (1 - \eta) \log \omega_{Bu}$$

Therefore, selected out-migration can change average county-level wages through both the direct effect on labor supply, $\frac{\partial \log \omega_B}{\partial L_B}$, and through a compositional effect that out-migration imposes on the skill distribution of those who do not migrate, $\frac{\partial \eta}{\partial L_B}$.

A.3.4 Theory: Roy Model of Selection among Non-Migrants

Recall from Section 5.1 that we are concerned that non-migrants in areas of high and low out-migration will be different if migration is selected. Using the potential outcomes notation from Section A.3.2, where $S_i = 1$ denotes non-migrants and $D_i = 1$ refers to an area of high out-migration, our concern is that migrant selection leads to differences across treatment status in average counterfactual outcomes:

$$\mathbb{E}[Y_{0i}|D_i = 1, S_i = 1] \neq \mathbb{E}[Y_{0i}|D_i = 0, S_i = 1]$$

In this section, we follow the setup from Borjas (1987), inspired by Roy (1951), to demonstrate that selection into migration can induce downwards bias into our estimates of the effect of out-migration on the income of those who do not migrate.

Assume that an individual i considers moving from location 0 to location 1, where the locations are indexed by j, with wages in each location given by

$$\ln w_{ij} = \mu_j + \epsilon_{ij} \tag{20}$$

The μ_j terms reflect observable determinants of wages in each location, while the ϵ_{ij} reflect unobservable individual deviations from that average, which can be thought of as the de-meaned value of the worker's skill in that location. Assume that ϵ_{i0} and ϵ_{i1} are both distributed standard normal with respective variances σ_j^2, are jointly normally distributed with covariance $\sigma_{0,1}$, and have correlation coefficient ρ.

Moving costs are given by $C = \pi w_o$. Dropping the i subscripts, an individual moves if:

$$\begin{aligned} \ln(w_1) - \ln(w_0 + C) &> 0 \\ \ln(w_1) - \ln(w_0 + \pi w_0) &> 0 \\ \ln(w_1) - \ln(w_0) - \ln(1 + \pi) &> 0 \\ \mu_1 + \epsilon_1 - \mu_0 - \epsilon_0 - \ln(1 + \pi) &> 0 \\ v = \epsilon_1 - \epsilon_0 &> \mu_0 + \pi - \mu_1 \end{aligned} \tag{21}$$

An individual moves if their wage, net of moving costs, will increase. Counties that have better connections to Northern cities with more manufacturing growth could be modeled as either having lower moving costs (lower π) or increased average destination wages (higher μ_1). In either case, $\mu_0 + \pi - \mu_1$ is lower for these counties, and more people migrate.

We want to know how $\mathbb{E}(\epsilon_0|v < \mu_0 + \pi - \mu_1)$, the average unobserved ability of those who do not migrate, and $\mathbb{E}(\ln w_0|v < \mu_0 + \pi - \mu_1)$, the average wage of those who do not migrate, is different in counties with different levels of $\mu_0 + \pi - \mu_1$ (and thus, in counties with different amounts of out-migration).

Re-writing by standardizing v, we have:

$$\mathbb{E}(\epsilon_0|v < \mu_0 + \pi - \mu_1) = \mathbb{E}\left(\epsilon_0 \middle| \frac{v}{\sigma_v} < \frac{\mu_0 + \pi - \mu_1}{\sigma_v} = z\right) = \sigma_0 \mathbb{E}\left(\frac{\epsilon_0}{\sigma_0} | \frac{v}{\sigma_v} < z\right) \tag{22}$$

Letting $s = \frac{v}{\sigma_v} \sim N(0, 1)$, ϵ_0 and s are jointly normally distributed, [42] so the conditional expectation $\mathbb{E}(\epsilon_0|s)$ makes use of the regression coefficient:

[42] This follows from ϵ_0 and v being jointly normally distributed, which, in turn, follows from ϵ_0 and ϵ_1 being jointly normally distributed.

$$\mathbb{E}\left(\frac{\epsilon_0}{\sigma_0}|\frac{v}{\sigma_v}\right) = \frac{1}{\sigma_0}\mathbb{E}\left(\epsilon_0|s\right) = \frac{1}{\sigma_0}\frac{\text{Cov}(\epsilon_0, s)}{\sigma_s^2}s = \frac{1}{\sigma_0}\frac{\text{Cov}(\epsilon_0, v)/\sigma_v}{1}\frac{v}{\sigma_v} = \rho_{\epsilon_o,v}\frac{v}{\sigma_v} \tag{23}$$

