Time to Accumulate: The Great Migration and the Rise of the American South

Dongkyu Yang*

[Most recent version here]
[Online Appendix here]

August 5, 2024

Abstract

The idea that labor scarcity can induce economic development has been long hypothesized at least since Hicks (1932) and Habakkuk (1962), but related evidence is still scarce, especially on non-agricultural development. In this paper, I empirically and quantitatively assess the role of the Second Great Migration (1940-1970) in the subsequent structural change in the American South between 1970 and 2010. Empirical results using shift-share instruments show that out-migration incentivized physical capital investment and capital-augmenting technical change, increasing capital and output per worker in both agriculture and manufacturing at least until 2010. Labor reallocated from agriculture to non-agriculture. I then develop and calibrate a dynamic spatial equilibrium model that allows substitution between factors of production, factor-biased technical change, and Heckscher–Ohlin forces in trade. The quantitative results indicate that the adjustments to the Second Great Migration could have contributed to 7% of the total decrease in agricultural employment between 1940 and 2010 in the South. The contribution analyses suggest that labor-capital substitution played a leading role, with capital-biased technical change and the quasi-Rybczynski effect playing important supplementary roles.

Keywords: Great Migration, Spatial equilibrium, Structural change, Heckscher-Ohlin.
JEL: N32, O11, R12, R23.

---

*Department of Economics, University of Colorado - Boulder. Dongkyu.yang@colorado.edu. I would like to thank my dissertation committee members, Richard Mansfield, Sergey Nigai, Taylor Jaworski, and Wolfgang Keller, for their guidance and support. I received valuable comments and suggestions from Ethan Artuc, Maggie Chen, Yeonha Jung, and Martin Rotemberg. All errors are my own. Please feel free to reach out to me with any comments.

Introduction

Economies undergo structural change from agriculture to non-agriculture as a crucial part of the development process. Given its importance, a large body of literature has pointed out various sources of structural change (Baumol, 1967; Caselli and Coleman II, 2001; Ngai and Pissarides, 2007; Alvarez-Cuadrado et al., 2017). Such an economic transformation is often accompanied by large internal migration, with its impacts felt on the origins and destinations (Tombe and Zhu, 2019; Bryan and Morten, 2019; Derenoncourt, 2022; Lagakos et al., 2023). While the role of economic transformation as a driver of migration is evident (Lewis, 1954; Harris and Todaro, 1970), migration as a source of structural change has received relatively less attention.[1] In this paper, I propose migration-induced labor scarcity and the accompanying capital accumulation as a source of structural change by capitalizing on one of the largest labor reallocation episodes in the United States' history: the Great Migration.

During the Great Migration (1910-1970), millions of Black and White Southerners left the South. By 1970, more than ten million Southern-born individuals lived outside the South. This historical episode is often divided into the first (1910-1930) and the second period (1940-1970), with the latter being much bigger. The Second Great Migration period is also characterized by rapid industrialization and structural change in the South. By 1940, the agricultural share of employment in the South (30%) was almost three times higher than the rest of the country ("the North"). However, the two economies converged by the 1980s in terms of industry employment share. Here, I focus on the second wave and investigate its long-run economic impacts.

The economics literature has long studied the influence of the Great Migration (Kirby, 1983; Boustan, 2010; Bazzi et al., 2023), while others investigated why the American South lagged in economic development and why it later caught up. (Whatley, 1985; Caselli and Coleman II, 2001; Depew et al., 2013). The maturing of the Southern economy has often been pointed out as a contributor to the Great Migration as well (Day, 1967; Grove and Heinicke, 2003; Boustan, 2016). This paper, on the other hand, takes an alternative view and aims to unveil a new link between the Great Migration and the transformation of the Southern economy by focusing on the role of physical capital accumulation.[2]

[1] Compared to the existing studies on labor mobility and structural change, such as Caselli and Coleman II (2001), Alvarez-Cuadrado and Poschke (2011), and Alonso-Carrera and Raurich (2018), I focus on the role of regional out-migration and resulting labor scarcity, rather than industry switching or out-migration from a specific industry.

[2] Contemporary (Raper, 1946) and later research (Peterson and Kislev, 1986; Holley, 2000; Hornbeck and Naidu, 2014) point out the potential role of labor scarcity on agricultural mechanization in the South. Instead, I focus on its role in non-agricultural development and structural change.

I first lay out a simple framework motivated by the theoretical results in macroeconomics literature (Acemoglu, 2002, 2007; Alvarez-Cuadrado et al., 2017) and trade (Rybczynski, 1955; Romalis, 2004). It features two regions, the North and the South, and two industries, agriculture and non-agriculture. Agriculture is assumed to substitute labor and capital more flexibly and is also labor-intensive.

I interpret the Great Migration as a large out-migration shock and take it as the economy-wide changes in capital-to-labor ratio. As labor becomes relatively more scarce, agriculture substitutes now more expensive labor with capital. The increase in capital usage in agriculture also induces technical change biased toward capital, further releasing labor. They are absorbed by local non-agriculture. Hence, the out-migration alone can stimulate labor reallocation across sectors. The above changes are driven by the assumption that the elasticity of substitution between labor and capital ("σ") is higher than one in agriculture but less than one in non-agriculture, consistent with estimates in Herrendorf et al. (2015), Oberfield and Raval (2021), and Caunedo and Keller (2024).

Following labor reallocation, non-agriculture is also incentivized to invest in capital due to the labor-capital complementarity. Nonetheless, capital accumulation may not materialize if the size of the industry is constrained by local demand. The open economy forces operate through a distinct channel: regional differences in factor abundance and industry differences in factor intensity. At least at the onset of the migration, the Southern economy can be characterized as labor abundant compared to the North, and Southern agriculture as labor intensive (Bateman and Weiss, 1981; Whatley, 1985; Wright, 1986). Hence, the decrease in labor endowment would have led to a contraction of agriculture but an increase in non-agriculture production. Such quasi-Rybczynski effect (Romalis, 2004) further allows non-agricultural capital accumulation through an accompanying expansion of non-agricultural production.

How much are these predictions relevant to the actual development of the Southern economy? As a next step, I examine the economic changes in the South between 1940 and 2010, using decadal panel datasets at the county level. Specifically, I estimate year-specific changes in outcomes after 1970, compared to their levels in 1940 and 1950, explained by variations in net out-migration rates.[3] The main explanatory variable, the county-level net out-migration rates between 1940 and 1970, is defined to be a negative value of the net migration rate during the 30-year period, calculated by the number of net out-migrants divided by the 1940 population.[4] It summarizes the

[3] To clarify the directions of the migration flows, I explicitly use the terminology in-migration and out-migration. I also use the terminologies "Blacks" and "Whites" for Americans with mainly African and European heritages.

[4] The baseline results use the total out-migration rates by combining Black and White out-migration. When separated by race, the overall results tend to be bigger in size but less precise for Black out-migration compared to White.

Second Great Migration at the county level. I view this variable as the simplest measure closest to the migration-driven changes in county-level capital-to-labor ratio, the shock used in the simple framework. Henceforth, I simply denote it as the out-migration rate. The estimation strategy compares the outcomes before and after the Second Great Migration between counties that had different levels of out-migration.

The major identification concerns are reverse causality, where the underlying forces in the South could have affected the regional out-migration, and omitted variable bias resulting from other accompanying changes. Hence, I use an instrumental variable strategy with an extensive list of control variables. First, the baseline strategy combines two sources of variation in predetermined shares (migration matrices between 1910 and 1940, separately for Blacks and Whites) and Northern pull factors (OLS in-migration prediction by race between 1940 and 1970) as shifts. The aim here is to isolate the variation in out-migration that is explained by what happened in the Northern destinations, rather than what happened in the Southern origins.[5]

Even though the estimation uses the component of out-migration explained by Northern pull factors, important changes in other dimensions in the South may confound the results. Hence, I include state-by-year and county fixed effects so that estimation relies on variation in relative change between counties within the same state. To further take into account the influences of initial differences within the same state, I include time-interacted values of (1) time-invariant county characteristics, (2) initial agriculture conditions, (3) New Deal investment variables, and (4) trade exposure.[6] In other words, the estimation compares the changes between counties within the same state in the same year with similar pre-Great Migration characteristics, assuming that these counties would have changed the same after 1970 in the absence of the Northern pull factor induced out-migration. I report robustness checks by using alternative shares, alternative shifts, and an alternative standard error, among others.

The baseline results suggest that relative labor scarcity from the Second Great Migration contributed to structural change, capital accumulation, and technology adoption in the South at least until 2010. I document that counties with more out-migration released more agricultural labor,

[5]In the Great Migration setting, Boustan (2010), Derenoncourt (2022), and Bazzi et al. (2023) construct instruments for Northern inflows using Southern push factors. I apply the same strategy in the opposite direction by predicting Southern outflows using the Northern pull factors.

[6]The time-invariant county characteristics include log land area, longitude, latitude, and 1940 values of log population and total farm acres. Agriculture conditions include 1940 values of the share of sharecroppers and acres harvested in cotton, tobacco, corn, and hay, as well as shares of farms in five different farm-size bins. The New Deal controls consist of spending on public work and the Agricultural Adjustment Administration (AAA) program. As a trade exposure, I include a measure of Japanese import penetration during the 1970s and 1980s.

used less farmland, but adopted more tractors and combines. Farm outputs stayed relatively stable. The above changes are consistent with agricultural mechanization from shrinking labor supply from a natural disaster (Hornbeck and Naidu, 2014) or an abrupt change in migration policy (Clemens et al., 2018) during the similar periods in the United States. Such findings also suggest that out-migration could have contributed to the rapid diffusion of tractors in the post-war South (Olmstead and Rhode, 2001), as the relative cost of labor and capital were a key determinant of the agricultural mechanization (Manuelli and Seshadri, 2014).

However, the adjustment to the out-migration did not end in agriculture. Non-agriculture results show that a 100% more out-migration at the county level increased manufacturing employment by 10% between 1970 and 2010, relative to its level in 1940 and 1950. There was an accompanying increase in manufacturing capital spending by 20%. Such changes, in turn, raised manufacturing value added and payroll by 12% and 14%, respectively. A similar development occurred in the local retail and wholesale sectors, with their employment increasing by 5% and 11%. The sales and payroll increased as well. As a result, agricultural employment share decreased while that of manufacturing increased. Although not precise, service employment share modestly increased.

Year-specific estimates further suggest that changes in employment were relatively stable between 1970 and 2010 for agriculture, while the increases in capital continued to grow until 2000 and were maintained at least until 2010 in both sectors. Such patterns can be rationalized by capital-biased technical change, with more efficient capital usage further incentivizing capital investment. As follows, the regions with higher out-migration rates primarily responded with physical capital investment, which might have paradoxically reduced the incentive for human capital accumulation. As a result, the level of education did not experience meaningful relative improvement. This suggests that out-migration provides an additional channel for why regions with a higher initial share of agriculture labor experienced faster growth and convergence, in addition to improvements in the average level of education in the South (Caselli and Coleman II, 2001).

In the final part of the paper, I construct a quantitative model featuring trade and migration by capitalizing on recent developments in dynamic spatial general equilibrium models (Eaton and Kortum, 2002; Caliendo et al., 2019; Kleinman et al., 2023; Fan et al., 2023). The model generalizes the simple framework into multiple periods and realistic geography. It considers two sets of industries, agriculture and non-agriculture, where non-agriculture is further divided into tradable and non-tradable sectors. All industries use two factors of production, labor and capital, with CES production structures. However, they are assumed to have different values of factor substitutability and intensity parameters. I parameterize the model based on the empirical estimates and findings from previous literature.

The baseline counterfactual results report that the Second Great Migration increased the United States' consumption welfare by 0.5% per capita by 1970, with the South experiencing a gain of 2.28%, whereas the North experienced a loss of 1.01%. The decomposition using the welfare effects suggests that factor substitution played the most important role, with trade adjustment and the directed technical change playing supplementary roles. As a response to relative labor scarcity from the Great Migration, the model predicts a decrease in labor but an increase in capital allocated to agriculture. Such adjustments are quantified to have reduced agricultural employment by 2 percentage points by 2010, equivalent to around 7% of the total decreases during this period.

Related literature. This paper extends several dimensions in the economics literature. First, it extends our understanding of the impacts of out-migration and, specifically, of the Great Migration. In terms of migration in general, recent literature identifies the influences of out-migration on origin through labor market upgrading (Akram et al., 2018), output mix adjustments (Lafortune et al., 2015), directed technological change (Andersson et al., 2022; San, 2023), labor/capital substitution (Hornbeck and Naidu, 2014; Clemens et al., 2018; Abramitzky et al., 2023), human capital investment (Theoharides, 2018; Caballero et al., 2023), and trade integration (Egger et al., 2024). Relatedly, this paper proposes a new channel on how out-migration can lead to structural change through the origin's re-optimization of its factor usages. The findings here also support promoting rural out-migration as a policy tool for correcting spatial misallocation of labor and capital that is still prevalent across the world (Caselli and Feyrer, 2007; Hsieh and Klenow, 2009; Adamopoulos and Restuccia, 2014; Gollin et al., 2014).

Economic history literature has long studied the impacts of the Great Migration on migrants themselves or receiving regions (Kirby 1983; Boustan 2010; Collins and Wanamaker 2015; Derenoncourt 2022; Bazzi et al. 2023; see Collins and Wanamaker (2022) for a review), while recent works also pay attention to the impacts on the Southern origin's political economy (Feigenbaum et al., 2020) or racial inequality and segregation (Clay et al., 2020; Chapel and Hung, 2024; Montrose, 2024). Among them, this project is closest to Hornbeck and Naidu (2014), who study the influences of the Great Mississippi Flood of 1927 on subsequent Black out-migration, a part of the First Great Migration, on agricultural mechanization until 1970. This paper, on the other hand, focuses on the influence of second flows on structural change and non-agricultural development in the post-war South.

Findings in this paper add to the structural change literature (Baumol, 1967; Caselli and Coleman II, 2001; Boppart, 2014; Porzio et al., 2022), especially on the studies that focus on the role of physical capital (Acemoglu and Guerrieri, 2008; Alvarez-Cuadrado et al., 2017; Alonso-Carrera and Raurich, 2018; Caunedo and Keller, 2024). I add empirical evidence on the mechanism underlying the structural change process (Michaels et al., 2012; Kline and Moretti, 2014; Fajgelbaum and Redding, 2018; Bustos et al., 2020; Dinkelman et al., Forthcoming) by documenting that labor scarcity could facilitate a "big push" out of labor-intensive equilibrium through capital deepening. Among them, this paper is closest to Bustos et al. (2020), who show that capital accumulation from agricultural productivity improvement led to structural change in Brazil. The major difference is that I emphasize the role of capital demand from the factor substitution channel, while Bustos et al. (2020) highlight the role of capital supply from the agricultural income shock.

Hence, this paper also contributes to the literature on why Southern economic development lagged behind the rest of the United States and later caught up (Whatley, 1985; Caselli and Coleman II, 2001; Grove and Heinicke, 2003; Bleakley, 2007; Depew et al., 2013; Jung, 2020). The results here support the hypothesis that the abundance of labor and the lack of physical capital in the South hampered economic advancement (Bateman and Weiss, 1981). Among them, this paper is closest to Caselli and Coleman II (2001), who study the role of structural change in North-South convergence through a quantitative model. However, while they emphasize the importance of education and human capital, this paper stresses the role of physical capital.

Finally, the quantitative framework developed in this paper contributes to the rapidly growing dynamic spatial equilibrium literature (Eaton and Kortum, 2002; Artuc et al., 2010; Caliendo et al., 2019; Kleinman et al., 2023; Fan et al., 2023; Eckert and Peters, 2023). Compared to the existing quantitative framework that features Hecksher-Ohlin forces (Chor 2010; Caron et al. 2014 and Burstein and Vogel 2017, for example), I incorporate recent innovations in modeling migration (Artuc et al., 2010; Caliendo et al., 2019), capital investment (Kleinman et al., 2023), and structural change (Fan et al., 2023) in realistic geography. The quantitative model can be extended to include multiple factors of production and multiple industries that are distinguished by different factor substitutability and factor intensity.

1 Historical Background

The Great Migration, roughly dated between 1910 and 1970, was one of the largest internal migration episodes in United States history. During this period, approximately six million Blacks left the American South in the pursuit of economic and educational opportunities and escaping oppressive systems symbolized by Jim Crow. Other millions of Southern-born Whites voted with their feet and moved out of the South in pursuit of better living conditions and economic prospects. This White Great Migration even exceeded in the total number of migrants.[7] This migration episode is often divided into the first (1910-1930) and the second flows (1940-1970), with the latter being much bigger. This study focuses on how the Second Great Migration shaped the economic outcomes in the South, with 1940 and 1970 as the start and end periods.

The start of the second wave coincided with the end of the Great Depression and the beginning of the Second World War, where increased labor demand in the North during the war boom and mobilization was unmet by international migration.[8] Instead, the Southern-born population started to migrate into the Northern and Western cities, leaving behind dire economic conditions at home. The migration flow continued after the war and remained at high levels until the 1960s. Two important changes in the U.S. could have contributed to the slowing flows after the 1960s. The Civil Rights Act of 1964 prohibited discrimination in public places and made living conditions in the South less harsh for the Blacks. Another important change was the Immigration and Nationality Act of 1965, also known as the Hart-Celler Act. The act was enacted in 1967 and abolished the quota system, significantly increasing international migration flows thereafter and until today.

Figure 1, Panel A, presents the log Black population in the South (red line) and the North (blue line).[9] The steep convergence between 1940 and 1970 give a hint of the sizes of migration flows between the two regions, where the share of Blacks in the South decreased from 76% to 51% during the 30-year period. There was accompanying out-migration of Whites: by 1940, 11% of Southern-born Whites lived outside the South, while the share increased to 20% by 1970 (Bazzi et al., 2023). However, the flow of migrants plateaued and reversed afterward; many gradually returned, and new migrants entered the South. Panel A reveals that the share of Blacks in the South again increased after 1990. This new migration flow is often dubbed the New Great Migration.

While many Southerners left their origin in pursuit of economic opportunity, the South's economy also matured during the same period. The next three Panels plot time trends on employment share in three major industries by region. Panels A through D highlight that the Southern economy rapidly caught up with the rest of the United States while simultaneously experiencing large out-

[7]Although it is hard to know the exact number of migrants, Gregory (2005) calculates that more than 27 million southerners left the South either permanently or temporarily over the course of the 20th century.

[8]International migration was largely shut down with the Immigration Act of 1924, which limited the number of immigrants allowed entry into the United States through a national origins quota. The Bracero program (1942-1964) is a notable exception during this period.

