Online Appendix

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

A Historical Background Appendix

Figure A1, Panel A, presents the log Black population in the South (red line) and the North (blue dashed line).[1] The steep convergence between 1940 and 1970 gives a hint of the sizes of migration flows between the two regions, where the share of Blacks living 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 hints 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-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.

Panels E and F present proxies for the changes in the capital-to-labor ratio in both agriculture and manufacturing. I measure agricultural mechanization using the number of tractors and combines[2] divided by the number of agricultural workers. For manufacturing, I use the reported manufacturing capital spending.[3] 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 mechanization.[4] Two channels explain the agricultural convergence: relative increases in the num-

[1]The definition of the South mainly follows the Census definition. It contains contain former Confederate states (South Carolina, Mississippi, Florida, Alabama, Georgia, Louisiana, Texas, Virginia, Arkansas, Tennessee, and North Carolina), Delaware, Kentucky, Maryland, Oklahoma, and West Virginia, but excludes the District of Colombia.

[2]Census of Agriculture started to collect the number of tractors in 1925, but combines in 1950. Hence, the 1940 value only includes the number of tractors.

[3]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.

[4]Historical account of the spatial diffusion of tractors accords with the trends in Panel E. By 1920, the absolute majority of American farms depend on horses and mules for power. Then initially, tractors were adopted in the Wheat Belt states (North Dakota, South Dakota, and Kansas) during the 1920s and diffused to the Corn Belt states (Iowa,

ber 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; Abramitzky et al., 2023). The main 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 likely have been 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 examines the relationship between large out-migration (Panel A) and regional convergence (Panels B and C). I especially focus on the potential role of capital deepening (Panels E and F) as a response to labor scarcity from out-migration.

B Simple Framework Appendix

This section presents a simplified model of a two-period, two-country small open economy framework with two industries and two factors of production—labor and capital by interpreting well-established theoretical results in relation to out-migration and subsequent economic development. 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 relative factor abundance and factor intensity. The interpretation introduced here can be applied to another setting where the initial economy contains a sufficient share of the workers in labor-intensive agriculture.

The production function for both agriculture and non-agriculture (denoted "*a*" and "*m*") takes the constant elasticity of substitution (CES) structure using labor L^s and capital K^s in sector s:

$$Y^s = \left(\rho^s (Z_L^s L^s)^{\frac{\sigma^s - 1}{\sigma^s}} + (1-\rho^s)(Z_K^s K^s)^{\frac{\sigma^s - 1}{\sigma^s}}\right)^{\frac{\sigma^s}{\sigma^s - 1}}, \tag{A1}$$

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 produc-

Illinois, and Nebraska) during the 1930s (Gross, 2018). In the South, tractors was rapidly employed in the post-World War 2 periods (Olmstead and Rhode, 2001).

tion function abstracts from Hicks-neutral and the exogenous components of the factor-augmenting technologies, as they are not needed for the core predictions.[5,6] For simplicity, I assume that technology cannot adapt in the first period when the shock occurs but can endogenously adjust in the second period through the directed technical change process (Acemoglu, 2002, 2007).

The CES production function allows for flexible factor usage with a restriction that the elasticity of substitution between labor and capital, σ, is constant. As described in the main text, the value of σ is assumed to be greater than one for agriculture but less than one for non-agriculture by following the CES elasticity estimates in the literature. The factor intensity is defined as the cost share of each factor. With the CES production, the cost share is determined in equilibrium by factor prices and the elasticity and share parameter. Specifically, the labor cost share ξ is given as $\xi = \rho^{\sigma} w^{(1-\sigma)} / (\rho^{\sigma} w^{(1-\sigma)} + (1-\rho)^{\sigma} r^{(1-\sigma)})$, with the wage rate w and the rental rate of capital r.

B.1 Closed-economy force: Factor substitutability

I start with a closed economy implication for the South, which 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 technological development.

As the population flows out, labor becomes more scarce and expensive relative to capital. 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-to-labor ratio, $k = (K^a + K^m)/(L^a + L^m)$. Note that the initial industry of the migrants is irrelevant to the changes in k. I assume that both factors are fully employed and perfectly mobile across sectors within the region. I interpret capital as mobile across sectors but a geographically immobile variable factor, such as local structures for production.

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

[5] In the real economy, 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 homothetic 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. The differences in technical growth between industries, on the other hand, lead to the classic Baumol (1967) effect. 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.

[6] In the main text, 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. The quantitative model allows Hicks-neutral technology and the exogenous components of the factor-augmenting technology.

$\sigma_M < 1$. 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-to-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 Section B.4. □

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-to-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):*

$$\frac{d(Z_K^{s\prime}/Z_L^{s\prime})^{\frac{\sigma^s-1}{\sigma^s}}}{d(K^s/L^s)} > 0, \tag{A2}$$

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-to-labor ratio induces technological change biased toward capital or labor depending on the value of σ. Remark 1 is equivalent to imposing additional assumptions on the technology environment as in Acemoglu (2007).[7] The CES function

[7] Acemoglu (2007) lays out a menu of different assumptions that are unrelated to, but needed in addition 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 through supplying intermediate goods.

meets the required assumption for the production side. The technical changes can be thought of as generated by learning-by-doing or directed R&D effort. Examples of learning by doing are Southern farmers becoming more proficient at operating farm machinery and local mechanics gaining expertise in repairing it. Directed R&D efforts include enhancing farm machines to better suit Southern crops.

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 with $\sigma > 1$. Hence, it becomes more profitable for the economy to focus on using capital more efficiently. 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 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 increase in the capital-to-labor ratio from the out-migration dynamically incentivizes the adoption of capital-augmenting technology, further raising the capital-to-labor ratio, k, in the second period. It again leads to the reallocation of labor from agriculture to non-agriculture.

B.2 Open-economy force: 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. The key idea is that relative factor abundance can determine the regional comparative advantage, and hence, the pattern of trade and production.

I introduce additional assumptions regarding factor abundance and factor intensity: the South is labor-abundant, characterized by a lower capital-to-labor ratio compared to the North (Figure A1, Panel E), and the Southern agricultural production is intensive in labor, with labor accounting for a larger share of production costs (Bateman and Weiss, 1981; Wright, 1986). In other words, labor is relatively cheaper in the South, giving it a comparative advantage in agriculture. I also assume that the relative factor intensity is not reversed from the factor reallocation. The resulting Prediction 3 states the quasi-Rybczynski effect (Romalis, 2004):[8]

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 relative contraction of the agricultural sector and a relative increase in non-agriculture production. Non-agriculture shares of labor and capital increase.*

The related proof and discussion are in Section B.4. The results follow from the changed pattern of comparative advantage. As the South lost labor, its comparative advantage in the labor-intensive sector decreased, incentivizing the economy to reallocate resources toward the non-agricultural sector, which had relatively gained a comparative advantage.

The open-economy forces predict a decrease in labor share in agriculture, as in the closed-economy forces. However, the open-economy forces relatively expand non-agriculture while shrinking agriculture, which has different implications for capital allocation. I summarize the common and competing predictions from these different perspectives in the next subsection. In the second period, economy-wide capital accumulation would lead to a further expansion of non-agriculture with a relative decrease in agriculture production.