Therefore, making use of the Law of Iterated Expectations, we have:

$$\begin{aligned}
&\mathbb{E}\left(\epsilon_0\middle|\frac{v}{\sigma_v} < z\right) \\
&= \sigma_0\mathbb{E}\left(\frac{\epsilon_0}{\sigma_0}\middle|\frac{v}{\sigma_v} < z\right) \\
&= \sigma_0\mathbb{E}\left[\mathbb{E}\left(\frac{\epsilon_0}{\sigma_0}\middle|\frac{v}{\sigma_v}\right)\middle|\frac{v}{\sigma_v} < z\right] \\
&= \sigma_0\mathbb{E}\left(\rho_{\epsilon_o,v}\frac{v}{\sigma_v}\middle|\frac{v}{\sigma_v} < z\right) \\
&= \sigma_0\rho_{\epsilon_o,v}\mathbb{E}\left(\frac{v}{\sigma_v}\middle|\frac{v}{\sigma_v} < z\right) \\
&= -\sigma_0\rho_{\epsilon_o,v}\frac{\phi(z)}{\Phi(z)}
\end{aligned} \tag{24}$$

Applying this result to average wages in the origin, we have:

$$\mathbb{E}\left(\ln w_0\middle|\frac{v}{\sigma_v} < z\right) = \mu_0 - \sigma_0\rho_{\epsilon_0,v}\frac{\phi(z)}{\Phi(z)} \tag{25}$$

The Inverse Mills Ratio, $\frac{\phi(z)}{\Phi(z)}$, is decreasing in z, but we also need to know the sign of $\sigma_0\rho_{\epsilon_o,v}$. Recalling that ρ is the correlation coefficient between ϵ_0 and ϵ_1, we note that:

$$\begin{aligned}
\sigma_0\rho_{\epsilon_o,v} &= \sigma_0\frac{\text{Cov}(\epsilon_0, v)}{\sigma_0\sigma_v} = \frac{\text{Cov}(\epsilon_0, \epsilon_1) - \text{Var}(\epsilon_0)}{\sigma_v} \\
&= \frac{\sigma_0\sigma_1}{\sigma_v}\left(\frac{\text{Cov}(\epsilon_0, \epsilon_1)}{\sigma_0\sigma_1} - \frac{\sigma_0}{\sigma_1}\right) = \frac{\sigma_0\sigma_1}{\sigma_v}\left(\rho - \frac{\sigma_0}{\sigma_1}\right)
\end{aligned} \tag{26}$$

Therefore, if $\rho > \frac{\sigma_0}{\sigma_1}$, that is, if the correlation between skills in the Southern origin and Northern destination is sufficiently high, then the average unobserved ability of stayers and the average wage of stayers are both lower in counties with lower z (more out-migration):

$$\frac{\partial\,\mathbb{E}(\epsilon_0|v < \mu_0 + \pi - \mu_1 = \sigma_v z)}{\partial z} > 0 \tag{27}$$

$$\frac{\partial\,\mathbb{E}(\ln w_0|v < \mu_0 + \pi - \mu_1 = \sigma_v z)}{\partial z} > 0 \tag{28}$$

$$\text{if } \rho > \frac{\sigma_0}{\sigma_1}$$

Intuitively, the conditions that we show are required for negative selection of stayers are the exact same as the conditions that produce positive selection of immigrants in the model of Borjas (1987). Referring back to the potential outcomes notation from the beginning of this section, we have shown above that $\mathbb{E}[\ln w_{0i}|D_i = 1, S_i = 1] < \mathbb{E}[\ln w_{0i}|D_i = 0, S_i = 1]$.

Note, though, that these selection results pertain to the level of income. Our main specifications use individual-level changes in income between 1940 and 1950. In the original model from Borjas (1987), changes in income would not be affected by selection.[43] We could instead consider a slight modification of the original model, in which:

$$\ln w_{ijt} = \mu_{jt} + \epsilon_{ijt} \tag{29}$$

The outcome we measure for non-migrants who stay in location $j = 0$ is then:

$$\ln w_{i,j=0,t=1} - \ln w_{i,j=0,t=0} = \Delta_t \ln w_{i,0} = (\mu_{0,1} - \mu_{0,0}) + (\epsilon_{i,0,1} - \epsilon_{i,0,0}) \tag{30}$$