[9] The definition of the South follows Boustan (2016) by mainly using the Census definition but classifying border states based on their migration history. The Southern states consist of former Confederate states, Kentucky, Oklahoma, and West Virginia, but exclude the District of Colombia, Maryland, and Delaware. The contiguous U.S. states not included in the South is classified as the North.

Figure 1: Great Migration and economic changes in the South.

Note: The figure presents Time trends in the selected variables for the North and the South. The definition of the region is described in Footnote 9. The number of the black population is calculated using Haines et al. (2010). The industry shares are calculated from 1% and 5% samples of the Population Census and American Community Survey (Ruggles et al., 2024b). The industry definition is based on the 1950 Census Bureau industrial classification system. Agriculture consists of agriculture, forestry, and fishing but excludes mining. The agriculture K/L (capital-to-labor) ratio is calculated from Haines et al. (2018). It is defined as the number of tractors and combines, divided by agriculture employment. The manufacturing (ΔK)/L ratio is calculated from HDES using manufacturing capital spending divided by manufacturing employment. The values in Panels A, E, and F are logged.

migration. By 1940, around 30% of employed Southerners were working in agriculture, forestry, and fishing, compared to 12% in the North. However, coinciding with the Second Great Migration, the two regions converged in terms of industry employment share at least by 1990.

Did out-migration induce regional convergence, or did the process of structural change generate out-migration? One persuasive and pervasive explanation is that agricultural development and mechanization drove the out-migration of workers into the North and into manufacturing (Grove and Heinicke, 2003; Boustan, 2016). It is natural that the agricultural development would have simultaneously decreased employment and led to out migrations. However, the influences may not be uni-directional, as labor scarcity from out-migration could have also incentivized endogenous responses in the Southern economy.

Panels E and F present the changes in the capital-to-labor ratio in agriculture and manufacturing. I measure agricultural mechanization using the number of tractors and combines[10] divided by the number of agricultural workers. For manufacturing, I use the reported manufacturing capital spending.[11] By 1940, Southern agriculture was much less mechanized and relied on labor-intensive practices, often dubbed as the "Old South" method of production (Wright, 1986). However, Panel E suggests that the South eventually caught up with the North in terms of agricultural mechanization. Two channels explain the agricultural convergence: relative increases in the number of tractors and combines (numerator) and relative decreases in employment (denominator). Although the decreases in employment would have contributed to the increase in the trend, the South may have also invested more in mechanization, as in other historical episodes in the United States (Hornbeck and Naidu, 2014; Clemens et al., 2018). Empirical analysis examines both channels.

Furthermore, Panel E shows that the Southern manufacturing invested more in capital per worker relative to the North. If the initial capital-to-labor ratio in the South had been low by 1950, the higher spending per worker in the South would have resulted in relative convergence in the manufacturing capital per worker. Note that if workers simply reallocated from agriculture to manufacturing in the South, capital per worker would have likely to be lower in the South. Instead, relatively higher capital spending suggests that the economy might have responded to labor scarcity by raising capital investment. This paper probes into the relationship between large out-migration (Panel A) and regional convergence (Panels B and C). I especially focus on the potential role of capital depending (Panels E and F) as a response to labor scarcity from out-migration.

2 A Model of Migration and Structural Change

This section outlines a simplified model of a two-period, two-country small open economy framework with two industries and two factors of production—labor and capital. I start with a static, closed-economy version of the model and subsequently add an additional period and another region to study dynamic and open-economy implications. The closed-economy model highlights the importance of factor substitutability and factor-augmenting technical change, whereas open-economy predictions emphasize the role of the Heckscher-Ohlin force in trade specialization through factor abundance and factor intensity. The aim here is to study the major mechanisms that can induce structural change in a model with two factors of production.

[10]Census of Agriculture started to collect the number of tractors in 1930, but combines in 1950. Hence, the 1940 value only includes the number of tractors.

[11]Census of Manufactures did not collect manufacturing capital stock during the early 20th century but started to report manufacturing capital spending in 1947 with a few exceptions (e.g. 1992). The value for 1990 is missing.

I proceed by interpreting well-established theoretical results in relation to the Great Migration and subsequent economic development. Specifically, I take the out-migration in the first period as given (the Great Migration) and analyze the implication of resulting relative labor scarcity on the economy and structural change.[12] Subsection 2.3 summarizes the core predictions. A similar interpretation can be applied to another setting where the initial economy contains a sufficient share of the agriculture employment in labor-intensive agriculture. The dynamic spatial general equilibrium model in Section 5 generalizes the model elements into realistic geography and multiple periods. I implicitly assume a constant consumption share with Cobb-Douglas utility. I allow non-homothetic preferences for the quantitative model (Section 5).

2.1 Closed-economy model: Factor substitutability

Agriculture versus non-agriculture. I start with a closed economy in one region, denoted South, with two industries: agriculture and non-agriculture (denoted "a" and "m"). The closed-economy model highlights the role of factor substitutability. Due to the differences in flexibility in combining labor and capital, the two factors reallocate across industries in the opposite direction from the common shock. The non-unitary elasticities also give rise to weak equilibrium biases in technical development (Acemoglu, 2002, 2007; Caselli, 2016).

The production function for each sector $s \in \{a, m\}$ is assumed to take the constant elasticity of substitution (CES) structure using labor L^s and capital K^s (Arrow et al., 1961):[13]

$$Y^s = \left(\rho^s (Z_L^s L^s)^{\frac{\sigma^s - 1}{\sigma^s}} + (1 - \rho^s)(Z_K^s K^s)^{\frac{\sigma^s - 1}{\sigma^s}}\right)^{\frac{\sigma^s}{\sigma^s - 1}}, \tag{1}$$

with labor- and capital-augmenting technologies, Z_L^s and Z_K^s. They are assumed to not exogenously grow but are endogenously affected by the changes in factor allocation. In other words, the production function abstracts from Hicks-neutral and the exogenous components of the factor-

[12]Empirical analysis aims to replicate "out-migration as given," at least in the perspective of the South, using shift-share instrumental variable (SSIV) strategy using pull factors of Northern destination.

[13]Appendix Section 1 adds land as another factor of production in agriculture. Land and capital, respectively, would represent geographically immobile fixed and variable factors of production, while labor can be viewed as a geographically mobile variable factor. I abstract from land in the main model as such an addition does not make much difference in terms of core predictions. Still, I outline an additional implication for labor-saving economic development (Acemoglu, 2010) that the introduction of land can generate.

augmenting technologies, as they are not needed for the core predictions.[14,15] For simplicity, I assume that technology cannot adapt in the first period when the shock occurs but can endogenously adjust in the second period. The adjustment process follows the directed technical change process (Acemoglu, 2002, 2007), as introduced shortly after.

The CES production function allows for flexible factor usage with a restriction that the elasticity of substitution between labor and capital, σ, is constant. The parameter ρ governs the labor intensity in production. The value of σ is assumed to be greater than one for agriculture but less than one for non-agriculture by following CES elasticity estimates in the literature. For instance, Herrendorf et al. (2015) estimate the value of σ for agriculture (1.58), manufacturing (0.80), and service (0.75), using the U.S. macro data for 1947-2010.[16] Oberfield and Raval (2021), by focusing on manufacturing, estimate the elasticity parameter at plant-level (0.3-0.5) and macro-level (0.5-0.7) using the U.S. Census of Manufactures for 1972-2007. Using global panel datasets, Boppart et al. (2023) estimate elasticity in contemporary agriculture to be 1.90. Hence, agriculture can be regarded as a flexible sector, while non-agriculture is relatively inflexible in factor usage.[17]

Out-migration and structural change. As the population flows out, labor becomes more scarce and expensive relative to capital. Such changes generate heterogeneous sectoral responses due to different factor usages. Here, I document how the share of capital allocated to the agriculture, $\kappa = K^a/(K^a + K^m)$, and labor share in agriculture, $\lambda = L^a/(L^a + L^m)$, responds to the change in regional capital-labor ratio, $k = (K^a + K^m)/(L^a + L^m)$. Note that the initial industry of the migrants is irrelevant for the changes in k. I interpret capital as a geographically immobile variable factor, such as local structures for production. I assume labor and capital are fully employed and perfectly mobile across sectors within the region.

[14] In real life, technological advances would consist of both Hicks-neutral and factor-augmenting (non-neutral) components, the latter of which could be either exogenous or endogenous. With the Cobb-Douglas demand, the growth of Hicks-neutral technology does not affect any predictions of the model as long as its growth rate is the same across industries. Differences in technical growth between industries, on the other hand, lead to classic Baumol (1967) effects. See Duernecker et al. (2023) for a recent theoretical treatment. The same growth rate of factor-augmenting technologies within an industry is equivalent to Hicks-neutral technology. However, differences in growth rates can act as another source of structural change. See Alvarez-Cuadrado et al. (2017) for related results.

[15] Empirical analysis aims to be consistent with such an abstraction through the parallel trend assumption. In other words, the strategy assumes that regions with different levels of out-migration changed the same in terms of Hicks-neutral technology and the exogenous components of factor-augmenting technologies. Section 3 illustrates the approach. The quantitative model in Section 6 allows Hicks-neutral technology and the exogenous components of the factor-augmenting technologies. Still, the baseline quantification focuses on the endogenous component.

[16] Using their methodology for measuring factor shares, Caunedo and Keller (2024) estimate the value of σ for agriculture (1.23), manufacturing (0.84), and service (0.74) between 1948-2020.

[17] Since Arrow et al. (1961), the CES elasticity estimates on the U.S. have tended to report a value significantly less than one (see León-Ledesma et al. (2010) for a review). However, the absolute majority of them focus on the aggregate economy or non-agriculture.

Predictions 1 and 2 reinterpret Alvarez-Cuadrado et al. (2017) and amend their results to examine the changes from out-migration and endogenous technology adoption. In order to obtain analytical results, I use $\sigma_A > 1$ but set $\sigma_M = 1$. The same result can be obtained with $\sigma_A = 1$ and $\sigma_M < 1$. The results would only be strengthened with $\sigma_A > 1$ and $\sigma_M < 1$.

Prediction 1 (Static response)**.** *Assume that the elasticity of substitution between labor and capital for the flexible sector (agriculture) is greater than one, $\sigma_A > 1$, while the elasticity of the inflexible sector is equal to one, $\sigma_M = 1$. As the economy-wide capital-labor ratio, k, increases, the fraction of capital allocated to the more flexible sector (agriculture) increases, while the fraction of labor decreases. In particular,*

$$\frac{\partial \kappa}{\partial k} = \frac{(1-\sigma)}{\sigma G(\kappa)k} > 0$$
$$\frac{\partial \lambda}{\partial k} = \left(\frac{\alpha}{1-\alpha}\right)\left(\frac{\lambda(\kappa)}{\kappa}\right)^2 \frac{\sigma-1}{\sigma G(\kappa)k} < 0.$$

where $G(\kappa) \equiv \left[\frac{1}{\sigma(1-\lambda(\kappa))} + \frac{1}{\lambda(\kappa)}\right]\left(\frac{\lambda(\kappa)}{\kappa}\right)\left(\frac{\alpha}{1-\alpha}\right) + \left[\frac{1}{\kappa} + \frac{1}{\sigma(1-\kappa)}\right]$.

Proof. See Online Appendix A. □

Prediction 1 focuses on factor reallocation channel. It clarifies how out-migration leads to an increase in capital allocated to agriculture while also inducing structural change out of agriculture. As labor becomes scarcer, the flexible sector substitutes now more expensive labor with capital, releasing labor and absorbing capital. Due to $G(\kappa)$, labor and capital shares always take the opposite direction from the changes in the capital-labor ratio.

Now, I posit a second period that allows endogenous technology adoption and capital investment. I use the prime notation ($'$) to denote the second period. The model abstracts from capital depreciation. First, I introduce the following remark:

Remark 1 (Remark on Prediction 1)**.** *Assuming that the South optimizes its levels of technology, the direction of technical change would exhibit weak equilibrium bias as follows (Acemoglu, 2007):*

$$\frac{\partial (Z_K^{s\prime}/Z_L^{s\prime})^{\frac{\sigma^s-1}{\sigma^s}}}{\partial (K^s/L^s)} > 0, \tag{2}$$

where the term Z_K^s/Z_L^s represents the relative level of capital- to labor-augmenting technology.

In other words, an increase in the sectoral capital-labor ratio induces technological change biased toward capital or labor depending on the value of σ. The technical changes can be thought of as generated by learning-by-doing (Arrow, 1962) or directed R&D efforts (Kennedy, 1964). Remark 1 is equivalent to imposing additional assumptions on the technology environment as in Acemoglu (2007).[18] The CES function meets the required assumption for the production side.

Two competing forces can influence the direction of technical change: price effects that are biased toward scarce factors and market size effects that benefit abundant factors. Given that labor and capital are gross substitutes in agriculture, increases in capital usage raise the relative profitability of capital-augmenting technology. As a result, the relative efficiency of capital increases. On the contrary, the price effects dominate in non-agriculture as two factors are gross complements. Note that Prediction 1 anticipates an increase in labor allocated to non-agriculture. Because both the value of elasticity and factor allocation take the opposite direction for non-agriculture, weak equilibrium bias would again favor capital in non-agriculture. Such technical changes further raise the economy-wide capital-to-labor ratio, leading to the following Prediction for the second period:

Prediction 2 (Dynamic response)**.** *Assume that the elasticity of substitution between labor and capital for the flexible sector (agriculture) is greater than one, $\sigma_A > 1$, while the elasticity of the inflexible sector is equal to one, $\sigma_M = 1$. With technology in both sectors exhibiting (dynamic) weak equilibrium bias, the fraction of capital allocated to the more flexible sector (agriculture) increases while the fraction of labor decreases.*

To sum up, the increases in the capital-to-labor ratio from the out-migration dynamically incentivize the adoption of capital-augmenting technology and capital accumulation, further raising the capital-to-labor ratio, k, in the second period. It leads to the same directions for the factor reallocation as in the first period. Hence, the simple framework features the out-migration as a source of "big push" as an endogenous outcome.

2.2 Open-economy model: Factor intensity

I introduce another region, the North, to investigate the open economy implications. Factor substitutability, the focus of the closed-economy predictions, does not yield direct implications for the open economy because trade depends on relative comparison to the North. However, a potential tension between labor and capital from factor intensity can give rise to the Heckscher-Ohlin force. For Prediction 3, I impose additional assumptions on factor abundance and factor intensity: the

[18] Acemoglu (2007) lay out a menu of different assumptions, unrelated to the above production structure, that can lead to equilibrium bias of technology. For instance, there could be a technologist monopolist that supplies technologies to good producers. She produces intermediate goods.

South is abundant in labor (Figure 1, Panel E), and agriculture is intensive in labor (Bateman and Weiss, 1981; Wright, 1986). I also assume that the relative factor intensity is not reversed from the factor reallocation.

Prediction 3 (Open economy)**.** *At constant relative goods prices, the decrease in the labor endowment in the South in the first period leads to a contraction of the agricultural sector with an absolute increase in non-agriculture production. Non-agriculture absorbs both labor and capital.*

Related proof and discussion are documented in Online Appendix Section A. As in the closed economy, open-economy forces predict a decrease in labor share in agriculture. However, the open-economy forces expand non-agriculture while shrinking agriculture, which leads to different implications for capital allocation. I summarize the common and competing predictions between the two types of economy in the next subsection. In the second period, if materialized, economy-wide capital accumulation would lead to a further expansion of non-agriculture while an absolute decrease in agriculture production.

2.3 Discussion

Summary of the predictions. Closed- and open-economy results rely on related but distinct assumptions. The closed-economy model focuses on differences in labor-capital substitutability, whereas the open-economy forces rely on differences in factor intensity between the industries.

First, there are common and non-competing predictions. In the first period, both types of models predict the reallocation of labor from agriculture to non-agriculture. The trade channel raises the non-agriculture production. Although the closed-economy model assumes that the industry share is fixed by the consumption share, there is no force from the closed economy that works vely increase through time the non-agricultural expansion. In the second period, the direction of technical change predicts a relative improvement in capital efficiency, which could result in capital accumulation and further expansion of the non-agriculture sector.

However, there are notable differences in terms of agricultural capital and production. In the static case, the increase in the economy-wide capital-to-labor ratio induces capital adoption in agriculture, while trade effects lessen it. Which effects would dominate depends on the strength of the closed-economy force (factor substitutability) and the Rybczynski effect (factor intensity). However, note that even if the agricultural share of capital in the economy decreases, the sectoral capital-labor ratio in agriculture would still increase with its elasticity of substitution greater than one. It will result in capital-biased technical change and capital accumulation, which may, in turn, increase agricultural production in the long run.

If agricultural output decreases, it favors the Hecksher-Ohlin channels, and one could expect accompanying decreases in agricultural capital. On the other hand, increases in agricultural capital are more consistent with closed-economy prediction, and hence, agricultural output is likely not to experience much change or even increase. In either case, the agricultural capital stock would progressively increase as time passes with capital-biased technical change.

The empirical analysis first checks the common and non-contradictory predictions. To be consistent with both frameworks, labor should reallocate from agriculture to non-agriculture, and non-agriculture production and capital should increase. Non-agriculture capital is also expected to be increased. On the other hand, the changes in agricultural capital and production are left as empirical questions. Then, I use quantitative analysis to assess the role of each component based on the changes in model outcomes.

Endogenous migration in the second period. The model takes the first period Southern out-migration as given, but it can allow endogenous migration in the second period. In the view of the standard migration settings (e.g. Rosen-Roback framework), the wage increases raise the value of living in the South in the second period, holding migration costs and amenities constant. Thus, the out-migration in the first period itself would decrease gross migration flow from the South to the North and increase the flow in the opposite direction. Still, the net effect is ambiguous and depends on the exact value of living in each region. However, if structural change and capital accumulation sufficiently raise the Southern wage, the net migration flow could be reversed in the second period.