B.3 Discussion

Summary of the predictions. Closed- and open-economy results rely on related but distinct assumptions. The closed-economy perspective focuses on the differences in labor-capital substitutability and factor reallocation across sectors, whereas the open-economy view relies on the differences in factor intensity between the industries and the changes in comparative advantages.

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 non-agricultural 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 against the relative expansion of the non-agricultural sector. 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.

[8]When a factor endowment increases, the Rybczynski theorem (Rybczynski, 1955) predicts a more than a proportional expansion of a sector which uses that factor more intensively through a magnification effect (Jones, 1965). However, his sharp prediction relies on factor price equalization, which may not hold in practice. Instead, I state a weaker version that requires a weaker set of assumptions.

However, there are notable differences in terms of agricultural capital and production. In the first period, the increase in the economy-wide capital-to-labor ratio induces capital adoption in agriculture, while trade effects lessen it. Which effects dominate depends on the relative strength of the closed-economy force (factor substitutability) and the quasi-Rybczynski effect (factor intensity). However, note that even if the agricultural share of capital in the economy decreases, the sectoral capital-to-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 Heckscher–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, with capital-biased technical change, the agricultural capital stock would progressively increase as time passes.

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. On the other hand, the changes in agricultural capital and production are left as empirical questions. I then use quantitative analysis to assess the potential contribution 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, 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 in the Second period 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.

B.4 Proofs

Predictions 1 and 2.

The proofs for Predictions 1 and 2 follow Proposition 1 of Alvarez-Cuadrado et al. (2017). This subsection outlines the key ideas of the proof. Let consumption share on non-agriculture ("m") be $0 < \gamma < 1$ in the Cobb-Douglas consumption allocation. Then, with total output Y and price index P, the economic allocation is given as:

$$Y^a = (1-\gamma)\frac{P}{p^a}Y \quad \text{and} \quad Y^m = \gamma\frac{P}{p^m}Y, \tag{A3}$$

where the subscript indexes sector. The price index P is normalized to one:

$$P = p^a \frac{Y^a}{Y} + p^m \frac{Y^m}{Y} = (1-\gamma)p^a + \gamma p^m \equiv 1. \tag{A4}$$

At any point in time, free mobility of capital and labor implies the equalization of the marginal value products across sectors within the region:

$$p^a \rho \left(\frac{Y^a}{K^a}\right)^{1/\sigma^a} = p^m \rho \left(\frac{Y^m}{K^m}\right)^{1/\sigma^m} = R, \tag{A5}$$

$$p^a (1-\rho) \left(\frac{Y^a}{K^a}\right)^{1/\sigma^a} A^{\frac{\sigma^a - 1}{\sigma^a}} = p^m (1-\rho) \left(\frac{Y^m}{L^m}\right)^{1/\sigma^m} A^{\frac{\sigma^m - 1}{\sigma^m}} = w. \tag{A6}$$

In order to obtain analytical results, I restrict non-agriculture to Cobb-Douglas by setting $\sigma_A = \sigma > 1$ and $\sigma_M = 1$. By combining Equations (A3), (A4), (A5), and (A6), the share of labor in agriculture, λ, and capital, κ, are each governed by the equilibrium relationships:

$$\lambda - (1-\lambda)^{1/\sigma} A^{\frac{1-\sigma}{\sigma}} K^{\frac{\sigma-1}{\sigma}} \frac{\kappa}{(1-\kappa)^{1/\sigma}} = 0, \tag{A7}$$

$$\kappa - (1-\kappa)^{1-\sigma} \frac{\gamma}{1-\gamma} \left(\frac{Y^a}{K}\right)^{\frac{\sigma-1}{\sigma}} = 0, \tag{A8}$$

where

$$\left(\frac{Y^a}{K}\right)^{\frac{\sigma-1}{\sigma}} = (1-\alpha)\left(\frac{(1-\lambda)A}{K}\right)^{\frac{\sigma-1}{\sigma}} + \alpha(1-\kappa)^{\frac{\sigma-1}{\sigma}}.$$

Combining Equations (A7) and (A8) yields the following relationship between the shares:

$$\lambda = \lambda(\kappa) = \frac{\gamma(1-\alpha)\kappa}{\kappa - \alpha\gamma} \text{ with } \frac{d\lambda}{d\kappa} = -\left(\frac{\alpha}{1-\alpha}\right)\left(\frac{\lambda(\kappa)}{\kappa}\right) < 0, \tag{A9}$$

which states that the share of labor and capital allocated to agriculture exhibits the opposite sign from a common shock.

Finally, manipulating Equations (A7) and (A9) and taking logarithms gives:

$$\frac{1-\sigma}{\sigma} \ln K = \frac{1}{\sigma} \ln\Big(1-\lambda(\kappa)\Big) - \ln\Big(\lambda(\kappa)\Big) + \frac{1-\sigma}{\sigma} \ln A + \ln\kappa - \frac{1}{\sigma}\ln(1-\kappa). \tag{A10}$$

The result for the capital share is obtained by totally differentiating Equation (A10):

$$\frac{\partial \kappa}{\partial k} = \frac{(1-\sigma)}{\sigma G(\kappa) k} > 0,$$

where $G(\kappa) \equiv \left[\frac{1}{\sigma(1-\lambda(\kappa))} + \frac{1}{\lambda(\kappa)}\right]\left(\frac{\lambda(\kappa)}{\kappa}\right)\left(\frac{\alpha}{1-\alpha}\right) + \left[\frac{1}{\kappa} + \frac{1}{\sigma(1-\kappa)}\right]$. Hence, the capital share on agriculture rises as the regional capital-to-labor ratio, k, increases. The result for the labor share follows from Equation (A9).

Prediction 3.

The classical proof using factor price equalization is first explained below. Consider unit costs and production constraints using input coefficients:

$$\begin{aligned} a_L^a w + a_K^a r &= p^a, \\ a_L^m w + a_K^m r &= p^m, \end{aligned} \tag{A11}$$

and

$$\begin{aligned} a_L^a A + a_K^a M &= L, \\ a_L^m A + a_K^m M &= K, \end{aligned} \tag{A12}$$

where the first two conditions are derived from the zero profit condition, and the last two are derived from the full employment assumption. The coefficient a_k^s is a derivative of unit cost functions with respect to factor prices and measures factor content of k on industry s at the initial equilibrium. The agriculture is more capital intensive if

$$\frac{a_K^a}{a_L^a} > \frac{a_K^m}{a_L^M}. \tag{A13}$$

The following derivation relies on the assumption that factor intensity is not reversed.