As we can see, this outcome is affected both by how one's ability to earn income in Southern counties changes with out-migration, which we are interested in, and individual deviations that may lead to selection bias. If we model $(\epsilon_{i,0,1} - \epsilon_{i,0,0}) = \Delta_t \epsilon_{i,0}$ the same way that we modeled ϵ_{i0} above, and dropping the i subscripts for consistency, our results from above directly apply as long as $\rho > \frac{\sigma_0}{\sigma_1}$:

$$\frac{\partial \mathbb{E}[\Delta_t \epsilon_0 \,|v < \mu_0 + \pi - \mu_1 = \sigma_v z]}{\partial z} > 0$$
$$\frac{\partial \mathbb{E}[\Delta_t \ln w_0 \,|v < \mu_0 + \pi - \mu_1 = \sigma_v z]}{\partial z} > 0$$

In either case, then, this selection effect would bias downwards our results of the effect of out-migration on changes in income for those who do not migrate :

$$\mathbb{E}[\Delta_t \ln w_{i0} \,|D_i = 1, S_i = 1] < \mathbb{E}[\Delta_t \ln w_{i0} \,|D_i = 0, S_i = 1] \tag{31}$$

[43] The original model does not include a time dimension, so it cannot explain changes in income between periods.

A.3.5 Simulation: Roy Model of Selection among Non-Migrants

We simulate the above model under the assumption that changes in wages are only caused by selection (that is, $\Delta_t \mu_0 = 0$). This assumption will allow us to isolate the effect of selection on our estimated regression coefficients. Our outcome of interest is:

$$\ln w_{i,j=0,t=1} - \ln w_{i,j=0,t=0} = \Delta_t \ln w_{i,0} = (\mu_{0,1} - \mu_{0,0}) + (\epsilon_{i,0,1} - \epsilon_{i,0,0}) = \Delta_t \epsilon_{i,0} \tag{32}$$

We assume that $\Delta_t \epsilon_0$ and $\Delta_t \epsilon_1$ are jointly normally distributed, and use the following variance-covariance matrix that allows $\rho > \frac{\sigma_0}{\sigma_1}$:

$$\begin{pmatrix} \Delta_t \epsilon_0 \\ \Delta_t \epsilon_1 \end{pmatrix} \sim \mathcal{N}\left(\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{bmatrix} 0.81 & 0.86 \\ 0.86 & 1 \end{bmatrix} \right)$$

Each individual in our sample of Black Southern residents in 1940 receives draws of $\Delta_t \epsilon_0$ and $\Delta_t \epsilon_1$ from the above distribution. We let moving costs vary by county to match the real number of migrants and stayers recorded for that county in our data. That is, our cutoff for migration, z, varies by county so that in each county the number of individuals for whom $\frac{v}{\sigma_v} < z$ is the number of non-migrating individuals that we observe in that county.

We then run the IV regression, equation 5, that we used to produce our main results of the effect of out-migration on non-migrants, on the sample of individuals with simulated values of $\frac{v = \Delta_t \epsilon_1 - \Delta_t \epsilon_0}{\sigma_v} < z$ (that is, non-migrants). In this regression, our outcome is $\Delta_t \epsilon_0$, which reflects simulated individual value-added in the Southern workforce. In the model, out-migration can only affect $\Delta_t \epsilon_0$ through selection.

In line with the theoretical model, we find that positive selection of migrants (negative selection of stayers) leads to a negative and statistically significant regression coefficient, displayed in Appendix Table A.30. These results provide further evidence for our argument that our estimates of a positive and statistically significant effect of out-migration on changes in income for those who do not migrate would be biased downward by positively-selected migration.

A.3.6 Shock-level Regression

Though the treatment variable x_j and county-city connections s_{jc} vary at the county level, in practice these variables are used at the individual level before being converted to the shock (city) level. The weighted shares and residuals for Northern city c are defined, respectively, below, where residuals ε_i come from the individual-level equation 5:

$$s_c = \sum_i s_{ijc},\ \tilde{\varepsilon}_c = \frac{\sum_i s_{ijc}\varepsilon_i}{\sum_i s_{ijc}} \tag{33}$$

The Stata program *ssaggregate* residualizes x_j and y_i from individual-level equation 5 on the vector of control variables. This gives us the residualized variables $y_i^{\perp}$ and $x_i^{\perp}$. These residualized variables are then converted to the shock level (where they are used in the regressions) by taking an exposure-weighted average:

$$\bar{y}_c^{\perp} = \frac{\sum_i s_{ijc} y_i^{\perp}}{\sum_i s_{ijc}}, \bar{x}_c^{\perp} = \frac{\sum_i s_{ijc} x_i^{\perp}}{\sum_i s_{ijc}} \tag{34}$$