3 Empirical Strategy

3.1 Data

First, I use the complete-count Census between 1910 and 1940 (Ruggles et al., 2024a,b) to generate county-level variables and to construct county-to-county level transition matrices for shift-share design. I use two sets of migration matrices, separately constructed for Blacks and Whites. The baseline migration share uses matched individuals between 1910 through 1940 using the Census Tree approach (Buckles et al., 2023). The Census Tree capitalizes on manual matches created by people doing research on their own family histories on FamilySearch.org. Buckles et al. (2023) then extend these links using both traditional and machine learning matching strategies. The datasets provide the largest matches among publicly available methods and also provide links for women and Blacks. Nonetheless, due to its inherent limitations, the linking method could overstate migration rates as incorrect matches would seem to be a move. On the contrary, it could understate them because migrants are harder to link. Hence, I also document the robustness in terms of migration share using the 1940 Census, which asks for state and country of residence 5 years ago. I use this information to construct an alternative migration matrix between 1935 and 1940.[19]

Secondly, I use county-level datasets: "*Historical, Demographic, Economic, and Social Data: The United States, 1790-2002*" (Haines et al., 2010), henceforth HDES, and agriculture Census compiled in Haines et al. (2018). HDES contains county and state-level information on agriculture, manufacturing, retail, and wholesale activity and various county characteristics. "*United States Agriculture Census, 1840 - 2012*" (Haines et al., 2018), henceforth Agriculture Census, contains rich information about county-level agricultural variables, such as farm output, farm value, and the number of tractors. I supplement the analysis using the county business patterns (Eckert et al., 2022; Census Bureau, 2023), henceforth CBP, and the County and City Data Book (Census Bureau, 2012). The CBP contains employment and information related to economic activity for detailed industry codes covering all counties in the United States. Given the long study periods, I match the aggregate values from the above datasets using the time-series data from the "*Historical Statistics of the United States*" (Carter et al., 2006) whenever possible.

The main sample is 1,148 counties in the South between 1940 and 2010. The definition of the South follows Boustan (2016) by mainly using the Census definition but classifying border states based on their migration history. The sample states contain former Confederate states,[20] Kentucky, Oklahoma, and West Virginia, but exclude the District of Colombia, Maryland, and Delaware. I denote counties in these states as "the South," while I use the terminology "the North" to denote counties in the contiguous U.S. outside the South. I restrict the sample to balanced counties for agriculture and manufacturing variables, except for the number of combines and manufacturing capital spending that started to be collected in 1950 and have less coverage. I also report the results using balanced counties for all main variables and by limiting the sample to former Confederate counties. Note that even within the balanced counties, the number of observations can

[19]For the migration between counties in the contiguous U.S., there are 9,622,404 ($= 3102 \times 3102$) possible combinations, including stayers. The Census Tree approach between 1910 and 1940 generates 1,748,472 (18.2%) non-zero migration cells among all possible flows, while the 1940 Census approach between 1935 and 1940 generates 830,892 (8.6%) non-zero cells.

[20]South Carolina, Mississippi, Florida, Alabama, Georgia, Louisiana, Texas, Virginia, Arkansas, Tennessee, and North Carolina.

differ between variables as some variables are not reported in specific Census years.

The values from different datasets are linked to the closest decadal year. For instance, the Census of Agriculture is taken every five years and was conducted in 1997 and 2002, for instance. For 2000 values, I take an average of the two nearby values. The CBP, on the other hand, is conducted annually, and its 2000 values are matched to the year 2000. The county border across different years is adjusted to 1990 boundaries (Eckert et al., 2020). Further details on data are described in Online Appendix Section B.

3.2 Estimating equation

Empirical analysis examines the relationship between out-migration between 1940 and 1970 with economic changes in the South after 1970. I adopt a shift-share design by combining a predetermined migration share and predicted numbers of Northern in-migrants, mainly using non-economic variables to instrument the county-level out-migration in the South. The identification relies on the parallel-trend assumption: counties with different levels of out-migration should have changed the same after 1970.

The baseline specification estimates year-specific differences between counties with different levels of net out-migration rates between 1940 and 1970, "$\left(Net\_Out\_Mig^{1940-1970}\right)_c$:"

$$Y_{c,t} = \beta_t \left(Net\_Out\_Mig^{1940-1970}\right)_c + \alpha_{s,t} + \alpha_c + \gamma X_{c,t} + \gamma_t X_c + \varepsilon_{c,t}. \tag{3}$$

The main explanatory variable summarizes different levels of out-migration during the Second Great Migration period. It is defined as a negative value of the net migration rate during the 30-year period, calculated by the number of net migrants divided by the 1940 population. The number of net migrants is calculated from decadal net migration rate estimates (Gardner and Cohen, 1992; Voss et al., 2005; Fuguitt et al., 2010; Winkler et al., 2013; Bowles et al., 2016). I also estimate a pooled version of equation (3) as a difference-in-differences estimator: it compares the outcomes before and after the Second Great Migration between counties that experienced different levels of out-migration.

The omitted base years are 1940 and 1950. Hence, the β_t captures the changes in outcome $Y_{c,t}$ relative to its levels in 1940 and 1950.[21] Compared to using only 1940 values as an initial condition, using both decadal years would minimize the influences of potential changes related

[21] A set of estimates $\{\hat{\beta}_{1970}, \cdots, \hat{\beta}_{2010}\}$ captures relative effects in 1970, $\cdots$, 2010, compared to its average value between 1940 and 1950.

to World War II during the 1940s (Jaworski and Yang, 2024). Compared to using all decadal years, I exclude the year 1960 because the values are expected to contain the influences of the out-migration in earlier periods. Hence, the estimates have an interpretation as a donut estimator. I also report the results by including 1960 values as a robustness check.

I include state-by-year fixed effects to account for national- and state-level time trends and county fixed effects to remove county-level time-invariant unobservable factors that may confound the results. Hence, the identifying variation uses differences in changes in outcome, relative to each county's base value, between counties that experienced different levels of out-migration within the same state in the same year.

As a set of time-varying controls, $X_{c,t}$, I use the log population and contemporaneous net migration rate. They play important roles in taking into account the changes in county sizes and the potential role of migration after 1970 that may be correlated with the Second Great Migration. However, one might be concerned about the presence of contemporaneous variables as they would have been affected by the main explanatory variable itself. Hence, I also report the results without these controls.

Even within the same state, counties with different initial conditions before 1940 may not have changed the same after the Great Migration period. Hence, I include extensive sets of time-interacted variables to ensure that the estimation relies on comparisons between counties in the same state with similar pre-Great Migration characteristics. Specifically, I include time-interacted values of (1) time-invariant county characteristics, (2) 1940 agriculture conditions, (3) New Deal variables, and (4) trade exposure.

First, I include log land area, longitude, latitude, and 1940 values of log population. They adjust for time-varying effects of initial county sizes and suitability for agriculture. Next, I include time-interacted values of agriculture variables, given that agriculture practices in the United States continued to be developed during the 20th Century. For instance, during this period, cotton production rapidly mechanized, while tobacco production continually declined, where both of them had been traditional cash crops in the South (Whatley, 1985; Holley, 2000; Jung, 2020). To minimize the role of such agricultural "push factors," I use 1940 values of the share of sharecroppers[22], total farm acres, and acres harvested in cotton, tobacco, corn, and hay, respectively.[23] I also include the

[22] In 1940, around 19% of agriculture workers in the Southern sample were sharecroppers. The share was as high as 37% in the Deep South states in 1940, especially pronounced in the Mississippi Delta. Such a prevalence may have influenced subsequent agricultural development (Day, 1967; Ferleger, 1993).

[23] Corn production took up the largest land in terms of acre harvested. Acre in hay is included as a proxy for livestock production.

shares of farms in five different farm-size bins, as initial farm size may have influenced the adoption of tractors or other agricultural practices (Grove and Heinicke, 2003; Manuelli and Seshadri, 2014).

I use two New Deal investment variables from Fishback et al. (2005), because different levels of investments during the 1930s may have affected subsequent economic development. For instance, Kline and Moretti (2014) finds the long-run positive impacts of the Tennessee Valley Authority investment in Southern manufacturing. Hence, I include the federal grants on public work and the payments to farmers through the Agricultural Adjustment Administration (AAA). The former captures the possible influences of the New Deal program in the form of public buildings and infrastructure (Schulman, 1994). AAA spending, on the other hand, paid farmers to take land out of production, with potentially distinct impacts than infrastructure investment (Depew et al., 2013).

A control for the trade exposure uses the Japanese import penetration measure from Batistich and Bond (2023), as the Japan shock is the most relevant trade shock in the late 20th Century, the main study period.[24] Regarding migration, first note that the contemporaneous net migration rate includes both internal and international migration. Furthermore, between the Johnson-Reed Act of 1924 and the Immigration and Nationality Act of 1965, U.S. international migration was largely restricted with a quota system. One notable exception during the Great Migration period was the Bracero program, a government-sponsored program that temporarily received Mexican workers for farm and railroad between the years 1942 and 1964. At least 4.2 million Mexicans entered the US through the program. The best available data on the direct measure of the program is state-level Bracero exposure digitized by Clemens et al. (2018). The state-by-year fixed effects capture relevant variation.

Baseline estimation is not weighted and the estimates capture the average outcome per county. The results are similar using the 1940 population as the weight, which represents the average outcomes per person in the initial period. All standard errors are clustered at the county level to take into account serial correlation within a county. As a robustness check, I report Conley standard errors that allow spatial correlation.

3.2.1 Shift-share instrument design

The main identification challenge is that the dependent variables could have reversely affected the out-migration rate in Southern county c. Omitted variables could also confound the estimates. The

[24] Note that the county fixed effect removes the level of trade exposure, and thus, the control variable needs to take into account the influences of the changes in trade exposure.

extensive sets of fixed effects and control variables are chosen to minimize such possibility. Still, to ameliorate remaining concerns, I limit the variations in Southern out-migration to the component that is explained by Northern pull factors.

The main empirical analysis utilizes a shift-share instrumental variable (SSIV) strategy that combines two sources of variation through fixed, predetermined migrant networks ("share") with pull factors of receiving cities ("shifts"). In the Great Migration setting, Boustan (2010), Derenoncourt (2022), and Bazzi et al. (2023) construct predicted shifts based on push factors of Southern origin to instrument the number of in-migrants to Northern destinations. I instead construct predicted shifts from Northern destinations to calculate the number of out-migrants in the origins.

In the zero-stage regression of in-migration prediction, I use OLS regression with latitude, longitude, log values of the total, Black, and White population, urbanization, median income, median rents, total housing units, 1940 values of the share of foreigners, Black and White, 1940 values of employment share and median occupational score, and Republican vote share of presidential elections between 1940 and 1972. I separately predict Black and White in-migration (Table A1). I also report the results with alternative shifts using actual in-migration rates as in Card (2001) and random forest prediction, an ensemble machine-learning technique based on the decision tree method.

Then, I allocate the predicted number of migrants back to Southern counties using pre-period migration shares. The baseline share uses matched individuals in 1910 through 1940 using the Census Tree approach (Buckles et al., 2023; Ruggles et al., 2024a). I check the robustness using the 1935-1940 migration from direct data on 1935 locations in the 1940 Census. The share is constructed separately for the Black and White migration to take into account heterogeneous migration patterns by race (Collins and Wanamaker, 2015). Finally, the first-stage regression instruments the actual out-migration rate by the constructed migration exposure from the North.

Discussion. The exclusion restriction of an SSIV can rely on either share exogeneity (Goldsmith-Pinkham et al., 2020) or shift exogeneity (Borusyak et al., 2022). I mainly discuss the identification strategy regarding the share, but the empirical strategy also takes into account the shift exogeneity.

Goldsmith-Pinkham et al. (2020) show that SSIV is numerically equivalent to a GMM estimator with the shares as a large set of instruments and a weight matrix constructed from shifts. Shares are allowed to be correlated with the levels of outcomes since the strategy asks whether differential exposure to common shocks leads to differential changes in the outcome. In this paper's setting, this condition requires migration linkages before 1940 (predetermined migration share) to be orthogonal to the changes in outcomes after 1970, conditional on observables. Note that county fixed effects isolate variation in changes and remove any time-invariant county characteristics that could have influenced the levels of migration linkages before 1940. The strategy can be viewed as a DiD-IV strategy with a parallel trend assumption.

In an alternative view, the identification of SSIV comes from exogenous shifts (Borusyak et al., 2022). Here, the empirical strategy uses a large set of Northern pull factors, which can plausibly assumed to be unrelated to the changes in Southern economic outcomes. Note that state-by-year fixed effects remove national- and state-level trends and, hence, any common shocks to the destinations and origins.

In applying SSIV, Jaeger et al. (2018) cautions the potentially confounding influences of serially correlated migration. While migration can induce both short—and long-term changes, the flows of migrants themselves tend to be correlated, making it hard to distinguish between longer-term adjustments and the influences of lagged migration. To limit such a concern, I restrict the attention only to longer-term changes induced by the 1940-1970 flow, which is unique in its size and breadth compared to the migration flows before or after (Section 1). Note that the estimation equation also controls for the contemporaneous net migration rates.

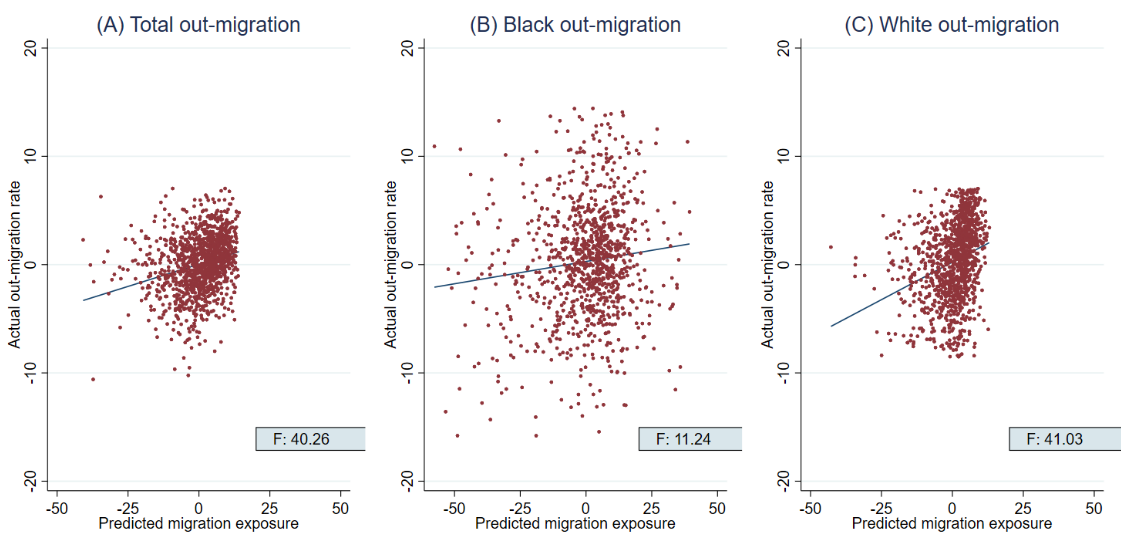

Figure 2: Scatter plot of the First-Stage.

Note: The figure presents first-stage regression results, with residualized F-statistics on the square box at the bottom right. The y-axis plots the net migration rates between 1940 and 1970 in Southern counties, and the x-axis is the migration exposure measure for years between 1940 and 1970, constructed by the SSIV strategy described in Section 3.2.1. The net migration rates are from Gardner and Cohen (1992) and Bowles et al. (2016).

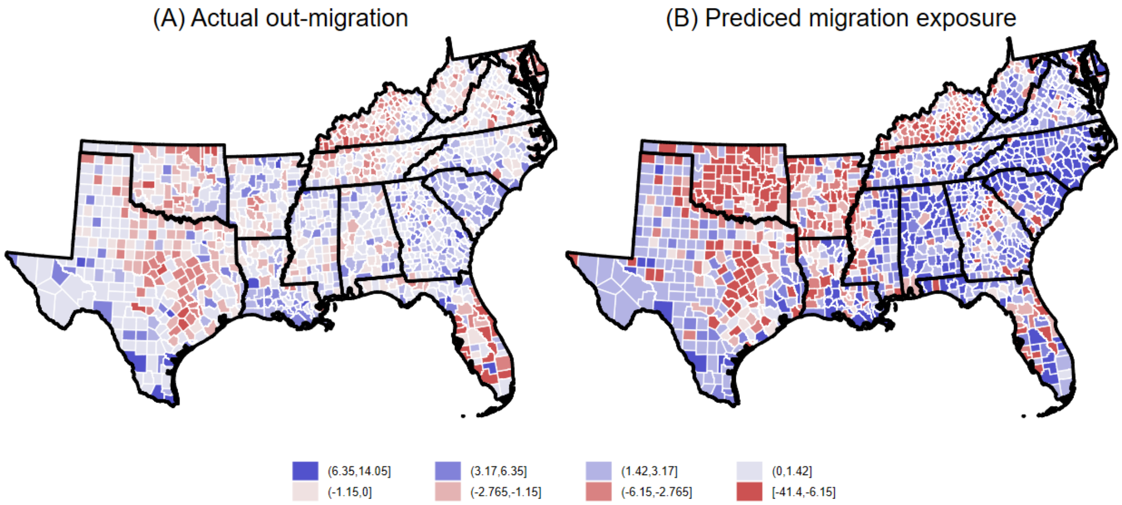

Figure 3: Map of the out-migration rate and predicted migration exposure.

Note: The figure presents the geographical distribution of the endogenous variable, actual out-migration, on Panel A. Panel B plots the excluded instrument, predicted migration exposure constructed by the SSIV design. Blue represents net out-migration, and red indicates net in-migration. Darker colors represent a higher level of absolute sizes. Both variables are not residualized.

3.2.2 First-stage results

Figure 2 shows the first-stage regression in Panel A, while Panels B and C separately report the results by race. The square box at the lower right reports residualized F-statistics with a full set of controls and the state fixed effect but without the county fixed effect. The baseline estimation combines both Black and White out-migration, as in Panel A, while the robustness check also separately reports the results by race, as in Panels B and C. Both are similar to the baseline outcomes using the total out-migration but less precise.

Overall, there exist strong correlations between actual out-migration rates and constructed migration exposure. However, the relationship tends to be weaker for Black out-migration. This could be due to the limitation of the linking approach in general, as it is harder to link historically similar Black names. An alternative approach using 1935-1940 migration shares exhibits a relatively stronger relationship for Blacks but overall weaker relationships when combined with Whites (Figure A1).

Figure 3 plots the map of the non-residualized values of the endogenous regressor, actual out-migration rate (Panel A), and the excluded instrument, predicted migration exposure (Panel B) on the same color scheme. While they are distributed similarly, the exact values differ by each county's experience on the levels of Northern pull factors.

3.2.3 Pretrend tests

While directly testing the validity of an instrumental strategy is, in general, not feasible, this section documents the robustness and limitations of the baseline strategy through commonly used statistical tests. For Table A2, I estimate Equation (3) on pre-period outcomes for variables that report pre-period information. The results examine whether pre-period changes (the values in 1920 and 1930 compared to 1940) in the main dependent variables are systemically correlated with the out-migration rate between 1940 and 1970.