By taking derivatives of Equation (A12) with respect to labor,

$$\begin{aligned} a_L^a \frac{\partial A}{\partial L} + a_K^a \frac{\partial M}{\partial L} &= 1, \\ a_L^m \frac{\partial A}{\partial L} + a_K^m \frac{\partial M}{\partial L} &= 0, \end{aligned} \tag{A14}$$

and applying the Cramer's rule yields:

$$\begin{aligned} \frac{\partial A}{\partial L} &= \frac{a_K^m}{a_L^a a_K^m - a_L^m a_K^a}, \\ \frac{\partial M}{\partial L} &= \frac{a_K^a}{a_L^a a_K^m - a_L^m a_K^a}. \end{aligned} \tag{A15}$$

Under the assumption that agriculture is labor intensive and the South is labor abundant, $\frac{\partial A}{\partial L} > 0$ and $\frac{\partial M}{\partial L} < 0$ should hold from Equation (A15). In other words, out-migration contracts agriculture and expands non-agriculture production. Note that Equation (A14) uses factor price equalization result where factor prices only depend on output prices, which is assumed to be constant. As long as the number of goods and the number of factors are equal, the same logic can be applied to an arbitrary number of goods and factors where an increase in endowment of one factor decreases the output of one good and increases the output of the others (Jones and Scheinkman, 1977).

Nonetheless, the factor price equalization is unlikely to hold in the real world. Romalis (2004), instead, lays out a model of two factors of production, multiple countries with continuous industries that differ in factor intensity. He derives the quasi-Rybczynski Theorem without the factor price equalization but with the introduction of trade costs and monopolistic competition, allowing regional differences in production costs. The quantitative model is closer to his formulation by incorporating the CES cost function into the Eaton and Kortum (2002) setting.

C Quantitative Framework Appendix

C.1 Initial equilibrium

Before performing the counterfactual analysis, I first construct bilateral trade flows between all regional pairs using the market access term approach (Donaldson and Hornbeck, 2016; Allen and Donaldson, 2022). The bilateral trade flows between importer n and exporter i in industry k can be expressed as a function of trade costs $\tau^s_{ni,t}$, price index $P^s_{n,t}$, expenditure $X^s_{n,t}$, and total income $Y_{n,t}$, with productivity $A^s_{i,t}$ and distance elasticity θ^k:

$$\begin{aligned} X^s_{ni,t} &= \left(\tau^s_{ni,t}\frac{w^s_{i,t}}{A^s_{i,t}}\right)^{-\theta^k}(P^s_{n,t})^{\theta^k} \times \varphi^k Y_{n,t} \\ &= T^s_{ni,t} \times \left(\frac{Y^s_{i,t}/Y_W}{(\Pi^s_{i,t})^{-\theta^k}}\right) \times \left(\frac{X^k_{n,t}/Y_W}{(P^s_{n,t})^{-\theta^k}}\right). \end{aligned} \tag{A16}$$

World income Y_W normalizes the output and expenditure. The second line follows from introducing the outward market access term $\Pi^s_{i,t} \equiv \left(\frac{w^s_{i,t}}{A^s_{i,t}}\right)^{-1}\left(\frac{Y^s_{i,t}}{Y_W}\right)$ and the effective trade costs $T^s_{ni,t} = (\tau^s_{ni,t})^{-\theta^k}$.

The outward market access term $\Pi_{i,t}$ captures how close exporter i is to the potential importers. The price index $P_{n,t}$ represents the inward trade market access and captures how close each importer n is to exporters. Using the two market access terms, the following proposition defines the initial equilibrium for trade.

C.2 Dynamic Exact Hat Algebra

I adopt the dynamic hat algebra approach of Caliendo et al. (2019) for the counterfactual analysis. Dynamic hat algebra calculates how allocations and prices change in a counterfactual economy relative to a baseline economy across space and time. It annihilates the need to recover time-invariant fundamentals and focuses on quantifying the changes in allocations and prices, given a new sequence of fundamentals.

First, I clarify fundamentals and prices. Geographic fundamentals $\{\bar{Z}^s_{i,t}, \bar{A}^{w,s}_{i,t}, \bar{A}^{k,s}_{i,t}, B_{i,t}, \tau^s_{ni,t}, \kappa_{ni,t}\}$ consists of fundamental productivity of firm, $\bar{Z}^s_{i,t}$; fundamental efficiency of labor and capital, $\bar{A}^{w,s}_{i,t}$ and $\bar{A}^{k,s}_{i,t}$; location amenities, $B_{n,t}$; and iceberg-type trade costs and migration costs, $\tau^s_{ni,t}$ and $\kappa_{ni,t}$. I use the terms Θ_t and $\bar{\Theta}$ to denote time-varying and time-constant fundamentals of the economy.

Prices and allocation $\{w^s_{i,t}, r^s_{i,t}, L_{i,t}, K_{i,t}, \mathbb{V}^l_{i,t}\}$ consists of location-sector specific factor prices: wage $w^s_{i,t}$ and rental rate of capital $r^s_{i,t}$; the measure of workers, $L_{i,t}$ and the stock of capital, $K_{i,t}$; and workers' value function, $\mathbb{V}^l_{i,t}$.

A dot notation expresses a variable in terms of changes over time: $\dot{Y}_{t+1} \equiv Y_{t+1}/Y_t$. A prime denotes the value at the counterfactual economy, and the change in the counterfactual economy can also be expressed in terms of time changes: $\dot{Y}'_{t+1} \equiv Y'_{t+1}/Y'_t$. Lastly, a hat variable, $\hat{Y}_t \equiv \dot{Y}'_t/\dot{Y}_t$, stands for the counterfactual time change relative to the baseline time changes for any variable Y.

In the baseline case, the saving rate of capitalists, $\varsigma_{n,t}$, is fixed at β using the log consumption utility, with the value of the intertemporal elasticity of substitution equaling to one (ψ=1). If $\psi > 1$, capitalists flexibly substitute consumption in different time periods, suggesting that the rise in the rental rate of capital induces a higher saving rate. In other words, the substitution effect dominates. If $0 < \psi < 1$, the income effect dominates, and the saving rate decreases with the rise in the rental rate. In a more general setting, the sectoral consumption share of capitalists can be allowed to vary and be determined by the shooting algorithm.

For production, the variable $\xi^s_{i,t} = (w^s_{i,t}L^s_{i,t})/(w^s_{i,t}L^s_{i,t} + r^s_{i,t}K^s_{i,t})$ represents the cost share of labor, and $\varphi^k_{is,t}$ the consumption shares of individuals in labor market *i*-*s* on sector *k* at time *t*. The factors of production are intra-temporarily reallocated across sectors until effective payments are equalized across sectors.