Panels A to C each show the outcomes on agriculture, manufacturing, and wholesale variables. Overall, they do not show a clear pattern and are statistically insignificant, suggesting there was no significant pretend that drove the main results. In other words, the baseline estimates do not simply capture the underlying correlation between the economic changes in the South and out-migration between 1940 and 1970.

One exception is agricultural employment (Panel A, column 1), which suggests that agricultural employment in 1940 was lower than its values in 1920 and 1930 for the counties that experienced higher levels of out-migration rate between 1940 and 1970. This could have been driven by a correlation between the changes in agricultural employment in the pre-period and migration patterns during the pre-period. To account for such a relationship, I add time-interacted pre-period values as an additional control for agriculture employment. Online Appendix Section D further documents overidentification tests using alternative shares and placebo tests using a random share and random shift.

4 Empirical Evidence

This section examines how the Second Great Migration shaped the subsequent economic development in the South. First, I study whether relative labor scarcity from out-migration led to agricultural mechanization. I then document novel findings on the relationships between out-migration and subsequent developments in manufacturing and services. The observed changes in the economy are discussed in terms of regional structural change.

4.1 Agriculture

Table 1 reports the estimation results on agricultural variables. All dependent variables are logged values and have semi-elasticity interpretation with respect to the out-migration rate. They compare

Table 1: Estimation results for agricultural variables.

	(1) Agriculture employment	(2) Number of farms	(3) Acres in farmland	(4) Number of tractors	(5) Number of combines	(6) Farm output	(7) Farm value per acre
(A) Out-migration rate (OLS)	-0.044***	-0.015***	0.001	0.014***	0.007**	0.015***	0.017***
Clustered s.e. (county)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)	(0.003)	(0.004)
Fixed effects	No	No	No	No	No	No	No
Controls	No	No	No	No	No	No	No
(B) Out-migration rate (SSIV)	-0.333***	-0.193***	-0.042**	0.197***	-0.019	0.232***	0.427***
Clustered s.e. (county)	(0.090)	(0.053)	(0.017)	(0.048)	(0.020)	(0.060)	(0.114)
Fixed effects	No	No	No	No	No	No	No
Controls	No	No	No	No	No	No	No
(C) Out-migration rate (SSIV)	-0.067***	-0.000	-0.029*	0.050**	0.082	-0.004	-0.012
Clustered s.e. (county)	(0.022)	(0.012)	(0.016)	(0.025)	(0.122)	(0.016)	(0.012)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No	No	No
(D) Out-migration rate (SSIV)	-0.074***	-0.014	-0.043**	0.066***	0.039	-0.016	-0.016
Clustered s.e. (county)	(0.026)	(0.014)	(0.018)	(0.024)	(0.168)	(0.024)	(0.013)
Conley s.e. (250km)	[0.016]	[0.012]	[0.015]	[0.020]	[0.103]	[0.019]	[0.010]
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F	14.76	14.99	14.99	15.03	8.56	14.99	14.99
Counties	1,148	1,148	1,148	1,148	1,103	1,148	1,148

Note: The table reports estimation results for agricultural variables using OLS on Panel A and SSIV on Panels B to D. All dependent variables are logged values and have semi-elasticity interpretation with respect to the out-migration rate. Panels A and B do not include any fixed effects or control variables. Panels C and D add state-by-year and county fixed effects and control variables described in Section 3.2. Each column reports the changes in the indicated outcome variable for the years 1970 to 2010, relative to the omitted years of 1940 and 1950, except for the number of combines with the omitted year of 1950. Robust standard errors are clustered by county and reported in parentheses. Panel D also reports Conley (1999) standard errors with 250 km (155 miles) as a cutoff in square brackets and the first-stage Kleibergen-Paap robust F-statistics. Stars represent: $* \ p < 0.10$, $* \ p < 0.05$, $* * * \ p < 0.01$.

the outcomes before (1940 and 1950) and after (1970 to 2010) the Second Great Migration between counties that experienced different levels of out-migration. Panels A and B, respectively, present the OLS and SSIV results without any fixed effects or control variables. They demonstrate the raw relationships between out-migration and outcome variables. Panel A shows that higher levels of out-migration are associated with less agricultural employment and fewer farms (Columns 1 and 2), but with higher numbers of tractors and combines (Columns 4 and 5), the measure of agricultural mechanization. Panel B tends to display similar patterns but with bigger magnitudes.

The OLS estimates capture any correlation between the dependent and explanatory variables, while the SSIV isolates the relationship from the explanatory variable explained by the differential exposure to the Northern migration pull factors. [25] In Panel A, the negative association between

[25]This interpretation follows the share view of Goldsmith-Pinkham et al. (2020). The strategy compares the changes in outcomes between counties with different predetermined migration shares before the Second Great Migration,

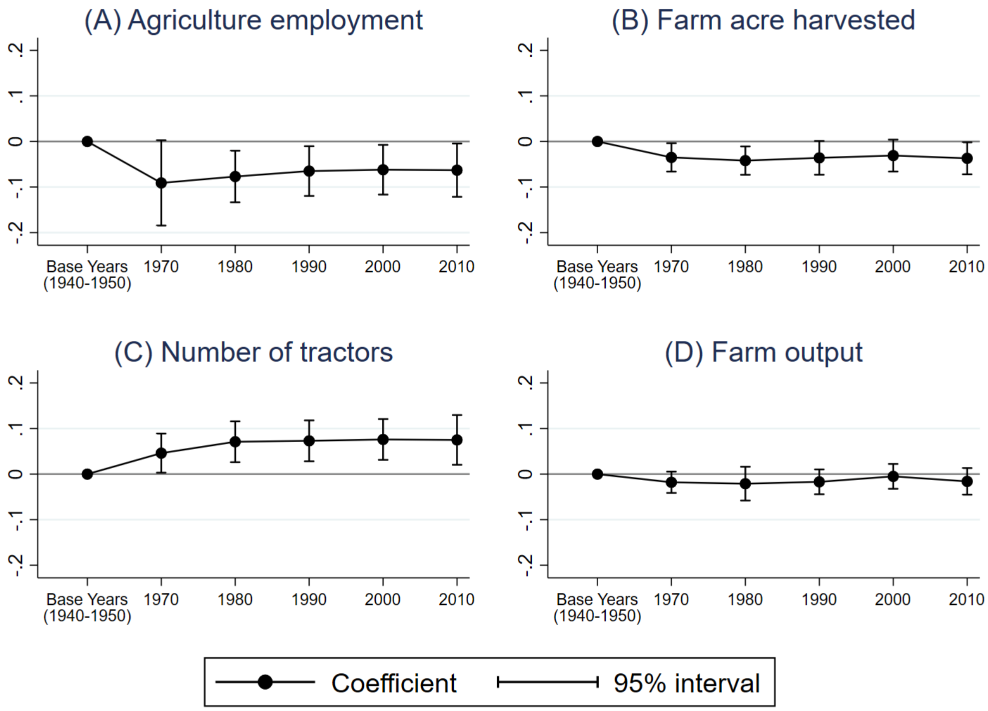

Figure 4: Time trends in agricultural estimates.

Note: The figure presents the estimates from Equation (3) with 95% confidence interval. The reported variables are agricultural employment (Panel A), land in acres in farmland (Panel B), the number of tractors (Panel C), and total farm output (Panel D). They correspond to the year-specific version of Table 1, Panel D with the sull set of fixed effects and control variables. The coefficients estimate the changes in the indicated outcome variable in each year, relative to the omitted years of 1940 and 1950. Robust standard errors clustered by county.

agricultural development and out-migration suggests that agricultural push factors, such as mechanization, played an important role in driving regional out-migration (Kirby, 1983; Boustan, 2010, 2016). However, such relationships are stronger when limiting attention to out-migration induced by Northern pull factors.

Panel C estimates relative changes in outcomes between counties within the same state in the same year by adding state-by-year and county fixed effects, where the reported estimate is an average across states and years. Panel D further adds control variables so that the comparison relies on counties with similar pre-Great Migration characteristics, where the parallel trend assumption is likely to hold. I report Kleibergen-Paap robust F[26] for Panel D. A comparison between Panels B,

weighted by common pull factors from Northern counties during the Second Great Migration period. Alternatively, the SSIV estimation can be viewed as isolating the relationship from the explanatory variable that is explained by Northern pull factors, weighted by fixed, predetermined migration share as an exposure (Borusyak et al., 2022).

[26]With one endogenous variable and one instrument, the value of robust F is equal to the value efficient F developed by Olea and Pflueger (2013).

C, and D indicates that the majority of the residual variations is captured by the set of fixed effects.

Table 1, Panel D, reports the baseline estimation results from the aggregated version of the Equation (3). Columns 1 to 3 show that a 100% increase in pull factors driven out-migration reduced agricultural employment by 7.4% and farm acres harvested by 4.3%[27] when compared to other counties in the same state with similar pre-Great Migration characteristics. However, the number of farms did not experience a meaningful relative change.

As the county-wide capital-to-labor ratio increased, agriculture substituted labor with capital,[28] proxied by the number of tractors and combines. Column 4 shows that higher out-migration induced relative adoption of tractors (6.6%). The number of combines could have increased as well (3.9%, Column 4).[29] As a result, the overall farm output is not much affected by different degrees of out-migration (Column 6). Similarly, Column 7 reports that the total value of farms, including the value of land, implements, and buildings, tended to experience only negligible changes.

Figure 4 plots the time trend in the estimates for agricultural employment (Panel A), acres in farmland (Panel B), the number of tractors (Panel C), and farm output (Panel D). I use the time-interacted migration exposures as instruments using equation (3). With a 100% higher out-migration rate, agricultural employment decreased by 9.1% in 1970 but recovered to 6.3% by 2010 (Panel A). The acres of farmland decreased by 3.5% in 1970 and stayed at a similar level at least until 2010 (Panel B).

The relative changes in the number of tractors were heightened until 2000 and maintained at least until the year 2010. It increased by 4.6% in 1970 and continued to grow by 7.6% in 2000. The influences of the out-migration wore off to 7.5% in 2010 (Panel C). The observed pattern of continued increases in tractor usage can be rationalized by the directed technical change (Hicks, 1932; Acemoglu, 2002, 2007). As agriculture uses less labor and more capital, agriculture becomes better at using capital, which in turn incentivizes further capital investment. For instance, the diffusion of tractors could have been initiated with a narrow application that can directly substitute labor but subsequently generalized to broader use (Gross, 2018). Finally, the relative null effects

[27] The decreased farmland could have been driven by the introduction of tractors. Tractors augment land by freeing up the land previously allocated to feed farm animals. Between 1930 and 1960, acres of cropland used to feed horses and mules decreased from 65 million acres to 5 million acres (Olmstead and Rhode, 2001). Alternatively, the results could have driven by complementarity between labor and land.

[28] Note that the main explanatory variable is the county-level out-migration rate, representing the total level of out-migration relative to the initial 1940 population regardless of the race or initial industry of the migrants.

[29] The result on the number of combines may be less precisely estimated compared to the number of tractors because of the differences in their usage. A tractor is essentially a power unit that can move agriculture equipment, and a combine is a combination of tractor and harvesting equipment. While tractors are universally used, combines are specialized to harvest field crops.

in farm output (Panel D) suggest that, at least in terms of agriculture, closed-economy forces from factor substitution are stronger than the Hecksher-Ohlin channel from the differences in factor intensity (Section 2.3).

Agricultural economics literature has documented that low labor costs can delay mechanization, while labor scarcity may induce the adoption of labor-saving technologies (see Gallardo and Sauer (2018) for a review). In the 20th United States setting, shrinking labor supply from a natural disaster (Hornbeck and Naidu, 2014) or an abrupt change in migration policy (Clemens et al., 2018; San, 2023) facilitated the adoption of labor-saving capital and technologies in agriculture. The relative labor shortages from pull factor-induced out-migration exhibit a similar pattern (Table 1, Panel D). Such labor scarcity might have also encouraged similar changes in non-agriculture.

4.2 Non-agriculture

Manufacturing. Table 2 reports the manufacturing results. All dependent variables are logged values. Panels A and B, respectively, report the OLS and SSIV estimates without any fixed effects or control variables. Panel A shows that higher levels of out-migration are associated with fewer manufacturing workers and lower levels of manufacturing development, represented by manufacturing capital spending, value added, and annual payroll. On the contrary, Panel B suggests that counties with more pull factor induced out-migration experienced subsequent manufacturing development after 1970. Such contrasting patterns suggest that the role of manufacturing development as migration push factors and the potential role of out-migration on manufacturing have opposite implications.

While agriculture mechanization would have likely pushed workers out of agriculture and out of more agrarian counties (Table 1, Panel A), manufacturing development would have pulled migrants. Hence, Table 2, Panel A shows that counties with more advanced manufacturing tended to experience less out-migration. On the contrary, the influences of pull factor induced out-migration, as evinced by positive associations in Panel B, could have incentivized labor reallocation into manufacturing and physical capital investment. Such disparities between Panels A and B demonstrate the role of the instrumental variable strategy.

To focus on changes in outcomes related to the out-migration, Panels C and D add state-by-year and county fixed effects and control variables so that estimation relies on comparisons between counties in the same state in the same year with similar pre-Great Migration characteristics. The contemporaneous variables condition the size of the counties. The baseline estimates in Panel D, Columns 1 to 5 reveal that out-migration modestly increased manufacturing employment and the levels of manufacturing development.

Recall that out-migration raises the capital-labor ratio. As a response, more flexible agriculture substituted now scarcer labor with capital, releasing workers from agriculture (Table 1, Panel D, Column 1). Some of this labor would have been reallocated to local manufacturing, increasing employment by 9.7% with a 100% higher out-migration rate. The number of establishments also increased, albeit to a lesser extent (Column 2, 5.9%), suggesting manufacturing firm size have grown as well. The relative increases in manufacturing employment incentivized further investment in manufacturing capital due to the labor-capital complementarity; Column 3 finds that relative capital spending increased by 19.9%, more than the relative increase in employment. As a result, manufacturing value added and annual payroll increased by about 11.6% and 13.5%

Table 2: Estimation results for manufacturing variables.

	(1) Manufacturing employment	(2) Manufacturing establishment	(3) Manufacturing capital spending	(4) Manufacturing value added	(5) Manufacturing annual payroll
(A) Out-migration rate (OLS)	-0.026***	-0.032***	-0.045***	-0.027***	-0.021***
Clustered s.e. (county)	(0.003)	(0.002)	(0.004)	(0.004)	(0.004)
Fixed effects	No	No	No	No	No
Controls	No	No	No	No	No
(B) Out-migration rate (SSIV)	0.163**	0.090***	0.746	0.560***	0.516***
Clustered s.e. (county)	(0.064)	(0.034)	(0.597)	(0.204)	(0.179)
Fixed effects	No	No	No	No	No
Controls	No	No	No	No	No
(C) Out-migration rate (SSIV)	0.049*	0.021	0.134*	0.070	0.083**
Clustered s.e. (county)	(0.028)	(0.020)	(0.072)	(0.045)	(0.039)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	No	No	No	No	No
(D) Out-migration rate (SSIV)	0.097**	0.059***	0.199**	0.116*	0.135**
Clustered s.e. (county)	(0.041)	(0.022)	(0.094)	(0.065)	(0.056)
Conley s.e. (250km)	[0.032]	[0.019]	[0.088]	[0.050]	[0.038]
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
First-stage F	14.19	14.37	10.16	13.05	13.63
Counties	1,148	1,148	1,115	1,148	1,148

Note: The table reports estimation results for manufacturing variables using OLS on Panel A and SSIV on Panels B to D. All dependent variables are logged values and have semi-elasticity interpretation with respect to the out-migration rate. Panels A and B do not include any fixed effects or control variables. Panels C and D add state-by-year and county fixed effects and control variables described in Section 3.2. Each column reports the changes in the indicated outcome variable for the years 1970 to 2010, relative to the omitted years of 1940 and 1950, except for the manufacturing capital spending with the omitted year of 1950. Robust standard errors are clustered by county and reported in parentheses. Panel D also reports Conley (1999) standard errors with 250 km (155 miles) as a cutoff in square brackets and the first-stage Kleibergen-Paap robust F-statistics. Stars represent: $* \, p < 0.10$, $** \, p < 0.05$, $*** \, p < 0.01$.

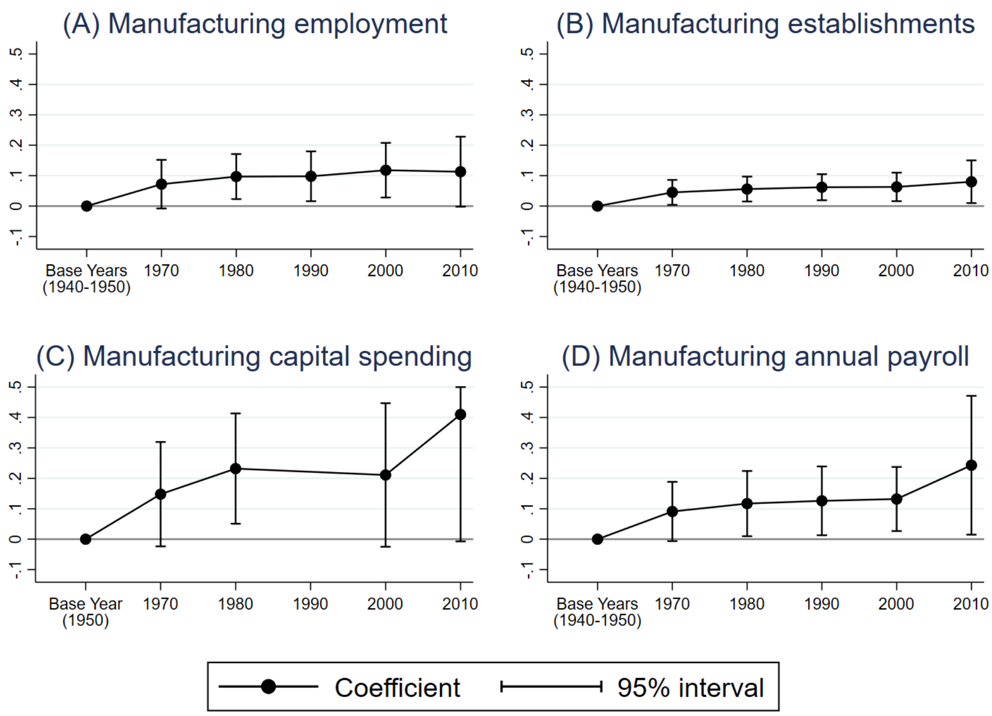

Figure 5: Time trends in manufacturing estimates.

Note: The figure presents the estimates from Equation (3) with 95% confidence interval. The confidence intervals are truncated above at 0.5 for visibility. The reported variables are manufacturing employment (Panel A), manufacturing capital spending (Panel B), manufacturing value added (Panel C), and annual payroll in manufacturing (Panel D). They correspond to the year-specific version Table 1, Panel D with fixed effects and control variables. The coefficients estimate the changes in the indicated outcome variable in each year, relative to the omitted years of 1940 and 1950. Robust standard errors clustered by county.

(Columns 4 and 5), less than the capital spending increase but more than employment. Hence, per worker values of manufacturing value added and payroll would have also increased.

Figure 5 presents the changes in manufacturing employment (Panel A), establishments (Panel B), capital spending (Panel C), and annual payroll (Panel D). Manufacturing employment increased by 7.2% in 1970, and continued to grow until 11.8% in 2000. The influences wore off to 11.3% in 2010. Other outcomes show similar patterns and demonstrate that overall results reported in Table 2 are maintained and even grew at least until 2000 or 2010.

The continued growth of manufacturing in more out-migrated South can be interpreted with the directed technical change and Heckscher-Ohlin framework (Section 2). As manufacturing absorbed labor, weak equilibrium bias in non-agriculture would have favored the complementary factor, capital (Acemoglu, 2007). The capital-biased technological growth would have further incentivized capital investment. Moreover, the Heckscher-Ohlin force suggests that if manufacturing is more capital-intensive, both labor scarcity and capital investment would lead to an expansion of its production through the Rybczynski effect.

Local nontradables. Table 3 documents wholesale and retail outcomes, which are used as proxies for local nontradable services. As in manufacturing, both retail and wholesale experienced positive growth from out-migration, with wholesale reporting stronger positive changes. For instance, with a 100% more out-migration, employment in wholesale increased by 12.6% and retail by 4.7% between 1970 and 2010, compared to their levels in 1940 and 1950, when compared to other counties in the same state with similar pre-Great Migration characteristics. Total sales in each sector grew by 11.7% and 5.1%. The number of establishments and sales also grew after the Great Migration. Figure 6 shows the year-specific changes in employment and total sales in wholesale and retail after 1970. They responded quickly, reaching close to their high levels by 1970, and stayed relatively stable afterward.

Table 3: SSIV estimation results for wholesale and retail.

Panel A. Wholesale				
	Wholesale employment	Wholesale establishment	Wholesale sales	Wholesale annual payroll
Out-migration rate	0.108***	0.074**	0.090*	0.092**
Clustered s.e. (county)	(0.036)	(0.030)	(0.048)	(0.046)
Conley s.e. (250km)	[0.030]	[0.021]	[0.043]	[0.041]
Fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
First-stage F	16.08	15.57	15.83	14.89
Counties	1,135	1,148	1,138	1,138
Panel B. Retail				
	Retail employment	Retail establishment	Retail sales	Retail annual payroll
Out-migration rate	0.046**	0.021*	0.033**	0.056***
Clustered s.e. (county)	(0.020)	(0.011)	(0.015)	(0.021)
Conley s.e. (250km)	[0.017]	[0.010]	[0.014]	[0.019]
Fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
First-stage F	14.97	14.99	14.97	14.96
Counties	1,148	1,148	1,148	1,148

Note: The table reports estimation results for wholesale variables in Panel A and retail in Panel B using the pooled specification of equation 3. All dependent variables are logged values and have semi-elasticity interpretation with respect to the out-migration rate. Each column reports the changes in the indicated outcome variable for the years 1970 to 2010, relative to the omitted years of 1940 and 1950. Robust standard errors in parentheses are clustered by county, and Conley (1999) standard errors in square brackets use 250 km (155 miles) as a cutoff. Kleibergen-Paap robust F-statistics are reported for baseline results in Panel D. Stars represent: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Figure 6: Time trends in wholesale and retail estimates.

Note: The figure presents the estimates from Equation (3) with 95% confidence interval. The reported variables are manufacturing employment (Panel A), manufacturing capital spending (Panel B), manufacturing value added (Panel C), and annual payroll in manufacturing (Panel D). They correspond to the year-specific version Table 1, Panel D with fixed effects and control variables. The coefficients estimate the changes in the indicated outcome variable in each year, relative to the omitted years of 1940 and 1950. Robust standard errors clustered by county.

Given that retail and wholesale use both factors of production, closed-economy forces would have generated similar changes as in manufacturing. However, the local nature of these industries implies that they are less governed by the open-economy the Heckscher-Ohlin force. Instead, the non-tradable sector was subjected to another channel: local spillover effects. The growth of labor payments in manufacturing and the possible increases in agricultural wages would have positively affected the growth of the local tradable industry.[30]

4.3 Structural change

This subsection examines how out-migration affected overall industry composition and the levels of education. Table 4, Panel A, reports the baseline results using employment share in each indus-

[30] Wholesale is closer to a tradable industry than retail because wholesale involves selling large quantities to other businesses. Retail sells products directly to consumers. Hence, the differences between the changes in the two industries give a hint of the role of trdability.

Table 4: SSIV estimation results for industry share and education.

Panel A. Structural change (employment share)

	(1) Agriculture	(2) Manufacturing	(3) Services	(4) Consumer services	(5) Producer services
Out-migration rate	-0.065**	0.086**	0.005	0.022	-0.009
Clustered s.e. (county)	(0.027)	(0.034)	(0.012)	(0.015)	(0.028)
Conley s.e. (250km)	[0.018]	[0.025]	[0.009]	[0.016]	[0.031]
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
First-stage F	13.78	13.25	14.53	13.56	15.87
counties	1,148	1,148	1,148	1,148	1,125

Panel B. Education

	(1) Median school year	(2) Share high school	(3) Share college	(4) Employment in education	(5) Education spending
Out-migration rate	-0.005	0.009	0.007	-0.028	-0.009
Clustered s.e. (county)	(0.004)	(0.007)	(0.033)	(0.018)	(0.006)
Conley s.e. (250km)	[0.005]	[0.005]	[0.029]	[0.014]	[0.005]
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
First-stage F	7.67	17.40	14.58	5.10	11.36
counties	1,148	1,148	1,148	1,146	1,148

Note: The table reports SSIV estimation results for employment shares (Panel A) and educational outcomes (Panel B). All results include state-by-year and county fixed effects and control variables described in Section 3.2. All dependent variables are logged values. Each column reports the changes in the indicated outcome variable for the years 1970 to 2010, relative to the omitted years of 1940 and 1950. The Census, the main dataset, does not report detailed service employment, and I supplement the analysis using CBP for Panel A, Columns 3 to 5. Consumer services are defined to be the 2017 NAICS classification in 42-45 and 61-72. Producer services include 51-56. Robust standard errors in parentheses are clustered by county, and Conley (1999) standard errors in square brackets use 250 km (155 miles) as a cutoff. The first-stage Kleibergen-Paap robust F-statistics are reported. Stars represent: $* \, p < 0.10$, $** \, p < 0.05$, $*** \, p < 0.01$.

try as dependent variables. Panel A, Columns 1 and 2 suggest that out-migration contributed to structural change out of agriculture and into manufacturing. A 100% increase in the out-migration rate resulted in a 7.8% decrease in agriculture employment share but a 7.4% increase in manufacturing share. Column 3 adds services, which experienced relatively small increases in employment share by 1.5%. Such an increase is mainly related to increases in consumer services (Column 4), while producer service tended to experience negative changes. A potential driver of such a disparity is discussed in Online Appendix Section B. The key idea is that producer services, which tend to be least capital-intensive among listed industries, could have been less affected by economic reallocation from the changes in the regional capital-to-labor ratio.

If a region with higher out-migration primarily responded with physical capital investment, the resulting capital deepening might have reduced the incentive for human capital accumulation. Panel B reports outcomes on county-level educational attainment to proxy human capital development. Overall results show that counties with higher out-migration rates did not increase the median school year (Column A) or high school and college graduates share (Columns B and C). Furthermore, they did not invest more in education, both in terms of spending (Column D) and employment (Column E). These results highlight that counties that experienced higher out-migration primarily responded by physical capital deepening, evinced by relative increases in agriculture and manufacturing capital relative to employment (Tables 1 and 2).

The results in Table 4 do not imply that the South failed to raise its overall level of education. On the contrary, the South experienced rapid growth in terms of educational attainment during and after the Great Migration period, on average.[31] Such overall increases in the level of education were pointed out as an important channel that facilitated structural change in the South (Caselli and Coleman II, 2001). However, the counties with different levels of pull factor induced out-migration did not experience relative growth in educational outcomes. Consequently, the role of out-migration, the main focus of this paper, provides a complementary explanation for Southern economic development.

4.4 Discussion and robustness

Here, I interpret the baseline estimates as the local average treatment effects (LATE) and introduce robustness checks documented in the Online Appendix Section D. In an ideal randomized setting, one might randomly allocate the number of migrants across counties and randomize who to migrate within each county. However, the SSIV isolates the component of out-migration induced by Northern pull factors, which could be different from variations in the randomized setting. Nonetheless, the LATE in this paper is not necessarily a weakness in terms of policy implications as the estimation primarily uses the variation generated by migrants who responded to outside incentives. In other words, the findings here could be more applicable to a setting where the government can incentivize people to move out of labor-abundant regions.

Table 5 presents falsification tests using public employment and payroll as dependent variables. The main mechanism in section 2 highlights the private industries' adjustments to the change in

[31] In the South, the share of college graduates for adults older than 25 years old increased from 6.5% in 1940 to 18.6% in 1990. The share of high school graduates also increased from 25.1% in 1950 to 71.4% in 1990. During the same period, college graduates increased from 7.2% to 21.1% and high graduates from 37.4% to 77.3% in the North.

relative factor abundance. On the contrary, the public sector would not have been much affected by such a change. Indeed, local government did not respond in terms of employment and annual payroll (Columns 1 and 2). Federal government employment even may have slightly decreased (Column 3). The results also suggest that the empirical results are not likely to be driven by government policies.

Online Appendix Section D reports the results using alternative approaches. I document robustness checks of main results in Tables 1, 2, and 3, (1) based on 1935-1940 migration matrix using 1935 locations in the 1940 Census, (2) by separately estimating by race, (3) by adopting alternative approaches for in-migration prediction (random forest algorithm and actual number of in-migrants), (4) by limiting sample to former confederate states, (5) using 1940 population-weighted regression, (6) by dropping the time-varying controls (contemporaneous population and net migration rate), and (7) by including 1960 in base years (1940, 1950, and 1960).

5 Quantitative Strategy

5.1 Roadmap to the quantitative model.

In this section, I construct a dynamic spatial general equilibrium model with multiple factors of production. The model is based on canonical models of trade and migration (Eaton and Kortum, 2002; Artuc et al., 2010; Caliendo et al., 2019) with capital accumulation (Kleinman et al., 2023)

Table 5: SSIV estimation results from the public sector (falsification test).

	(1) Local government employment	(2) Local government annual payroll	(3) Federal government employment
Out-migration rate	0.004	0.009	-0.021
Clustered s.e. (county)	(0.006)	(0.008)	(0.028)
Conley s.e. (250km)	[0.006]	[0.010]	[0.028]
First-stage F	10.49	11.90	10.35
County	1,148	1,148	1,148

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

Note: The table reports SSIV estimation results for federal government employment (Panel A), local government employment (Panel B) and annual payroll (Panel C). All results include state-by-year and county fixed effects and control variables described in Section 3.2. Robust standard errors in parentheses are clustered by county, and Conley (1999) standard errors in square brackets use 250 km (155 miles) as a cutoff. The first-stage Kleibergen-Paap robust F-statistics are reported. All dependent variables are logged values.

and structural change (Fan et al., 2023). Here, I extend the aforementioned frameworks by introducing the role of factor substitution and factor intensity in driving structural change and economic allocation. I do so by generalizing the key forces highlighted in Section 2 into multiple periods and realistic geography. Table A3 summarizes the model elements.

The economy consists of a set of discrete locations ($i = 1, \ldots, R$) and three industries: agriculture and tradable and nontradable non-agriculture ($s = A$, M and S). Time is discrete and indexed by t. All sectors use two factors of production, labor and capital, and they are assumed to have CES production structures. Agriculture and tradable non-agriculture are subject to forces arising from factor substitution and trade. The local nontradable sector is only affected by factor substitution, but instead, it is affected by local spillover effects. Tradable non-agriculture consists of manufacturing and (tradable) production services, and it can be regarded as a goods-producing sector. On the contrary, nontradable non-agriculture can be viewed as consumer services that are locally provided.[32] For simplicity, I also refer to them as manufacturing and services.[33]

Compared to Section 2, I directly model economic agents and general equilibrium conditions. There are two types of infinitely-lived agents: workers and capitalists ($\mathscr{T} = L$, and K), each supplying labor and capital. Capitalists are geographically immobile and own the depreciable capital stock in their location, and they make forward-looking decisions over consumption and investment. Workers do not have access to investment technology and live hand to mouth, but they are geographically mobile, subject to migration costs.[34]

I make two simplifications for tractability. First, to obtain analytical expressions for trade and migration with realistic geography, I rely on extreme value distribution assumptions. The resulting expenditure and migration shares take the standard gravity structure. Second, for model calculation, I follow the dynamic exact hat-algebra approach (Caliendo et al., 2019) to eliminate the need to recover counterfactual-invariant fundamentals of the model.

[32] Although the simple framework focuses on the tradable sectors, the addition of a nontradable sector in the quantitative model helps to position agriculture and manufacturing in the economy, as noted in Eaton and Kortum (2002).

[33] The model can further separate tradable non-agriculture into two types of tasks: physical-capital-intensive and human-capital-intensive tasks, where the former uses both factors while the latter only uses labor (Online Appendix Section A). The human-capital-intensive tasks would represent the portion of the economy not affected (if nontradable) or mechanically negatively affected (if tradable) by labor scarcity.

[34] The baseline model assumes that Black and White workers are perfectly substitutable, but a model extension can consider potentially different productivity by race and imperfect substitutability through an additional layer of CES composite of labor. I abstract from this distinction as there is no clear contrast between the influences from Black and White out-migration (Table OA4 and OA5).

5.2 Preferences and factor supply

Preferences. Workers' welfare is defined as the discounted sum of the infinite path of flow utility, which is defined over consumption indirect utility via the log-utility function:

$$U(C_{i,t}^{k}) = \log\left(C(e_{i,t}^{k}, P_{i,t})\right). \tag{4}$$

Workers inelastically supply their labor and earn wages at competitive market rates. Following Boppart (2014) and Fan et al. (2023), individuals' consumption preferences are in the non-homothetic Price-Independent Generalized Linear (PIGL) class. It represents the structural change on the demand side, while the changes in factor allocations capture the structural change on the supply side. The indirect utility of consumption for an individual with expenditure e facing local price of P_i takes the form:

$$C(e, P_i) = \frac{1}{\varepsilon}\left(\frac{e}{(P_i^a)^{\phi^A}(P_i^m)^{\phi^M}(P_i^s)^{\phi^s}}\right)^{\varepsilon} - \sum_{s\in\{A,M,S\}} \nu^s \ln P_i^s, \tag{5}$$

over sectoral value-added CES aggregates of varieties from all regions. I use $P_i \equiv (P_i^a)^{\phi^A}(P_i^m)^{\phi^M}(P_i^s)^{\phi^s}$ as local price index with $\sum_{s\in\{A,M,S\}} \phi^s = 1$. If $\nu^s = 0$ for all sectors and $\varepsilon = 1$, the consumption utility reduces to a Cobb-Douglas utility with consumption share ϕ^s allocated to each sector s. The income elasticity parameter, $\varepsilon \in (0,1)$, is interpreted as the Engel elasticity. The larger the Engel elasticity, the stronger the effect of real income on demand. As incomes grow to infinity, the consumption share on each good converges to $\phi^s \in (0,1)$ as a consumption asymptote.

By applying Roy's identity to the indirect utility function, the consumption share is given as:

$$\varphi^s(P_i, e) = \phi^s + \nu^s\left(\frac{e}{P_i}\right)^{-\varepsilon}. \tag{6}$$

An individual's consumption share depends on the price index in region i and her income. The consumption share on necessity declines as workers' real income rises. Workers do not have access to the investment technology and their labor income equals total expenditure in each period. Regional aggregate demand is derived by summing up individual demand in each location.

Intratemporal labor supply. In each period, factors of production are freely mobile across sectors within the region. To rationalize the observed difference in factor prices, I introduce Roy-type machinery by modeling that a worker has a different ability in each sector, as in Eckert and Peters (2023). Concretely, a worker supplies a_i^s efficiency units to sector s that are drawn from a sector-specific Frechet distribution with dispersion parameter ζ^L, $P(a_i^s \leq a) = \exp(-(a/A_i^{L,s})^{-\zeta^L})$. The level of $A_i^{L,s}$ represents the fundamental level of labor-augmenting technology.

Each worker chooses a sector by maximizing her income, and the employment share is:

$$s_i^{L,s} = (w_i^s/\bar{w}_i)^{\zeta^L} \text{ where } \bar{w}_i = \left((w_i^a)^{\zeta^L} + (w_i^m)^{\zeta^L} + (w_i^s)^{\zeta^L}\right)^{1/\zeta^L}, \tag{7}$$

where the dispersion parameter governs the sectoral labor supply. The measure of workers in sector s is then given as $L_i^s = s_i^s L_i$, and the effective unit of workers supplied can be written as:

$$\tilde{L}_i^s = \Gamma_{\zeta^L} A_i^{L,s} (w_i^s/\bar{w}_i)^{\zeta^L - 1} L_i. \tag{8}$$

I assume that labor efficiency consists of exogenous fundamental $\bar{A}$ and endogenous components $f(\cdot)$ from regional factor allocations:

$$A_{i,t}^{L,s} = \bar{A}_{i,t}^{L,s} \times (L_{i,t\text{-}1}^s, K_{i,t\text{-}1}^s), \tag{9}$$

where workers take this term as given. The endogenous component incorporates the dynamic weak equilibrium bias in a reduced form way (Section 2). It captures how the previous factor allocation influences the development of factor-augmenting technologies. For instance, it can be viewed as a function capturing the outcomes of learning-by-doing given the regional factor abundance as in Foster and Rosenzweig (1995). Alternatively, it can be thought of as generated by regional technology developers in shadow as in Acemoglu (2002, 2007). In contrast, the exogenous component $\bar{A}$ is not influenced by the changes in regional factor allocation.