Proposition 1 (Temporary Equilibrium). Given the allocation of the temporary equilibrium at t, $\{L_t, K_t, X_t, \mathbb{S}_t, \varphi_t\}$, the solution to the temporary equilibrium at $t+1$ for a given change in $\dot{L}_{t+1}$, $\dot{K}_{t+1}$, and $\dot{\Theta}_{t+1}$ does not require information on the level of fundamentals at t, Θ_t or $\bar{\Theta}$. In particular, the set of solutions is obtained as the solution to the following system of non-linear equations:

$$\dot{x}^s_{i,t+1} = \left((\xi^s_{i,t})(\dot{w}^s_{i,t+1})^{1-\sigma^s} + (1-\xi^s_{i,t})(\dot{r}^s_{i,t+1})^{1-\sigma^s} \right)^{\frac{1}{1-\sigma^s}}, \tag{T1}$$

$$\dot{P}^s_{n,t+1} = \left[\sum_{i=1}^{N} \mathbb{S}^s_{ni,t} \left(\dot{\tau}^s_{ni,t+1} \dot{x}^s_{i,t+1} \right)^{-\theta^s} \right]^{-1/\theta^s}, \tag{T2}$$

$$\mathbb{S}^s_{ni,t+1} = \mathbb{S}^s_{ni,t}\left(\frac{\dot{\tau}^s_{ni,t+1}\dot{x}^s_{i,t+1}}{\dot{P}^s_{n,t+1}}\right)^{-\theta^s}, \tag{T3}$$

$$s^{w,s}_{i,t+1} = s^{w,s}_{i,t}\left(\frac{\dot{A}^{w,s}_{i,t+1}\dot{w}^s_{i,t+1}}{\dot{\bar{w}}_{i,t+1}}\right) \quad \text{and} \quad s^{k,s}_{i,t+1} = s^{k,s}_{i,t}\left(\frac{\dot{A}^{k,s}_{i,t+1}\dot{R}^s_{i,t+1}}{\dot{\bar{R}}_{i,t+1}}\right), \tag{T4}$$

$$\dot{L}^s_{i,t+1} = \dot{A}^{w,s}_{i,t+1}\dot{s}^{w,s}_{i,t+1}\dot{L}_{i,t+1} \quad \text{and} \quad \dot{K}^s_{i,t+1} = \dot{A}^{k,s}_{i,t+1}\dot{s}^{k,s}_{i,t+1}\dot{K}_{i,t+1}, \tag{T5}$$

$$E^s_{n,t+1} = \dot{w}^s_{n,t+1}\dot{L}^s_{n,t+1}w^s_{n,t}L^s_{n,t} \quad \text{and} \quad E^k_{n,t+1} = (1-\beta)\sum_{s\in\{a,m,l\}}\dot{R}^s_{n,t+1}\dot{K}^s_{n,t+1}R^s_{n,t}K^s_{n,t} \tag{T6}$$

$$\varphi^k_{is,t+1} = \phi^k + \left(\varphi^k_{is,t} - \phi^k\right)\left(\frac{\dot{w}^s_{i,t+1}}{\dot{P}_{i,t+1}}\right)^{-\varepsilon}, \tag{T7}$$

$$\dot{Y}^s_{i,t+1}Y^s_{i,t} = \sum_{n=1}^{N}\mathbb{S}^s_{ni,t+1}\left(\sum_{s=1}^{S}\varphi^s_{ns,t+1}E^s_{n,t+1} + \varphi^k_{ns,t+1}E^k_{n,t+1}\right), \tag{T8}$$

$$\dot{w}^s_{n,t+1}\dot{L}^s_{n,t+1}w^s_{n,t}L^s_{n,t} = \xi^s_{i,t+1}\dot{Y}^s_{i,t+1}Y^s_{i,t} \quad \text{and} \quad \dot{r}^s_{n,t+1}\dot{K}^s_{n,t+1}r^s_{n,t}K^s_{n,t} = (1-\xi^s_{i,t+1})\dot{Y}^s_{i,t+1}Y^s_{i,t}, \tag{T9}$$

$$K_{i,t+1} = \beta\bar{R}_{i,t}K_{i,t}. \tag{T10}$$

The next proposition shows how to calculate a dynamic sequence of the economy. I study how migration flows, $\mathbb{M}$, vary by the new sequence of temporary equilibrium. I define $\mathbb{V}^k_{n,t} \equiv \exp(V^k_{n,t})$.

Proposition 2 (Sequential Equilibrium). Given an initial allocation of the economy, $(L_0, K_0, X_0, \mathbb{S}_0, \varphi_0, \mu_{-1})$, and an anticipated convergent sequence of time changes in fundamentals, $\{\dot{\Theta}_t\}_{t=\tau+1}^{\infty}$ with $\lim_{t\to\infty}\dot{\Theta}_t = 1$, the solution to the sequential competitive equilibrium in time differences does not require information in the level of fundamentals, $\{\Theta_t\}_{t=0}^{\infty}$, where all agents are assumed to learn the full sequence shock between τ and $\tau+1$. In particular, the changes in migration shares and capital stocks are obtained as the solution to the following system of non-linear equations:

$$\mathbb{M}_{in,t+1} = \frac{\mathbb{M}_{in,t}\left(\dot{\mathbb{V}}_{i,t+2}\right)^{\beta/\eta}}{\sum_{l=1}^{N}\mathbb{M}_{il,t}\left(\dot{\mathbb{V}}_{l,t+2}\right)^{\beta/\eta}} \tag{S1}$$

$$\dot{\mathbb{V}}^k_{n,t+1} = \dot{B}_{n,t+1}\dot{C}_n\left(\dot{L}_{t+1}, \dot{\Theta}_{t+1}\right)\left(\sum_{i=1}^{N}\sum_{s=1}^{k}\mathbb{M}_{ni,t}(\dot{\mathbb{V}}_{i,t+2})^{\beta/\eta}\right)^{\eta}, \tag{S2}$$

$$L_{n,t+1} = \sum_{i=1}^{N}\mathbb{M}_{in,t}L_{i,t}, \tag{S3}$$

where the changes consumption indirect utility, $C_n(\dot{L}_{t+1}, \dot{\Theta}_{t+1})$, is constructed from the solution of the temporary equilibrium given $\{\dot{L}_t, \dot{\Theta}_t\}_{t=1}^{\infty}$:

$$C_{n,t+1} = \frac{1}{\varepsilon}\left(\frac{\dot{\bar{w}}_{n,t+1}}{\dot{P}_{n,t+1}}\frac{\bar{w}_{n,t}}{P_{n,t}}\right)^{\varepsilon} - \sum_{s=1}^{k} \nu^s \ln \dot{P}_{n,t+1}^s P_{n,t}^s.$$

I first recover wages at $t = 1$ consistent with the initial period economic allocation. The first term in RHS is the real wage component and the second term is the non-homothetic price adjustments.

Stationary Equilibrium. A stationary equilibrium is a sequential competitive equilibrium such that $\{\mathbf{L}_t, \mathbb{M}_t, \mathbb{V}_t, \mathbf{w}_t, \mathbf{K}_t, \mathbf{r}_t\}_{t=0}^{\infty}$ is constant for every t. A stationary equilibrium in this economy is a situation in which no aggregate variables change over time. It requires fundamentals to be constant for all t. In a stationary equilibrium, individuals continually move from one market to another, but inflows and outflows balance.

D Additional Empirical Results

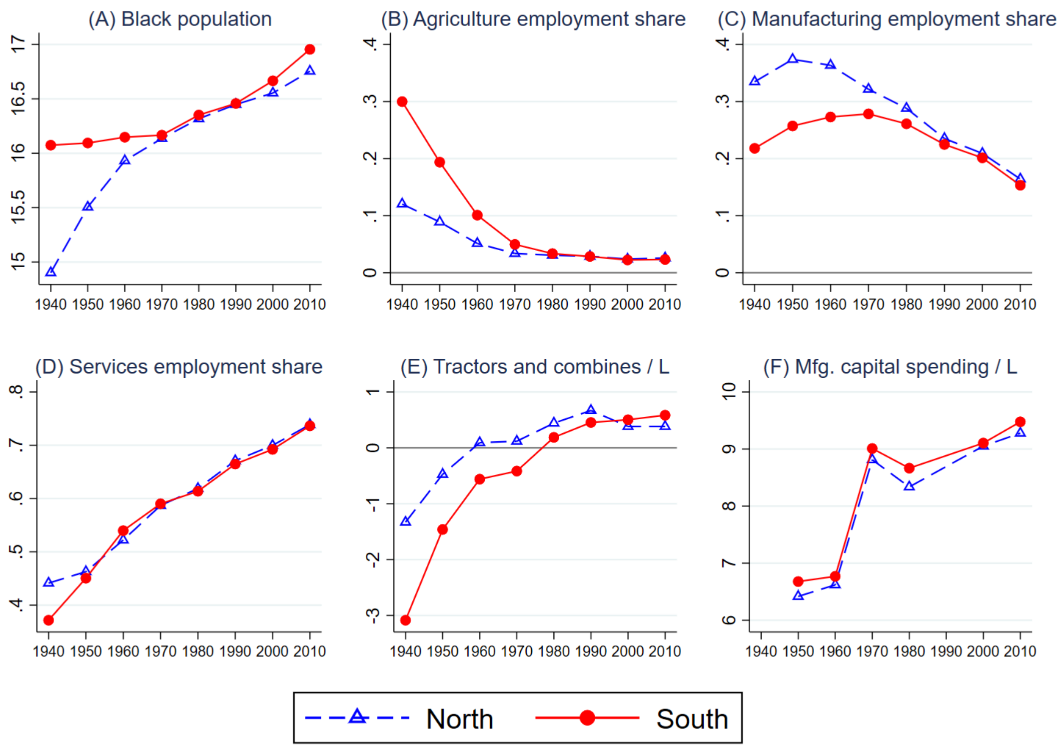
D.1 Robustness check

Tables A8, A9, A10, A11, and A12 replicate the main results (Tables 2, 3, and 4) using alternative empirical strategies with (1) the alternative share based on 1935-1940 migration matrix from the 1940 Census (Table A8), (2) alternative approaches for in-migration prediction (by using actual in-migration rates in Table A9 or by using the random forest prediction method in Table A10), (3) subsample consisting former confederate States, and (4) additional time-interacted controls that can at least partially take into account potentially correlated migration exposure across Southern counties (Table A12). I do so by constructing a weighted average of the Northern migration exposures as an additional control. I use pre-period migration linkages within the South during the 1910-1940 period as weights, as in Borusyak et al. (2023).

I mainly discuss the cases when the results differ from the baseline results. Table A8 reports the estimates using the alternative migration share between 1935 and 1940. As demonstrated using overidentification tests documented in Table A4, the baseline migration share and the alternative 1935-1940 share tend to produce similar results. However, there are notable differences for manufacturing value-added, with a slightly negative estimate, and for retail variables, which are less precisely estimated. The difference between the baseline method and the 1935-1940 migration matrix may be due to more sparse migration cells in the 1935-1940 share, as pointed out in the main text. The results are similar in the other robustness exercises.

E Additional Tables and Figures.

Figure A1: 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 values in Panels A, E, and F are logged. The definition of the region is described in Footnote 1, where "the North" encompasses both Northern and Westerns states. Panel A, the number of the Black population, is calculated using Haines et al. (2010). Panels B to D, the industry employment 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. Panel E, the number of tractors and combines per agriculture employment, is calculated from Haines et al. (2018). Panel F, the manufacturing capital per manufacturing employment, is calculated from Haines et al. (2010).

Table A1: Zero-stage in-migration prediction.

	(1)	(2)	(3)	(4)	(5)	(6)
	1940-1950		1950-1960		1960-1970	
	Black	White	Black	White	Black	White
Latitude	-5.550	-0.154	-7.302*	-0.377	-3.029*	-1.355***
	(3.605)	(0.227)	(3.937)	(0.487)	(1.828)	(0.226)
Longitude	-4.769***	0.035	-2.099	-0.412*	-1.936**	-0.248**
	(1.702)	(0.107)	(1.907)	(0.236)	(0.914)	(0.113)
Log population	67.285	3.636	-117.904	21.049	12.575	13.916**
	(73.652)	(4.636)	(106.951)	(13.216)	(49.977)	(6.172)
Log black population	-15.663***	0.101	-63.246***	1.234**	-4.314**	0.490*
	(3.423)	(0.215)	(4.044)	(0.500)	(2.076)	(0.256)
Log white population	-71.899	-3.410	162.889	-29.611**	-11.011	-13.347**
	(74.179)	(4.669)	(108.079)	(13.356)	(50.658)	(6.256)
Urbanization	-0.987***	0.057***	-0.813***	0.251***	-0.307**	-0.029*
	(0.279)	(0.018)	(0.301)	(0.037)	(0.137)	(0.017)
Median income	-3.979	0.016	26.244***	15.626***	14.718***	9.442***
	(4.364)	(0.275)	(8.928)	(1.103)	(4.577)	(0.565)
Log housing units	12.854**	-0.294	14.702**	0.819	2.392	-1.024***
	(5.283)	(0.333)	(5.736)	(0.709)	(2.646)	(0.327)
Median rent	-3.305	-0.064	-7.525	5.714***	-11.123***	5.218***
	(5.549)	(0.349)	(7.098)	(0.877)	(3.430)	(0.424)
1940 Share naturalized	3.413	-0.559***	6.187*	-0.407	5.571***	0.467**
	(3.138)	(0.198)	(3.422)	(0.423)	(1.573)	(0.194)
1940 Share foreigner	-1.768	-0.365	3.385	0.259	-1.126	-0.185
	(3.633)	(0.229)	(3.977)	(0.491)	(1.840)	(0.227)
1940 Employment rate	3.579**	0.874***	1.025	-0.163	1.328	0.469***
	(1.622)	(0.102)	(1.802)	(0.223)	(0.840)	(0.104)
1940 Occupational score	0.036	0.018**	0.227*	0.015	-0.061	-0.020***
	(0.117)	(0.007)	(0.130)	(0.016)	(0.060)	(0.007)
Republican vote share (1944)	-0.055	-0.072***	0.332	0.174***	-0.246	-0.041
	(0.438)	(0.028)	(0.484)	(0.060)	(0.224)	(0.028)
Republican vote share (1948)	-1.028**	0.032	-0.827*	0.012	-0.252	0.087***
	(0.449)	(0.028)	(0.492)	(0.061)	(0.227)	(0.028)
Republican vote share (1952)	0.478	-0.155***	0.052	0.004	0.600*	-0.007
	(0.632)	(0.040)	(0.695)	(0.086)	(0.320)	(0.039)
Republican vote share (1956)	0.101	0.234***	-0.664	-0.135*	-0.233	0.083**
	(0.577)	(0.036)	(0.633)	(0.078)	(0.291)	(0.036)
Republican vote share (1960)	-0.312	0.002	0.494	0.133**	0.090	0.044
	(0.480)	(0.030)	(0.525)	(0.065)	(0.242)	(0.030)
Republican vote share (1964)	0.169	-0.011	-0.533	-0.153***	-0.035	-0.133***
	(0.412)	(0.026)	(0.453)	(0.056)	(0.211)	(0.026)
Republican vote share (1968)	1.429**	0.028	1.636**	0.100	0.631**	0.129***
	(0.600)	(0.038)	(0.660)	(0.082)	(0.305)	(0.038)
Republican vote share (1972)	-0.090	-0.017**	0.071	0.005	-0.048	-0.016*
	(0.134)	(0.008)	(0.147)	(0.018)	(0.068)	(0.008)
State FE	Yes	Yes	Yes	Yes	Yes	Yes
Observation	3,102	3,102	3,092	3,092	3,080	3,080
R^2	0.152	0.234	0.333	0.261	0.114	0.525