Spatial mobility. Individuals make forward-looking decisions over which region n to live in the next period, taking the expected value of future utilities $\mathbb{V}$ and migration costs κ as given:

$$\mathbb{V}_{i,t}^s = U(C_{i,t}^L, B_{i,t}) + \max_{\{n\}} \left\{\beta \mathbb{E}[\mathbb{V}_{n,t+1}] - \kappa_{ni,t} + \eta u_{n,t}\right\}, \tag{10}$$

where the idiosyncratic preference shocks, u, follow the Type I extreme value distribution. The parameter η scales the variance of the shock and $1/\eta$ has an interpretation of migration elasticity.

The solution to the above dynamic problem yields the migration share proportional to the cost and elasticity-adjusted utility, compared to that of all other possible destinations. By expressing the expected value of the worker's value function as v, the migration share is:

$$\mathbb{M}_{in,t} = \frac{\exp\left(\beta\mathbb{E}_t(v_{n,t+1} - \kappa_{ni})/\eta\right)}{\sum_{j=1}^{N}\exp\left(\beta\mathbb{E}_t(v_{j,t+1} - \kappa_{nj,t})/\eta\right)}, \tag{11}$$

which, combined with the initial population, yields labor market distributions in the next period.

5.2.1 Capitalists

Capitalists' problem. Geographically immobile capitalists of measure zero in each location choose their consumption and investment to maximize the expected present value of their consumption utility, subject to their budget constraint:

$$v^K_{i,t} = \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \log(C^K_{i,t}) \text{ subject to } \bar{r}_{i,t}K_{i,t} = P_{i,t}\left(C^K_{i,t} + K_{i,t} - (1-\delta)K_{i,t}\right), \tag{12}$$

where the consumption takes the log utility form, and the superscript K indexes capitalists. I assume that capitalists consume the Cobb-Douglas composite of the three industries, $C^K_{i,t} \equiv (C^a_i)^{\phi^A} \times (C^m_i)^{\phi^M}(C^s_{in})^{\phi^s}$, with the CES aggregates of varieties at the lower level. Compared to the workers' problem, capitalists' consumption share equals the consumption asymptote. This simplification can be viewed as assuming that the capitalists have surpassed the income threshold to reach the asymptotes. An extension that allows the PIGL preference to capitalists is documented in Online Appendix Section C.[35]

Given log utility, the wealth and substitution effects cancel out, and capitalists save a fixed share of their investment income:

$$I_{i,t} = \frac{\beta}{1+\beta}\bar{r}_{i,t}K_{i,t}, \tag{13}$$

where $\bar{r}_{i,t}$ is the effective average net return on capital. The gross return on capital can be written as $\bar{R}_{i,t} \equiv 1 - \delta + \bar{r}_{i,t}/P_{i,t}$ with the depreciation rate δ. The capital is geographically immobile once installed and depreciates gradually at a constant rate δ. The investment goods combine goods from all sectors with the asymptotic consumption share.

Intratemporal capital supply. In each period, the regional capital stock is allocated across the

[35]However, such a complication does not allow a fixed consumption rate in this paper's setting. Instead, the path of consumption is governed by the Euler equation and transversality condition. A shooting algorithm can be adopted with initial values of capital, initial guesses of expenditure, and steady-state values of capital.

sectors, as in labor, by assuming the role of capital efficiency drawn from a Frechet distribution. The intratemporal capital allocation across sectors is given by the share $s^{K,s}_{i,t} = (R^s_{i,t}/\bar{R}_{i,t})^{\zeta^K}$, with the rental rate of capital $R^s_{i,t}$ and its effective average $\bar{R}_{i,t}$. The resulting effective capital for each sector is then given as $K^s_{i,t} = \Gamma_{\zeta^K} A^{K,s}_{i,t}(R^s_{i,t}/\bar{R}_{i,t})^{\zeta^K-1}K_{i,t}$. As in labor-augmenting technology, the fundamental efficiency consists of two components:

$$A^{K,s}_i = \bar{A}^{K,s}_i \times g(L^s_{i,t\text{-}1}, K^s_{i,t\text{-}1}), \tag{14}$$

where the endogenous component $g(\cdot)$ again depends on regional factor allocation.

5.3 Production.

In each region i, a representative local firm in each sector s uses the following CES technology:

$$Y^s_i = Z^s_i \left(\rho^s_i (A^{L,s}_i L^s_i)^{\frac{\sigma^s-1}{\sigma^s}} + (1-\rho^s_i)(A^{K,s}_i K^s_i)^{\frac{\sigma^s-1}{\sigma^s}}\right)^{\frac{\sigma^s}{\sigma^s-1}}. \tag{15}$$

The production side is given similarly to Section 2, with the addition of Hicks-neutral technology Z^s_i. This component is a realization of the Frechet distribution with local fundamental $\bar{Z}^s_i$ and the shape parameter θ^s, as in the standard Eaton and Kortum (2002) setting. The local fundamental $\bar{Z}^s_i$ is assumed not to be affected by the changes in factor allocation. The parameter σ^s governs the elasticity of substitution between factors, while ρ^s_i controls the region-specific labor intensity.

Trade and Market Clearing. The price of each industry k in importer n is determined as the minimum unit cost across all regions:

$$p^k_{n,t} = \min_{1\le i\le N}\left\{\frac{x^k_{i,t}\tau^k_{ni,t}}{Z^k_{i,t}}\right\}, \tag{16}$$

where the term inside Equation (16) is the factory-gate price of one unit of goods multiplied by the iceberg-type trade costs, $\tau^k_{ni,t}$. The unit cost function follows from the CES production structure:

$$x^s_{i,t} = \min\left(w^s_{i,t}A^{L,s}_{i,t}L^s_{i,t} + r^s_{i,t}A^{K,s}_{i,t}K^s_{i,t}\right) = \left((\rho^s_i)^{\sigma^s}(w^s_{i,t})^{1-\sigma^s} + (1-\rho^s_i)^{\sigma}(r^s_{i,t})^{1-\sigma^s}\right)^{\frac{1}{1-\sigma^s}}. \tag{17}$$

The changes in factor allocation and resulting changes in factor prices affect the regional comparative advantage through Equation (17).[36]

[36]As labor becomes more expensive, a sector that relies more on labor will lose its comparative advantage and

With the distributional assumption on productivity, the bilateral expenditure share is given by:

$$\mathbb{S}_{ni,t}^{s} = \frac{\left(x_{i,t}^{s}\tau_{ni,t}^{s}/\bar{Z}_{i,t}^{s}\right)^{-\theta^{s}}}{\sum_{j=1}^{N}\left(x_{j,t}^{s}\tau_{nj,t}^{s}/\bar{Z}_{j,t}^{s}\right)^{-\theta^{s}}}, \tag{18}$$

where the denominator can be interpreted as inward market access of region n. It also represents the sectoral price index up to a constant. Finally, the region-industry-level price indices are aggregated by consumption share asymptotes to yield the regional price index.

Goods market clearing implies that regional total expenditure, X_n, equals regional total income, I_n, where income consists of labor payments from workers and capital payments from capitalists. Market clearing conditions for the factor market are then given by $I_{n,t}^{k} = \sum_{i=1}^{N}\mathbb{S}_{ni,t}^{k}X_{n,t}^{k}$.

5.4 Taking the model to the data

Table 6 summarizes the model parameters. In terms of the dynamic spatial equilibrium setting, relatively distinctive features of the model is the introduction of CES production functions and factor-augmenting technologies. For related parameters, I capitalize on estimates in the literature and empirical findings in Section 4. First, I use the CES production function estimates in Herrendorf et al. (2015), who uses U.S. macro data between 1947 and 2010. The model here introduces the same production structure in the value-added form. The study periods align except for 1940-1946, when the required data is unavailable. Using their estimates, the values of σ^{s} for agriculture and two non-agriculture sectors are set to 1.58, 0.80, and 0.75, respectively. The values suggest that agriculture is flexible, while the non-agriculture sectors are inflexible in factor usage.

Herrendorf et al. (2015) also report the average production weights in capital, ρ^{s}.[37] However, the production weights are location-specific, and I need to take into account the regional differences in production practices. For instance, between 1940 and 1970, the number of tractors and combines per agricultural worker in the North was two to three times higher than in the South. However, a major issue is that there is no detailed information on the value of capital by sector in this period.

Instead, I adjust the capital intensity estimates by regional proxy for the capital-to-labor ratio. I use the variations in the number of tractors and combines per agricultural worker as a proxy for the regional differences in agricultural capital-to-labor ratio. For non-agriculture, I recover the values

experience a decrease production. It is the quasi-Rybczynski effect as in Romalis (2004).

[37] I exclude the land share in production to focus on the distinction between labor and physical capital.

of manufacturing capital stock using the perpetual inventory method and calculate the regional capital-to-labor ratio. I then apply these regional variations to adjust the labor intensity parameter ρ^s Admittedly, they impose heroic assumptions, and I report alternative results using the range of parameter values. The resulting weights imply that agriculture tended to be labor-intensive in the South but capital-intensive in the North, consistent with a historical account of regional differences in agricultural practices during the Great Migration period.

With the listed values of σ^s, the direction of technical changes would be biased toward capital in all industries from the out-migration shock (Section 2). Hence, I focus on constructing the capital efficiency functions. First, I newly estimate the measure of capital efficiency using output per capital stock as the dependent variables, with employment, capital, and their squared values as additional controls using equation (3). Table AX reports the results for agriculture and manufacturing, along with the estimates on employment and capital variables. They correspond to moment conditions in each year for the changes in factor allocation and the measure of capital efficiency. With the parallel trend assumption, I interpret the observed changes as the endogenous component of capital-augmenting technologies in each industry. The factor substitutability parameters are taken from the labor dispersion parameter from Eckert and Peters (2023), estimated using 1880-1920 U.S. data. Finally, I calibrate the capital depreciation rate and trade elasticities by following Hulten and Wykoff (1981)[38] and Nigai (2016).

The consumption side and migration estimates follow Yang (2024),[39] where I estimate the PIGL parameters and migration elasticity in the 20th Century U.S. setting. The values of the preference elasticities imply that agriculture is a necessity and two non-agriculture sectors are luxuries, with nontradable non-agriculture having the highest income elasticity. Within non-agriculture, the tradable sector is closer to normal goods. The baseline migration elasticity over a 10-year span is estimated to be $\eta = 0.84$, which suggests a higher migration response compared to 1.88 at the annual frequency (Artuc et al., 2010) and 5.34 at quarterly (Caliendo et al., 2019). I set the discount rate β for the decennial interval to be 0.67, assuming a yearly discount rate of 4%.

5.5 Baseline economy

For the baseline counterfactual analysis, I fix the migration flows during the Second Great Migration as given and use them as a shock. Specifically, I use the migration flows between 1940 and

[38] They estimate the average depreciation rate of structures to be 3.7%, with a range of 1.9% to 5.6%.

[39] In this paper, I study the spillover effects of the Dust Bowl through a standard dynamic spatial equilibrium model.

Table 6: Parameters for Quantitative Analysis.

Definition	Parameter	Comment
Panel (A) Productivity parameters		
EIS between labor and capital	$\sigma = (1.58, 0.80, 0.75)$	Herrendorf et al. (2015)
Average capital weights in production	$\bar{\rho} = (0.53, 0.29, 0.34)$	Herrendorf et al. (2015)
Regional Capital efficiency (agriculture)	$g^A(\cdot) = \exp\left((-0.85)\hat{L}_{it}^A + (0.65)\hat{K}_{it}^A\right)$	Moment condition
Regional Capital efficiency (non-agriculture)	$g^M(\cdot) = \exp\left((1.23)\hat{L}_{it}^M + (-0.70)\hat{K}_{it}^M\right)$	Moment condition
Labor and capital substitutability	$(\zeta^L, \zeta^K) = (6.9, 6.9)$	Eckert and Peters (2023)
Hicks-neutral productivity distribution	$\theta = (12, 6.5, \infty)$	Nigai (2016)
Capital depreciation rate	$\delta = 0.34$	Hulten and Wykoff (1981)
Panel (B) Utility parameters		
Asymptotic consumption share	ϕ = (0.01, 0.33, 0.66)	Moment condition
Preference elasticity	ν = (1.27, -0.27, -1.0)	Moment condition
Engel elasticity	$\eta = 0.39$	Estimation
Migration elasticity	$\chi = 0.83$	Estimation
Discount rate	$\beta = 0.67$	Set to $(0.96)^{10}$

1970 that are explained by the Northern pull factors as the baseline shock. I add the predicted out-migration rate back to its origin and restrict the migration from the South to the North. I still allow endogenous migration between 1970 and 2010. To close the model, I add an additional period that corresponds to the year 2020, where the economy is assumed to reach a steady state.

The model calculation adopts the dynamic exact hat algebra approach (Online Appendix Section 3). The method calculates the time changes in economic allocation given the shock in time change. The use of the time change terms annihilates the need to recover the majority of the time-invariant components of the model because they are canceled out during calculation. I first run the quantitative model without the shock to calculate a baseline economy with the actual history. I then calculate a counterfactual economy in the absence of the Great Migration, represented as the predicted value of the out-migration rate. The differences between the two are interpreted as the impacts of the Second Great Migration. Unless mentioned otherwise, the reported effects measure the outcomes in 1970, the end period of the Great Migration.

6 Quantitative Results

In this section, I first outline counterfactual outcomes in terms of welfare effects and introduce contribution analysis. Here, I focus on the changes in the workers' consumption welfare (Equation 4), and I refer to it simply as welfare.[40] Although this section does not directly report the capitalists' welfare, I discuss it in terms of changes in capital rents. Finally, I show how the Great Migration shaped economic distributions through time and geography.

6.1 Welfare effects and contribution analysis

The baseline counterfactual analysis shows that the Second Great Migration increased the United States consumption welfare by 0.5% per capita by 1970. The South experienced a gain of 2.28%, while the North a loss of 1.01%. Table 7 reports the baseline welfare effect for the contiguous U.S. and the contribution of each model element in generating welfare. I examine the contribution of factor substitution, trade adjustment, and directed technical change in response to the Second Great Migration. Specifically, I use consumption welfare as the criteria since it is the main outcome of the model, summarizing all functions and interactions of the model elements.[41]

Table 7, Row A, reports the welfare effect for the baseline and restricted scenarios. As a benchmark, I run a scenario where all three adjustment mechanisms are held fixed. In other words, I fix the share of labor and capital allocated to each industry, trade share, and the level of factor-augmenting technology to the baseline level in the absence of the Great Migration. Column 5 reports that the fully restricted model yields a welfare effect of -3.0% from the Great Migration. I then run a constrained model separately for each channel by turning off one model component at a time. The difference in welfare between the baseline model and the constrained model, divided by

[40] The consumption welfare effect measures the changes in real income with non-homothetic adjustments.

[41] Although this process is not a formal decomposition analysis, the exercise provides a useful gauge of which model element is driving the simulation results. For instance, Chor (2010) conducts a similar contribution analysis using how the welfare effects change by turning off each model component to quantify the relative importance of different sources of comparative advantage.

Table 7: Welfare effect and the contribution of each model element.

	(1) Baseline results	(2) No factor substitution	(3) No trade adjustment	(4) No directed technical change	(5) Without all adjustments
A. Consumption welfare effect	+0.49%	-1.92%	0.08%	-0.09%	-2.99%
B. Contribution of each channel	-	[69.3%]	[11.8%]	[16.7%]	[100%]

Note: This table shows the consumption welfare effect by scenario in Row A and the contribution of each model element in Row B. The baseline analysis quantifies the impact of the Great Migration using the predicted out-migration rate using equation (3). The contribution analysis compares the difference between the welfare effects of the full model and a constrained version by turning off each model component. The residual welfare effects that are not taken into account by the constrained models are denoted as the interaction effect.

the difference in welfare between the baseline model and the fully restricted model, is interpreted as the contribution of each model element.

Row B, Columns 2 to 4, presents the contribution analysis results. As shown in Column 2, factor substitution takes into account the lion's share of the response to the Great Migration, driving 69.3% of the adjustments. The trade adjustment and directed technical change played supplementary roles, and each contributed to 11.8% and 16.7%. The remaining 2.2% is generated by the interaction effects.

Here, the trade adjustment measures how the changes in trade share alleviated the potential welfare loss. Given that changes in trade share are driven by the changes in relative factor prices, the trade adjustment channel can be interpreted as capturing the quasi-Rybczynski effect. Given the capital intensity estimates (Herrendorf et al., 2015) and capital-labor ratio adjustment, the differences in factor intensity and resulting Hecksher-Ohlin forces played a less pivotal role.

6.2 Distribution of the shock

Figure 7 displays the geographical distribution of the shock and model outcomes. The map highlights the relationship between the shock and the main mechanism captured through the model. Panel A shows the predicted out-migration rate to the North, the shock used in the counterfactual analysis. Red means higher levels of predicted out-migration relative to yellow. The Northern states are denoted as grey. The Deep Southern states and Oklahoma experienced the highest level of Northern pull factor induced migration to the North, while Texas and Florida's average out-migration rates are predicted to be relatively low.

Panels B to D plot the changes in labor wage, capital rents, and capital-to-labor ratio in 1970. The effect size is defined as the percentage point changes in the outcomes due to the Great Migration. Panel B shows the changes in aggregate capital-to-labor ratio. It contains the mechanical change from the labor decrease from the Great Migration and the endogenous response from capital accumulation. The Southern states increased in capital more relative to labor, while the opposite pattern held for the North, especially for the states that received larger migrants, such as California. The absolute value of the correlation between the wage changes and predicted out-migration is 0.15. While labor scarcity drove the model outcomes, the interaction between labor scarcity and the stated model elements determined the exact results.

Panels C and D report the impacts of the Great Migration on factor prices. Red indicates a decrease, while blue means an increase, with a darker color representing a larger absolute size.