Note: This table reports zero-stage in-migration predictions using OLS regression at the county level. The dependent variable is the net migration rate by race and decade, and the explanatory variables are as listed. The sample includes all counties in the U.S., and all specifications include state fixed effects. Robust standard errors are clustered by county and reported in parentheses. Stars represent: $** \ p < 0.05$, $** \ p < 0.01$, $*** \ p < 0.001$.

Table A2: SSIV second-stage estimates for main outcomes.

Panel A. Agriculture

	(1) 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
Out-migration rate (2SLS, 1%)	-0.011*	-0.006*	-0.023***	0.042***	0.014	-0.004	0.005
Clustered s.e. (County)	(0.007)	(0.004)	(0.006)	(0.010)	(0.020)	(0.004)	(0.003)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F	26.59	24.62	24.36	25.40	18.48	14.64	26.76
Counties	1,096	1,096	1,096	1,090	1,058	1,090	1,090

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Out-migration rate (2SLS, 1%)	0.022**	0.026***	0.040	0.024	0.043**
Clustered s.e. (County)	(0.011)	(0.007)	(0.028)	(0.021)	(0.017)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
First-stage F	23.03	24.33	15.54	21.07	22.23
Counties	1,096	1,096	1,065	1,096	1,096

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Out-migration rate (2SLS, 1%)	0.047***	0.040***	0.026*	0.035**	0.021***	0.010***	0.015***	0.022***
Clustered s.e. (County)	(0.011)	(0.010)	(0.014)	(0.015)	(0.006)	(0.003)	(0.005)	(0.007)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F	27.46	24.93	27.57	22.70	25.07	25.06	25.07	25.04
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096

Note: The table reports the second-stage estimation results, with county-year as the unit of observation. Panels A to C correspond to the baseline results in Panel D of Tables 2 to 4 with the full set of fixed effects and control variables. Robust standard errors are clustered by county and reported in parentheses. Stars represent: $* \; p < 0.10$, $** \; p < 0.05$, $*** \; p < 0.01$.

Table A3: 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
Migration exposure (SSIV, 1std)	0.051***	0.009	0.014	0.029	0.038***	-0.004
Clustered s.e. (county)	(0.013)	(0.011)	(0.018)	(0.021)	(0.012)	(0.013)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.422	0.411	0.201	0.136	0.478	0.298
First-stage F	12.97	12.97	15.99	16.71	12.97	12.97
Counties	1,096	1,096	1,095	1,090	1,096	1,096

Panel B. Manufacturing

	(1) Employment	(2) Number of establishments	(3) Value added	(4) Annual payroll	(5) Intermediate goods	(6) Revenue
Migration exposure (SSIV, 1std)	-0.058	0.022	-0.041	-0.066	-0.007	-0.021
Clustered s.e. (county)	(0.041)	(0.019)	(0.033)	(0.042)	(0.042)	(0.036)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.172	0.204	0.207	0.197	0.195	0.216
First-stage F	14.36	13.05	12.69	12.69	12.69	12.69
Counties	1,035	1,058	994	994	994	994

Panel C. Wholesale

	(1) Wholesale employment	(2) Wholesale establishment	(3) Wholesale sales	(4) Wholesale annual payroll
Migration exposure (SSIV, 1std)	-0.035	-0.001	-0.036	0.003
Clustered s.e. (county)	(0.038)	(0.030)	(0.033)	(0.036)
Fixed effects	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes
R-squared	0.133	0.125	0.146	0.183
First-stage F	14.52	14.83	14.77	13.13
Counties	947	1,060	953	943

Note: The table reports estimation results using Equation (3) on pre-period outcomes (1920 and 1930 for agriculture and manufacturing, 1930 for wholesale variables), with county-year as the unit of observation. Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1920 and 1930, relative to the omitted year of 1940. Robust standard errors are clustered by county and shown in parentheses, and the first-stage Kleibergen-Paap robust F-statistics are reported. Counties with zero values in the dependent variable before log transformation are dropped from the sample. Stars represent: $* \ p < 0.10$, $** \ p < 0.05$, $*** \ p < 0.01$.

Table A4: Overidentification tests using alternative shares (second-stage estimates).

Panel A. Agriculture

	(1) 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
Out-migration rate (2SLS, 1%)	-0.012*	-0.010**	-0.024***	0.039***	0.019	-0.016**	0.005
	(0.006)	(0.004)	(0.006)	(0.011)	(0.020)	(0.008)	(0.003)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F	13.69	13.16	13.16	13.22	10.93	13.16	13.16
Counties	1,096	1,096	1,096	1,096	1,060	1,096	1,096
Sargan-Hansen J	0.43	0.15	1.74	0.17	1.78	0.03	0.00
Sargan-Hansen p-value	0.51	0.70	0.19	0.68	0.18	0.87	1.00

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Out-migration rate (2SLS, 1%)	0.020*	0.026***	0.044	0.019	0.039**
	(0.011)	(0.007)	(0.027)	(0.019)	(0.016)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
First-stage F	12.38	12.82	8.00	11.04	11.67
Counties	1,096	1,096	1,065	1,096	1,096
Sargan-Hansen J	1.79	0.03	1.90	4.59	4.20
Sargan-Hansen p-value	0.18	0.87	0.17	0.03	0.04

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Out-migration rate (2SLS, 1%)	0.052***	0.042***	0.029**	0.037***	0.019***	0.009***	0.013***	0.020***
	(0.012)	(0.010)	(0.014)	(0.014)	(0.005)	(0.003)	(0.004)	(0.006)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
First-stage F	14.40	13.25	14.39	12.46	13.16	13.16	13.16	13.15
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096
Sargan-Hansen J	5.89	0.74	1.32	0.60	12.05	2.35	13.59	15.43
Sargan-Hansen p-value	0.02	0.39	0.25	0.44	0.00	0.13	0.00	0.00

Note: The table reports two-stage estimates with two migration exposures (two excluded instruments), separately constructed from (1) 1910-1940 migration share from matched Census (Ruggles et al., 2024a; Buckles et al., 2023) and (2) 1935-1940 migration share from 1940 Census record. The unit of observation is county-year. Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by county-level net out-migration rates, relative to the omitted base years of 1940 and 1950. Robust standard errors are clustered by county and shown in parentheses, and the first-stage Kleibergen-Paap robust F-statistics are reported. Heteroskedasticity-robust Sargan-Hansen J statistics under the null of constant effects and related p-values are added. Stars represent: $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table A5: SSIV estimation results from the public sector (reduced-form, falsification test).

	(1) Local government employment	(2) Local government annual payroll	(3) Federal government employment
Migration exposure (SSIV, 1std)	-0.015	-0.018*	0.019
Clustered s.e. (county)	(0.009)	(0.010)	(0.021)
Conley s.e. (250km)	[0.009]	[0.008]	[0.021]
R-squared	0.404	0.187	0.219
First-stage F	17.11	23.81	16.85
County	1,096	1,096	1,096

Note: The table reports SSIV reduced-form estimates using Equation (3) 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. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2000 by standardized Northern migration exposure, relative to the omitted year of 1960. 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: $*$ $p < 0.10$, $**$ $p < 0.05$, $***$ $p < 0.01$.

Table A6: Placebo test - Random share.

Panel A. Agriculture

	(1) 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
Migration exposure	-0.000	-0.007	-0.011*	0.000	0.039	-0.015	0.005
	(0.015)	(0.006)	(0.006)	(0.020)	(0.035)	(0.013)	(0.007)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.457	0.625	0.339	0.438	0.087	0.968	0.640
First-stage F	0.02	0.02	0.05	0.00	0.04	2.49	0.02
Counties	1,096	1,096	1,096	1,090	1,058	1,090	1,090

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Migration exposure	-0.015	-0.006	-0.108	-0.070	-0.023
	(0.024)	(0.010)	(0.111)	(0.049)	(0.041)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.156	0.411	0.160	0.144	0.173
First-stage F	0.00	0.06	0.28	0.14	0.15
Counties	1,096	1,096	1,065	1,096	1,096

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Migration exposure	0.013	0.013	0.048	0.031	0.009	-0.002	0.013*	0.013
	(0.016)	(0.011)	(0.044)	(0.037)	(0.008)	(0.006)	(0.007)	(0.010)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.269	0.435	0.141	0.130	0.659	0.727	0.679	0.623
First-stage F	0.17	0.12	0.16	0.04	0.04	0.04	0.04	0.05
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096

Note: The table reports reduced-form estimates from Equation (3). The placebo migration exposure is created by using randomly shuffled North-South migration linkages as shares. Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by standardized Northern migration exposure, relative to the omitted base years of 1940 and 1950. 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 A7: Placebo test - Random shift.

Panel A. Agriculture

	(1) 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
Migration exposure	-0.021	-0.003	0.003	0.004	0.003	-0.030*	-0.029
	(0.017)	(0.015)	(0.014)	(0.014)	(0.053)	(0.016)	(0.020)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.457	0.625	0.338	0.438	0.087	0.968	0.641
First-stage F	5.07	7.76	5.75	5.15	4.76	0.03	3.38
Counties	1,096	1,096	1,096	1,090	1,058	1,090	1,090

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Migration exposure	0.060**	0.039***	-0.005	0.035	0.045
	(0.029)	(0.013)	(0.030)	(0.035)	(0.030)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.157	0.413	0.160	0.144	0.173
First-stage F	2.95	3.66	2.58	2.63	2.62
Counties	1,096	1,096	1,065	1,096	1,096

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Migration exposure	0.035	0.037**	-0.002	0.015	0.030*	0.009	0.017	0.030*
	(0.030)	(0.018)	(0.030)	(0.030)	(0.016)	(0.008)	(0.012)	(0.018)
Fixed effects	0.041	0.039**	0.005	0.022	0.032*	0.011	0.019	0.033*
	(0.028)	(0.018)	(0.029)	(0.028)	(0.018)	(0.008)	(0.013)	(0.020)
R-squared	0.270	0.436	0.141	0.130	0.660	0.727	0.679	0.624
First-stage F	2.23	2.61	2.29	1.58	4.31	4.31	4.31	4.30
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096

Note: The table reports reduced-form estimates from Equation (3). The placebo migration exposure is created by using randomly shuffled Northern in-migration predictions as shifts. Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by standardized Northern migration exposure, relative to the omitted base years of 1940 and 1950. 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 A8: Robustness - Alternative share (1935-1940 migration).

Panel A. Agriculture

	(1) 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
Migration exposure	-0.036**	-0.002	-0.025	0.070***	0.120	-0.019	0.002
	(0.015)	(0.012)	(0.026)	(0.022)	(0.074)	(0.028)	(0.009)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.457	0.625	0.340	0.441	0.088	0.968	0.639
First-stage F	5.42	4.78	5.31	4.52	12.80	5.45	5.32
Counties	1,096	1,096	1,096	1,090	1,058	1,090	1,090

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Migration exposure	0.020	0.051***	0.122**	-0.003	0.032
	(0.025)	(0.015)	(0.055)	(0.040)	(0.032)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.156	0.414	0.160	0.144	0.173
First-stage F	5.28	5.40	3.36	4.77	4.74
Counties	1,096	1,096	1,065	1,096	1,096

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Migration exposure	0.152***	0.091***	0.084**	0.105***	0.013	0.010*	0.006	0.011
	(0.043)	(0.018)	(0.035)	(0.039)	(0.010)	(0.006)	(0.009)	(0.010)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.278	0.441	0.141	0.130	0.659	0.727	0.679	0.623
First-stage F	4.85	5.25	4.81	7.48	5.34	5.34	5.34	5.33
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096

Note: The table reports reduced-form estimates from Equation (3), using an alternative migration exposure constructed from the share based on "location 5 years ago" recorded in the 1940 Census (Ruggles et al., 2024a). Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by standardized Northern migration exposure, relative to the omitted base years of 1940 and 1950. 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 A9: Robustness - Alternative shift (actual number of in-migrants).

Panel A. Agriculture

	(1) 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
Migration exposure	-0.048**	0.008	-0.049***	0.109***	-0.017	0.008	0.018
	(0.020)	(0.011)	(0.012)	(0.034)	(0.072)	(0.020)	(0.013)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.457	0.625	0.344	0.443	0.087	0.968	0.640
First-stage F	18.01	17.70	17.87	18.18	13.96	1.22	17.97
Counties	1,096	1,096	1,096	1,090	1,058	1,090	1,090

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Migration exposure	0.083**	0.084***	0.170*	0.115	0.163***
	(0.037)	(0.016)	(0.101)	(0.079)	(0.059)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.157	0.416	0.161	0.144	0.175
First-stage F	15.81	16.52	12.53	14.52	14.82
Counties	1,096	1,096	1,065	1,096	1,096

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Migration exposure	0.140***	0.101***	0.076	0.090*	0.044***	0.016**	0.031**	0.042**
	(0.031)	(0.023)	(0.049)	(0.051)	(0.015)	(0.008)	(0.013)	(0.018)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.274	0.441	0.141	0.130	0.660	0.727	0.680	0.624
First-stage F	21.07	16.79	21.19	18.07	17.43	17.43	17.43	17.41
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096

Note: The table reports reduced-form estimates from Equation (3), using an alternative migration exposure constructed from actual in-migration rates in Northern counties rather than predicted in-migration rates. Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by standardized Northern migration exposure, relative to the omitted base years of 1940 and 1950. 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 A10: Robustness - Alternative shift (random forest prediction).