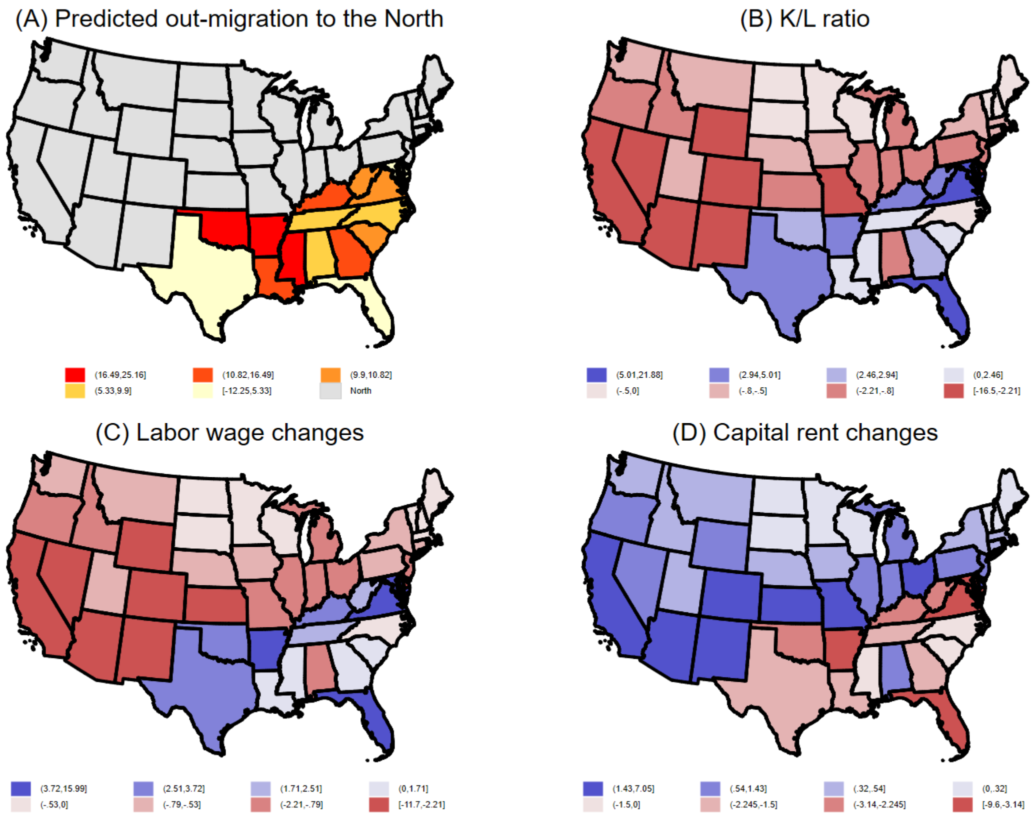

Figure 7: Maps of the Great Migration and outcome variables.

Note: The figure shows the geographic distributions of the shock and outcome variables in the contiguous United States. Panel A plots the predicted out-migration during 1940 and 1970 using equation (3). Red means higher levels of predicted out-migration relative to yellow. Panels B to D map the simulation outcome evaluated in the year 1970 for the aggregate capital-to-labor ratio, wage of labor, and rental rate of capital. Red indicates a decrease, while blue means an increase, with a darker color representing a larger absolute size. The predicted out-migration between 1940 and 1970 is used as a shock that represents the Second Great Migration.

Compared to patterns in Panel B, Panel C documents a similar distribution of wage changes, and Panel D shows the opposite pattern in terms of the rental rate of capital. The region that accumulated more capital relative to labor tended to increase wages but decrease the rental rate of capital.

Next, Figure 8 plots the changes in economic distribution between 1940 and 2010. It shows the changes in the share of labor (Panel 1), capital (Panel 2), and consumption spending (Panel 3) allocated to agriculture and non-agriculture. The Figure separately reports the results for the South (Row A) and the North (Row B). The non-agriculture is further separated by tradability.

Panel 1, Row A, suggests that the Great Migration led to a structural change in labor allocation. The relative labor scarcity incentivized the flexible sector, agriculture, to substitute labor with capital. Concretely, the model calculates that the Great Migration and the following adjustment

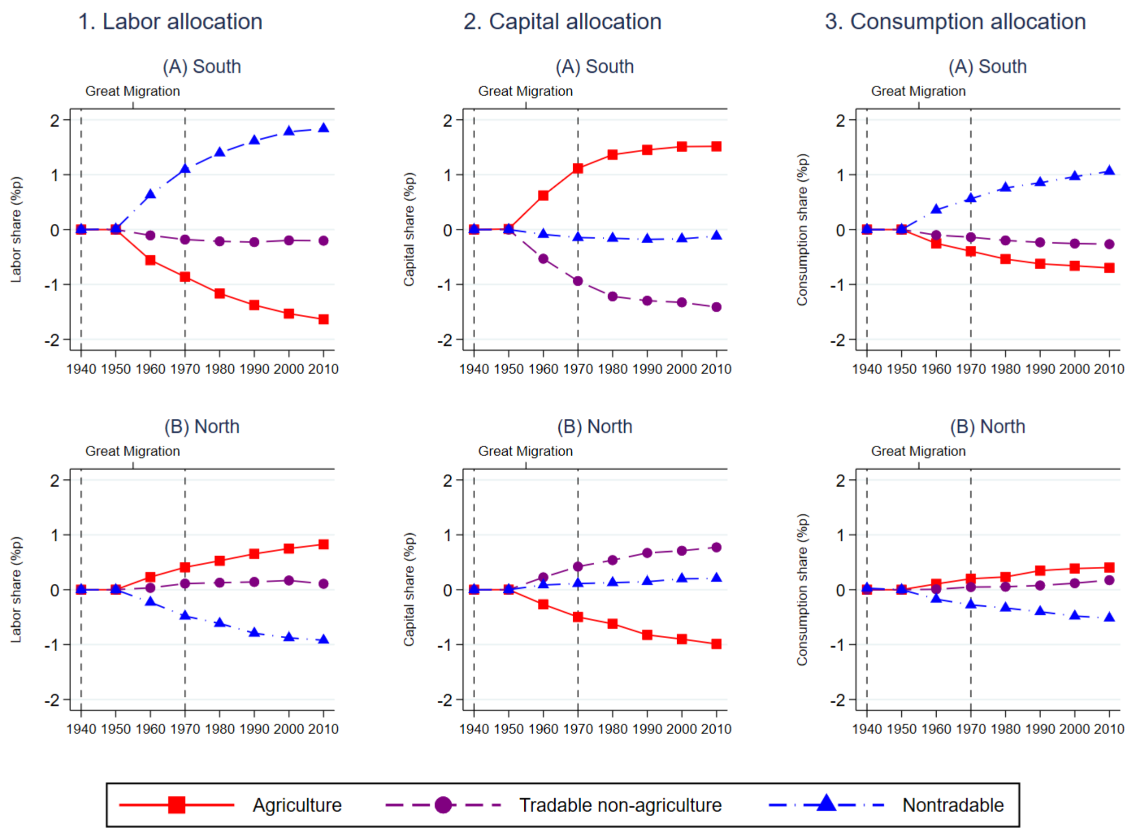

Figure 8: Simulated Changes in Economic Allocation by Region, 1940-2010.

Note: The figure shows the time trends in sectoral allocation of labor, capital, and consumption between 1940 and 2010. The effect sizes are defined as the percentage point changes in outcomes in each region. The red straight line represents agriculture, and the dashed line indicates non-agriculture. Tradable and nontradable non-agriculture are denoted purple and blue, respectively. The predicted out-migration between 1940 and 1970 is used as a shock that represents the Second Great Migration.

decreased agricultural employment share by around 2% (Panel 1, Row A). Such a decrease constituted around 7% of the total decrease in agricultural employment during the study period, given that the agricultural employment share decreased from 30% to 2% in the South between 1940 and 2010. Hence, the Great Migration played a supplementary yet important role in reallocating labor out of agriculture. On the contrary, the share of capital allocated to agriculture increased (Panel 2, Row A).

In the long run, agricultural labor has been reallocated to the nontradable sector, driven by the income effects depicted in Panel 3. The relative increase in capital raised the workers' income levels and, subsequently, increased the consumption share allocated to luxury goods provided by local nontradables. Hence, the Great Migration could have also contributed to the development of the local nontradable sector in the long run.

7 Conclusion

In 2019, around 1.2 billion people worked in the agricultural sector globally, constituting approximately 28% of the employed population, with a significant portion residing in rural areas in low-income countries (Davis et al., 2023). Still to this day, the agricultural sector in developing countries is characterized by low labor productivity compared to the non-agricultural counterpart, due to ineffective use of labor (Gollin et al., 2014), land (Adamopoulos and Restuccia, 2014; Chen et al., 2023), as well as capital and intermediate inputs (Gollin and Udry, 2021; Foster and Rosenzweig, 2022; Boppart et al., 2023). Hence, policies that can facilitate agriculture productivity improvement and structural change out of agriculture could yield substantial gains.

This paper proposes a new perspective on the economic development of the American South during the 20th century by focusing on the role of labor scarcity in inducing capital accumulation and capital-augmenting technical change. In response to the out-migration, flexible agriculture substituted labor with capital, while open economy forces depressed the size of labor-intensive agriculture. The following labor-capital reallocation induced structural change out of agriculture, expansion of non-agriculture, and capital-biased technical change. The core lesson here is that regional out-migration can induce factor reallocation to facilitate economic development.

However, physical capital accumulation could have paradoxically limited the extent of human capital improvements, as the increase in the price of labor reduces the incentive to invest for future wage increases. Given the continued rise in the importance of human capital relative to physical capital (Goldin and Katz, 2008), a supplementary policy on human capital may also be beneficial.

References

Abramitzky, Ran, Philipp Ager, Leah Boustan, Elior Cohen, and Casper W. Hansen, "The Effect of Immigration Restrictions on Local Labor Markets: Lessons from the 1920s Border Closure," *American Economic Journal: Applied Economics*, 2023, *15* (1), 164–91.

Acemoglu, Daron, "Directed Technical Change," *The Review of Economic Studies*, 10 2002, *69* (4), 781–809.

—, "Equilibrium Bias of Technology," *Econometrica*, 2007, *75* (5), 1371–1409.

—, "When Does Labor Scarcity Encourage Innovation?," *Journal of Political Economy*, 2010, *118* (6), 1037 – 1078.

— **and Veronica Guerrieri**, "Capital deepening and nonbalanced economic growth," *Journal of Political Economy*, 2008, *116* (3), 467–498.

Adamopoulos, Tasso and Diego Restuccia, "The Size Distribution of Farms and International Productivity Differences," *American Economic Review*, June 2014, *104* (6), 1667–97.

Akram, Agha Ali, Shyamal Chowdhury, and Ahmed Mushfiq Mobarak, "Effects of Emigration on Rural Labor Markets," *Unpublished Manuscript*, 2018.

Alonso-Carrera, Jaime and Xavier Raurich, "Labor mobility, structural change and economic growth," *Journal of Macroeconomics*, 2018, *56* (April), 292–310.

Alvarez-Cuadrado, Francisco and Markus Poschke, "Structural change out of agriculture: Labor push versus labor pull," *American Economic Journal: Macroeconomics*, 2011, *3* (3), 127–158.

_ , Ngo Van Long, and Markus Poschke, "Capital-labor substitution, structural change, and growth," *Theoretical Economics*, 2017, *12* (3), 1229–1266.

Andersson, David, Mounir Karadja, and Erik Prawitz, "Mass Migration and Technological Change," *Journal of the European Economic Association*, 05 2022, *20* (5), 1859–1896.

Arrow, Kenneth J., "The Economic Implications of Learning by Doing," *The Review of Economic Studies*, 1962, *29* (3), 155–173.

_ , Hollis B. Chenery, B. S. Minhas, and Robert M. Solow, "Capital-Labor Substitution and Economic Efficiency," *The Review of Economics and Statistics*, 1961, *43* (3), 225–250.

Artuc, Erhan, Shubham Chaudhuri, and John McLaren, "Trade Shocks and Labor Adjustment: A Structural Empirical Approach," *American Economic Review*, 2010, *100* (3), 1008–1045.

Bateman, Fred and Thomas Weiss, "A Deplorable Scarcity: The Failure of Industrialization in the Slave Economy.," 1981.

Batistich, Mary Kate and Timothy N Bond, "Stalled Racial Progress and Japanese Trade in the 1970s and 1980s," *The Review of Economic Studies*, 01 2023, *90* (6), 2792–2821.

Baumol, William J., "Macroeconomics of Unbalanced Growth: The Anatomy of Urban Crisis," *The American Economic Review*, 1967, *57* (3), 415–426.

Bazzi, Samuel, Andreas Ferrara, Martin Fiszbein, Thomas Pearson, and Patrick A Testa, "The Other Great Migration: Southern Whites and the New Right," *The Quarterly Journal of Economics*, 03 2023, *138* (3), 1577–1647.

Bleakley, Hoyt, "Disease and Development: Evidence from Hookworm Eradication in the American South," *The Quarterly Journal of Economics*, 2007, *122* (1), 73–117.

Boppart, Timo, "Structural Change and the Kaldor Facts in a Growth Model With Relative Price Effects and Non-Gorman Preferences," *Econometrica*, 2014, *82* (6), 2167–2196.

_ , Patrick Kiernan, Per Krusell, and Hannes Malmberg, "The Macroeconomics of Intensive Agriculture," *Unpublished Manuscript*, 2023.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," *Review of Economic Studies*, 2022, *89* (1), 181–213.

Boustan, Leah, *Competition in the Promised Land: Black Migrants in Northern Cities and Labor Markets*, Princeton: Princeton University Press, 2016.

Boustan, Leah Platt, "Was postwar suburbanization "white flight"'? Evidence from the black migration," *Quarterly Journal of Economics*, 2010, *125*, 417–443.

Bowles, Gladys K., James D. Tarver, Calvin L. Beale, and Everette S. Lee, "Net Migration of the Population by Age, Sex, and Race, 1950-1970," 2016. ICPSR08493-v2.

Bryan, Gharad and Melanie Morten, "The aggregate productivity effects of internal migration: Evidence from Indonesia," *Journal of Political Economy*, 2019, *127* (5), 2229–2268.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley Wilbert, "Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project," 2023. Working Paper.

Burstein, Ariel and Jonathan Vogel, "International Trade, Technology, and the Skill Premium," *Journal of Political Economy*, 2017, *125* (5), 1356–1412.

Bustos, Paula, Gabriel Garber, and Jacopo Ponticelli, "Capital Accumulation and Structural Transformation," *The Quarterly Journal of Economics*, 01 2020, *135* (2), 1037–1094.

Caballero, Maráa Esther, Brian Cadena, and Brian K Kovak, "The International Transmission of Local Economic Shocks Through Migrant Networks," *Journal of International Economics*, 2023, *145.*

Caliendo, Lorenzo, Maximiliano Dvorkin, and Fernando Parro, "Trade and Labor Market Dynamics: General Equilibrium Analysis of the China Trade Shock," *Econometrica*, 2019, *87* (3), 741–835.

Card, David, "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration," *Journal of Labor Economics*, 01 2001, *19* (1), 22–64.

Caron, Justin, Thibault Fally, and James R. Markusen, "International Trade Puzzles: A Solution Linking Production and Preferences," *The Quarterly Journal of Economics*, 2014, *129* (3), 1501–1552.

Carter, Susan B., Scott Sigmund Gartner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright, "Farm Employment: 1910–1999," 2006.

Caselli, Francesco, *Technology Differences over Space and Time.* CREI Lectures in Macroeconomics, MIT Press, 2016.

_ **and James Feyrer**, "The Marginal Product of Capital," *The Quarterly Journal of Economics*, 2007, *122* (2), 535–568.

_ **and Wilbur John Coleman II**, "The U.S. Structural Transformation and Regional Convergence: A Reinterpretation," *Journal of Political Economy*, 2001, *109* (3), 584–616.

Caunedo, Julieta and Elisa Keller, "Capital-Embodied Structural Change," *Unpublished Manuscript*, 2024.

Census Bureau, "County and City Data Book [United States] Consolidated File: County Data, 1947-1977," September 2012.

—, "County Business Patterns," 2023. Accessed: 2023-05-21.

Chapel, Jack and Yi-Ju Hung, "The Great Migration's Impact on Southern Inequality," *Unpublished Manuscript*, 2024.

Chen, Chaoran, Diego Restuccia, and Raul Santaeulàlia-Llopis, "Land Misallocation and Productivity," *American Economic Journal: Macroeconomics*, April 2023, *15* (2), 441–65.

Chor, David, "Unpacking sources of comparative advantage: A quantitative approach," *Journal of International Economics*, 2010, *82* (2), 152–167.

Clay, Karen, Ethan J Schmick, and Werner Troesken, "The Boll Weevil's Impact on Racial Income Gaps in the Early Twentieth Century," *NBER Working Paper No. 27101*, 2020.

Clemens, Michael A., Ethan G. Lewis, and Hannah M. Postel, "Immigration Restrictions as Active Labor Market Policy: Evidence from the Mexican Bracero Exclusion," *American Economic Review*, June 2018, *108* (6), 1468–87.

Collins, William J. and Marianne H. Wanamaker, "The Great Migration in Black and White: New Evidence on the Selection and Sorting of Southern Migrants," *The Journal of Economic History*, 2015, *75* (4), 947–992.

— **and** —, "African American Intergenerational Economic Mobility since 1880," *American Economic Journal: Applied Economics*, July 2022, *14* (3), 84–117.

Conley, T.G., "GMM estimation with cross sectional dependence," *Journal of Econometrics*, 1999, *92* (1), 1–45.

Davis, Benjamin, Esin Mane, Lale Yalçın Gurbuzer, Giulia Caivano, Natalia Piedrahita, Karen Schneider, Nabila Azhar, Mahdi Benali, Naveed Chaudhary, Riad Rivera, Ramanan Ambikapathi, and Paul Winters, "Estimating global and country-level employment in agrifood systems," Technical Report 23-34, Food and Agriculture Organization of the United Nations, Rome 2023.

Day, Richard H., "The Economics of Technological Change and the Demise of the Sharecropper," *American Economic Review*, 1967, *57* (3), 427–449.

Depew, Briggs, Price V. Fishback, and Paul W. Rhode, "New deal or no deal in the Cotton South: The effect of the AAA on the agricultural labor structure," *Explorations in Economic History*, 2013, *50* (4), 466–486.

Derenoncourt, Ellora, "Can You Move to Opportunity? Evidence from the Great Migration," *American Economic Review*, 2022, *112* (2), 369–408.

Dinkelman, T, G Kumchulesi, and M Mariotti, "Labor Migration and the Structure of Rural Labor Markets," *Review of Economics and Statistics*, Forthcoming.

Duernecker, Georg, Berthold Herrendorf, and Ákos Valentinyi, "Structural Change within the Services Sector and the Future of Cost Disease," *Journal of the European Economic Association*, 05 2023, *22* (1), 428–473.

Eaton, Jonathan and Samuel Kortum, "Technology, Geography, and Trade," *Econometrica*, 2002, *70* (5), 1741–1779.

Eckert, Fabian and Michael Peters, "Spatial Structural Change," *NBER Working Paper Series 30489*, 2023.

_ , Andrés Gvirtz, Jack Liang, and Michael Peters, "A Method to Construct Geographical Crosswalks with an Application to Us Counties since 1790," *Unpublished Manuscript*, 2020.

_ , Ka leung Lam, Atif Mian, Karsten Muller, Rafael Schwalb, and Amir Sufi, "The Early County Business Pattern Files: 1946-1974," Technical Report, National Bureau of Economic Research 2022.