Panel A. Agriculture

	(1) 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
Migration exposure	-0.068***	-0.005	-0.064***	0.149***	-0.053	-0.006	0.022*
	(0.021)	(0.012)	(0.013)	(0.035)	(0.095)	(0.021)	(0.013)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.458	0.625	0.348	0.448	0.087	0.968	0.640
First-stage F	22.87	23.37	23.06	22.56	15.54	4.80	22.38
Counties	1,096	1,096	1,096	1,090	1,058	1,090	1,090

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Migration exposure	0.096**	0.100***	0.197**	0.119	0.179***
	(0.038)	(0.016)	(0.095)	(0.078)	(0.059)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.158	0.418	0.161	0.144	0.175
First-stage F	20.00	21.45	16.05	18.35	19.04
Counties	1,096	1,096	1,065	1,096	1,096

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Migration exposure	0.167***	0.126***	0.073	0.102**	0.062***	0.025***	0.046***	0.065***
	(0.031)	(0.022)	(0.048)	(0.049)	(0.015)	(0.008)	(0.013)	(0.018)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.277	0.444	0.141	0.130	0.662	0.728	0.680	0.626
First-stage F	23.99	21.57	24.15	20.21	22.55	22.55	22.54	22.53
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096

Note: The table reports reduced-form estimates from Equation (3), using an alternative migration exposure constructed from the predicted number of in-migrants based on a random forest algorithm instead of OLS regression. Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by standardized Northern migration exposure, relative to the omitted base years of 1940 and 1950. 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 A11: Robustness - Subsample (former confederate).

Panel A. Agriculture

	(1) 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
Migration exposure	-0.042*	-0.001	-0.053***	0.107***	0.050	-0.008	0.014
	(0.022)	(0.010)	(0.013)	(0.028)	(0.061)	(0.020)	(0.012)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.455	0.645	0.341	0.443	0.094	0.967	0.662
First-stage F	23.76	22.96	23.12	22.94	18.37	9.22	23.12
Counties	868	868	868	866	847	866	866

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Migration exposure	0.059*	0.069***	0.143*	0.098	0.159***
	(0.034)	(0.016)	(0.077)	(0.066)	(0.053)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.165	0.447	0.186	0.164	0.172
First-stage F	19.86	22.19	13.46	18.04	19.08
Counties	868	868	844	868	868

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Migration exposure	0.133***	0.098***	0.056	0.078*	0.059***	0.027***	0.045***	0.065***
	(0.028)	(0.021)	(0.045)	(0.044)	(0.013)	(0.008)	(0.012)	(0.016)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.312	0.494	0.154	0.141	0.738	0.797	0.751	0.701
First-stage F	26.50	22.16	26.47	21.83	22.68	22.68	22.68	22.65
Counties	860	868	860	860	868	868	868	868

Note: The table reports reduced-form estimates from Equation (3) by limiting the sample to former Confederate states (Texas, Arkansas, Louisiana, Tennessee, Mississippi, Alabama, Georgia, Florida, South Carolina, North Carolina, and Virginia). Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by standardized Northern migration exposure, relative to the omitted base years of 1940 and 1950. 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 A12: Robustness - Adding pre-period migration share weighted Northern exposure as controls.

Panel A. Agriculture

	(1) 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
Migration exposure	-0.040	-0.022*	-0.075***	0.140***	0.041	-0.033	0.017
	(0.024)	(0.012)	(0.015)	(0.028)	(0.063)	(0.024)	(0.013)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.457	0.626	0.363	0.452	0.087	0.968	0.641
First-stage F	25.37	23.74	23.52	24.49	18.43	14.63	25.77
Counties	1,096	1,096	1,096	1,090	1,058	1,090	1,090

Panel B. Manufacturing

	(1) Employment	(2) Number of establishment	(3) Capital spending	(4) Value added	(5) Annual payroll
Migration exposure	0.084**	0.091***	0.121	0.084	0.154***
	(0.038)	(0.018)	(0.088)	(0.072)	(0.057)
Fixed effects	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes
R-squared	0.157	0.417	0.161	0.144	0.174
First-stage F	22.44	23.57	15.06	20.64	21.77
Counties	1,096	1,096	1,065	1,096	1,096

Panel C. Wholesale and retail

	(1) Wholesale emp.	(2) Wholesale est.	(3) Wholesale sales	(4) Wholesale payroll	(5) Retail emp.	(6) Retail est.	(7) Retail sales	(8) Retail payroll
Migration exposure	0.176***	0.140***	0.099*	0.128**	0.074***	0.034***	0.054***	0.081***
	(0.031)	(0.022)	(0.050)	(0.049)	(0.015)	(0.008)	(0.013)	(0.018)
Fixed effects	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
R-squared	0.281	0.447	0.141	0.130	0.663	0.729	0.681	0.627
First-stage F	26.71	24.08	26.86	22.25	24.10	24.09	24.09	24.07
Counties	1,083	1,096	1,086	1,086	1,096	1,096	1,096	1,096

Note: The table reports reduced-form estimates from Equation (3) by adding time interacted values of the weighted average of Southern counties' Northern migration exposures as additional control variables. The weight is proportional to the pre-period (1910-1940) migration linkages within the South, as in Borusyak et al. (2023). Panels A to C correspond to the baseline results in Tables 2 to 4 with the full set of fixed effects and control variables. Each column reports the changes in the indicated outcome variable in logs for the years 1970 to 2010 by standardized Northern migration exposure, relative to the omitted base years of 1940 and 1950. Robust standard errors are clustered by county and reported in parentheses, and the first-stage Kleibergen-Paap robust F-statistics are reported. Stars represent: $* \; p < 0.10$, $** \; p < 0.05$, $*** \; p < 0.01$.

References

Abramitzky, Ran, Philipp Ager, Leah Platt 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*, 2002, *69* (4), 781–809.

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

Allen, Treb and Dave Donaldson, "Persistence and Path Dependence in the Spatial Economy," *Unpublished Manuscript*, 2022.

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

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

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.

Borusyak, Kirill, Rafael Dix-Carneiro, and Brian K. Kovak, "Understanding Migration Responses to Local Shocks," June 2023. Unpublished Manuscript.

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

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.

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.

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

Donaldson, Dave and Richard Hornbeck, "Railroads and American economic growth: A "market access" approach," *Quarterly Journal of Economics*, 2016, *131*, 799–858.

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

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.

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. V3.

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.

Jones, Ronald W., "The Structure of Simple General Equilibrium Models," *Journal of Political Economy*, 1965, *73* (6), 557–572.

_ and José A. Scheinkman, "The Relevance of the Two-Sector Production Model in Trade Theory," *Journal of Political Economy*, 1977, *85* (5), 909–935.

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.

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," 2024. [dataset].

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

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