Egger, Dennis, Benjamin Faber, Ming Li, and Wei Lin, "Rural-Urban Migration and Market Integration," *Unpublished Manuscript*, 2024.

Fajgelbaum, Pablo and Stephen J. Redding, "Trade Structural Transformation and Development: Evidence from Argentina 1869-1914," *Journal of Political Economy*, 2018, *130* (5), 1249–1318.

Fan, Tianyu, Michael Peters, and Fabrizio Zilibotti, "Growing Like India—the Unequal Effects of Service-Led Growth," *Econometrica*, 2023, *91* (4), 1457–1494.

Feigenbaum, James J, Soumyajit Mazumder, and Cory B Smith, "When Coercive Economies Fail: The Political Economy of the US South After the Boll Weevil," Working Paper 27161, National Bureau of Economic Research 2020.

Ferleger, Louis, "Sharecropping contracts in the late-nineteenth-century South," *Agricultural History*, 1993, *67* (3), 31–46.

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

Foster, Andrew D. and Mark R. Rosenzweig, "Learning by Doing and Learning from Others: Human Capital and Technical Change in Agriculture," *Journal of Political Economy*, 1995, *103* (6), 1176–1209.

_ and Mark Rosenzweig, "Are There Too Many Farms in the World? Labor Market Transaction Costs, Machine Capacities, and Optimal Farm Size," *Journal of Political Economy*, 2022, *130* (3), 636–680.

Fuguitt, Glenn V., Calvin L. Beale, and Paul R. Voss, “County-Specific Net Migration Estimates, 1980-1990 [United States],” 2010. ICPSR26761-v1.

Gallardo, R Karina and Johannes Sauer, “Adoption of Labor-Saving Technologies in Agriculture,” *Annual Review of Resource Economics*, 2018, *10*, 185–206.

Gardner, John and William Cohen, “Demographic Characteristics of the Population of the United States, 1930-1950: County-Level,” 1992.

Goldin, Claudia and Lawrence F. Katz, *The Race Between Education and Technology*, Harvard University Press, 2008.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, “Bartik Instruments: What, When, Why, and How,” *American Economic Review*, 2020, *110* (8), 2586–2624.

Gollin, Douglas and Christopher Udry, “Heterogeneity, Measurement Error, and Misallocation: Evidence from African Agriculture,” *Journal of Political Economy*, 2021, *129* (1), 1–80.

_ , David Lagakos, and Michael E. Waugh, “The agricultural productivity gap,” *Quarterly Journal of Economics*, 2014, *129* (2), 939–993.

Gregory, James N., *The Southern Diaspora: How the Great Migrations of Black and White Southerners Transformed America*, Chapel Hill: University of North Carolina Press, 2005.

Gross, Daniel P., “Scale versus scope in the diffusion of new technology: evidence from the farm tractor,” *RAND Journal of Economics*, 2018, *49* (2), 427–452.

Grove, Wayne A. and Craig Heinicke, “Better Opportunities or Worse? The Demise of Cotton Harvest Labor, 1949-1964,” *The Journal of Economic History*, September 2003, *63* (3), 736–767.

Habakkuk, H. J., *American and British Technology in the Nineteenth Century: The Search for Labor-Saving Inventions*, Cambridge, UK: Cambridge University Press, 1962.

Haines, Michael, Price Fishback, and Paul Rhode, “United States Agriculture Data, 1840 - 2012,” 2018. Inter-university Consortium for Political and Social Research [distributor].

Haines, Michael R., Inter university Consortium for Political, and Social Research, “Historical, Demographic, Economic, and Social Data: The United States, 1790-2002,” 2010. Version 3.

Harris, John R. and Michael P. Todaro, “Migration, Unemployment and Development: A Two-Sector Analysis,” *The American Economic Review*, 1970, *60* (1), 126–142.

Herrendorf, Berthold, Christopher Herrington, and Ákos Valentinyi, “Sectoral Technology and Structural Transformation,” *American Economic Journal: Macroeconomics*, October 2015, *7* (4), 104–33.

Hicks, John R., *The Theory of Wages*, London: Macmillan, 1932.

Holley, Donald, *The Second Great Emancipation: The Mechanical Cotton Picker, Black Migration, and How They Shaped the Modern South*, Fayetteville, AR: University of Arkansas Press, 2000.

Hornbeck, Richard and Suresh Naidu, "When the Levee Breaks: Black Migration and Economic Development in the American South," *American Economic Review*, March 2014, *104* (3), 963–90.

Hsieh, Chang-Tai and Peter J. Klenow, "Misallocation and Manufacturing TFP in China and India," *The Quarterly Journal of Economics*, 11 2009, *124* (4), 1403–1448.

Hulten, Charles R. and Frank C. Wykoff, "The Estimation of Economic Depreciation Using Vintage Asset Prices: An Application of the Box-Cox Power Transformation," *Journal of Econometrics*, 1981, *15* (3), 367–396.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," NBER Working Paper 24285, National Bureau of Economic Research February 2018.

Jaworski, Taylor and Dongkyu Yang, "Did War Mobilization Cause Regional Development?," *Unpublished Manuscript*, 2024.

Jung, Yeonha, "The long reach of cotton in the US South: Tenant farming, mechanization, and low-skill manufacturing," *Journal of Development Economics*, 2020, *143*, 102432.

Kennedy, Charles, "Induced Bias in Innovation and the Theory of Distribution," *The Economic Journal*, 09 1964, *74* (295), 541–547.

Kirby, Jack Temple, "The Southern Exodus, 1910-1960: A Primer for Historians," *The Journal of Southern History*, Nov 1983, *49* (4), 585–600.

Kleinman, Benny, Ernest Liu, and Stephen J. Redding, "Dynamic Spatial General Equilibrium," *Econometrica*, 2023, *91* (2), 385–424.

Kline, Patrick and Enrico Moretti, " Local Economic Development, Agglomeration Economies, and the Big Push: 100 Years of Evidence from the Tennessee Valley Authority," *The Quarterly Journal of Economics*, 2014, *129* (1), 275–331.

Lafortune, Jeanne, José Tessada, and Carolina González-Velosa, "More hands, more power? Estimating the impact of immigration on output and technology choices using early 20th century US agriculture," *Journal of International Economics*, 2015, *97* (2), 339–358.

Lagakos, David, Ahmed Mushfiq Mobarak, and Michael E. Waugh, "The Welfare Effects of Encouraging Rural–Urban Migration," *Econometrica*, 2023, *91* (3), 803–837.

Lewis, W. Arthur, "Economic Development with Unlimited Supplies of Labour," *The Manchester School*, 1954, *22* (2), 139–191.

León-Ledesma, Miguel A., Peter McAdam, and Alpo Willman, "Identifying the Elasticity of Substitution with Biased Technical Change," *American Economic Review*, September 2010, *100* (4), 1330–57.

Manuelli, Rodolfo E. and Ananth Seshadri, "Frictionless Technology Diffusion: The Case of Tractors," *American Economic Review*, April 2014, *104* (4), 1368–91.

Michaels, Guy, Ferdinand Rauch, and Stephen J. Redding, "Urbanization and structural transformation," *Quarterly Journal of Economics*, 2012, *127* (2), 535–586.

Montrose, Paige, "Without Deliberate Speed: The Effects of Southern Out-Migration on School Desegregation," *Unpublished Manuscript*, 2024.

Ngai, L. Rachel and Christopher A. Pissarides, "Structural change in a multisector model of growth," *American Economic Review*, 2007, *97* (1), 429–443.

Nigai, Sergey, "On Measuring the Welfare Gains from Trade Under Consumer Heterogeneity," *Economic Journal*, June 2016, *126* (593), 1193–1237.

Oberfield, Ezra and Devesh Raval, "Micro Data and Macro Technology," *Econometrica*, 2021, *89* (2), 703–732.

Olea, José Luis Montiel and Carolin Pflueger, "A Robust Test for Weak Instruments," *Journal of Business Economic Statistics*, 2013, *31* (3), 358–369.

Olmstead, Alan L. and Paul W. Rhode, " Reshaping the Landscape: the Impact and Diffusion of the Tractor in American Agriculture, 1910-1960 ," *The Journal of Economic History*, 2001, *61* (3), 663–698.

Peterson, Willis and Yoav Kislev, "The Cotton Harvester in Retrospect: Labor Displacement or Replacement?," *The Journal of Economic History*, 1986, *46* (1), 199–216.

Porzio, Tommaso, Federico Rossi, and Gabriella Santangelo, "The Human Side of Structural Transformation," *American Economic Review*, 2022, *112* (8), 2774–2814.

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

Romalis, John, "Factor proportions and the structure of commodity trade," *American Economic Review*, 2004, *94* (1), 67–97.

Ruggles, Steven, Matt A. Nelson, Matthew Sobek, Catherine A. Fitch, Ronald Goeken, J. David Hacker, Evan Roberts, and J. Robert Warren, "IPUMS Ancestry Full Count Data: Version 4.0," dataset 2024.

_ , Sarah Flood, Matthew Sobek, Daniel Backman, Annie Chen, Grace Cooper, Stephanie Richards, Renae Rodgers, and Megan Schouweiler, "IPUMS USA: Version 15.0," dataset 2024.

Rybczynski, T. M., "Factor Proportions and Comparative Advantage," *Economica*, 1955, *22* (88), 336–341.

San, Shmuel, "Labor Supply and Directed Technical Change: Evidence from the Termination of the Bracero Program in 1964," *American Economic Journal: Applied Economics*, January 2023, *15* (1), 136–63.

Schulman, Bruce J., *From Cotton Belt to Sunbelt: Federal Policy, Economic Development, and the Transformation of the South, 1938-1980*, Durham, NC: Duke University Press, 1994.

Theoharides, Caroline, "Manila to Malaysia, Quezon to Qatar: International Migration and its Effects on Origin-Country Human Capita," *Journal of Human Resources*, 2018, *53* (4), 1022–1049.

Tombe, Trevor and Xiaodong Zhu, "Trade , Migration , and Productivity : A Quantitative Analysis of China," *American Economic Review*, 2019, *109* (5), 1843–1872.

Voss, Paul R., Scott McNiven, Roger B. Hammer, Kenneth M. Johnson, and Glenn V. Fuguitt, "County-Specific Net Migration by Five-Year Age Groups, Hispanic Origin, Race, and Sex, 1990-2000: [United States]," 2005. ICPSR04171-v1.

Whatley, Warren C, "A History of Mechanization in the Cotton South: The Institutional Hypothesis," *The Quarterly Journal of Economics*, 1985, *100* (4), 1191–1215.

Winkler, Richelle, Kenneth Johnson, Cheng Cheng, Paul Voss, and Katherine J. Curtis, "County-Specific Net Migration by Five-Year Age Groups, Hispanic Origin, Race, and Sex: 2000-2010," 2013. ICPSR34638-v1.

Wright, Gavin, *Old South, New South: Revolutions in the Southern Economy since the Civil War*, New York: Basic Books, 1986.

Yang, Dongkyu, "A General Equilibrium Investigation of the American Dust Bowl," *Unpublished Manuscript*, 2024.

Appendix.

Table A1: Zero-stage in-migration prediction.

	(1) Black in-migration rate		(2) White in-migration rate	
Latitude	-1.967*	(0.971)	-0.316***	(0.092)
Longitude	-1.607***	(0.476)	0.034	(0.045)
Log population	58.642***	(8.790)	-7.388***	(0.834)
Log white population	-45.825***	(8.613)	8.138***	(0.818)
Log black population	-26.708***	(0.971)	0.290**	(0.092)
Urbanization	-0.379***	(0.072)	-0.020**	(0.007)
Non-white share	-0.346*	(0.168)	0.035*	(0.016)
Median income	-3.446	(1.766)	1.432***	(0.168)
Log housing units	8.291***	(1.397)	-0.994***	(0.133)
Median rent	-0.122	(1.814)	2.727***	(0.172)
Share foreign (1940)	1.300	(0.976)	-0.293**	(0.093)
Share employment (1940)	0.637	(0.440)	0.342***	(0.042)
Occupational score (1940)	-0.020	(0.033)	-0.010**	(0.003)
Republican vote share (1944)	-0.300*	(0.118)	-0.017	(0.011)
Republican vote share (1948)	-0.499***	(0.121)	0.072***	(0.011)
Republican vote share (1952)	0.784***	(0.169)	-0.078***	(0.016)
Republican vote share (1956)	-0.201	(0.155)	0.076***	(0.015)
Republican vote share (1960)	-0.027	(0.129)	0.054***	(0.012)
Republican vote share (1964)	0.025	(0.110)	-0.059***	(0.010)
Republican vote share (1968)	0.707***	(0.158)	0.008	(0.015)
Republican vote share (1972)	-0.006	(0.036)	-0.004	(0.003)
N	27,643		27,643	
R^2	0.246		0.249	

Note: This table reports zero-stage in-migration prediction for the Northern origins using OLS regression. Robust standard errors are clustered by county and reported in parentheses. The dependent variable, net migration rates by race, are from Gardner and Cohen (1992) and Bowles et al. (2016). Stars represent: $* \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$.

Table A2: Exclusion restriction - Pretrend tests.

Panel A. Agriculture

	(1) Employment	(2) Number of farms	(3) Acres in farmland	(4) Number of tractors	(5) Farm output	(6) Farm value per acre
Out-migration rate	0.026**	-0.006	-0.006	0.007	0.003	-0.011
	(0.013)	(0.005)	(0.009)	(0.016)	(0.009)	(0.007)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F	9.56	9.56	12.75	14.07	9.56	9.56
Counties	1,148	1,148	1,147	1,140	1,148	1,148

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Value added	(4) Annual payroll
Out-migration rate	0.002	-0.003	0.012	0.016
	(0.026)	(0.026)	(0.022)	(0.023)
Fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
First-stage F	11.19	13.54	11.82	11.82
Counties	1,110	1,088	1,043	1,043

Panel C. Wholesale

	(1) Wholesale employment	(2) Wholesale establishment	(3) Wholesale sales	(4) Wholesale annual payroll
Out-migration rate	-0.022	-0.022	0.000	-0.012
	(0.016)	(0.021)	(0.015)	(0.018)
Fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
First-stage F	13.93	13.66	13.66	13.06
County	1,114	994	1,001	992

Note: The table reports estimation results using Equation (3) on pre-period outcomes (1920 and 1930). Panels A to C correspond to the baseline results in Tables 1 to 3 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable for the years 1920 and 1930, relative to the omitted years of 1940. Robust standard errors are clustered by county and reported in parentheses, and the first-stage Kleibergen-Paap robust F-statistics are reported. Stars represent: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A3: Quantitative Environment.

Definition	Parameter
Panel (A) Production technology	
Production function	$Y^s_{i,t} = Z^s_{i,t}\left(\rho^s_i (A^{L,s}_{i,t} L^s_{i,t})^{\frac{\sigma^s-1}{\sigma^s}} + (1-\rho^s_i)(A^{K,s}_{i,t} K^s_{i,t})^{\frac{\sigma^s-1}{\sigma^s}}\right)^{\frac{\sigma^s}{\sigma^s-1}}$
Unit cost	$x^s_{i,t} = \left((\rho^s_i)^{\sigma^s}(w^s_{i,t})^{1-\sigma^s} + (1-\rho^s_i)^{\sigma}(r^s_{i,t})^{1-\sigma^s}\right)^{\frac{1}{1-\sigma^s}}$
Panel (B) Workers	
Intratemporal preferences	$U(C^k_{n,t}) = \log\left(C(e^k_{n,t}, P_{n,t})\right)$
Consumption utility	$C(e, P_n) = \frac{1}{\varepsilon}\left(e/P_i\right)^{\varepsilon} - \sum_s \nu^s \ln P^s_n$
Consumption share	$\varphi^s(P_i, e) = \phi^s + \nu^s\left(e/P_i\right)^{-\varepsilon}$
Intertemporal preferences	$\mathbb{V}^s_{i,t} = U(\cdot) + \max_{\{n\}}\left\{\beta\mathbb{E}[\mathbb{V}_{n,t+1}] - \kappa_{ni,t} + \eta\varepsilon_{n,t}\right\}$
Average wage	$\bar{w}_i = \left((w^a_i)^{\zeta^L} + (w^m_i)^{\zeta^L} + (w^s_i)^{\zeta^L}\right)^{1/\zeta^L}$
Sectoral labor allocation	$\tilde{L}^s_i = \Gamma_{\zeta^L} A^{L,s}_i (w^s_i/\bar{w}_i)^{\zeta^L-1} L_i$
Panel (C) Capitalists	
Intertemporal preferences	$v^K_{i,t} = \mathbb{E}_t \sum_{t=0}^{\infty} \beta^t \log(C^K_{i,t})$
Budget constraint	$\bar{r}_{i,t} K_{i,t} = P_{i,t}\left(C^K_{i,t} + K_{i,t} - (1-\delta)K_{i,t}\right)$
Average rental rate	$\bar{r}_i = \left((r^a_i)^{\zeta^K} + (r^m_i)^{\zeta^K} + (r^s_i)^{\zeta^K}\right)^{1/\zeta^K}$
Investment rule	$I_{i,t} = \beta/(1+\beta)\bar{r}_{i,t} K_{i,t}$
Sectoral capital allocation	$\tilde{K}^s_i = \Gamma_{\zeta^K} A^{K,s}_i (R^s_i/\bar{R}_i)^{\zeta^K-1} K_i$
Panel (D) General equilibrium	
Expenditure share	$\mathbb{S}^s_{ni,t} = \frac{\left(x^s_{i,t}\tau^s_{ni,t}/Z^s_{i,t}\right)^{-\theta^s}}{\sum_{l=1}^{N}\left(x^s_{l,t}\tau^s_{nl,t}/Z^s_{l,t}\right)^{-\theta^s}}$
Goods market clearing	$Y^s_{i,t} = \sum_{n=1}^{N} \mathbb{S}^s_{ni,t}\left(\sum_{s=1}^{S}\varphi^s_{ns,t}E^s_{n,t} + \varphi^K_{ns,t}E^K_{n,t}\right)$
Migration flow	$\mathbb{M}_{in,t} = \frac{\exp\left(\beta\mathbb{E}_t(v_{n,t+1} - \kappa_{ni})/\eta\right)}{\sum_{j=1}^{N}\exp\left(\beta\mathbb{E}_t(v_{j,t+1} - \kappa_{nj,t})/\eta\right)}$

Note: The table summarizes the main model elements described in Section 5. The subscript i and n index region. For expenditure and migration share, i denotes the exporter and origin, respectively, and n represents the importer and destination. The superscript $s \in \{a, m.s\}$ denote industry.