The Origins of Regional Specialization*

Josh Morris-Levenson[†] Marta Prato[‡]

December 2022

Abstract

Why do US states specialize in different sectors? We document that employment specialization is highly persistent, which suggests that specialization may deviate from natural advantage, and reallocation could increase aggregate output. We develop a quantitative spatial model in which workers move across state-by-sector labor markets in response to exogenous changes in local fundamentals, as well as mobility frictions and labor market-specific idiosyncratic skills. We quantify the model with historical Census microdata that tracks workers' joint regional and sectoral mobility, which yield novel estimates of regional and sectoral mobility frictions as well as new evidence that workers carry state-, sector-, and pair-specific skills. We find limited scope for aggregate misallocation, but substantial idiosyncratic misallocation due to mobility frictions. Migration costs are the main barrier to workers' reallocation, but the benefit of lower frictions comes from new sector-specific opportunities.

Keywords: Regional Specialization, Migration, Comparative Advantage

*We are grateful to Rodrigo Adão, Jonathan Dingel, Erik Hurst, and Esteban Rossi-Hansberg for their guidance. We thank Adrien Bilal, Greg Kaplan, and Milena Almagro for helpful conversations; Greg Howard for an insightful discussion; and participants in the UEA 2021 North America Meeting, UChicago Applied Macro Seminar, Capital Theory Working Group, and Trade Working Group for helpful comments. Any omissions or errors are our own.

[†]Cornerstone. [‡]Yale university, Email: marta.prato@yale.edu

1 Introduction

US states specialize in different sectors in terms of employment, with services on the coasts, manufacturing in the Midwest, and agriculture in the Plains States. Moreover, these patterns are highly persistent, with significant correlations dating back 120 years (Table 1). We study the sources of persistence in regional specialization in order to shed light on the present distribution of economic activity. If specialization at a given point in time reflects the persistent effects of historical factors rather than contemporaneous natural advantage, then it may be possible to increase aggregate output and welfare by reallocating sectors across space. In addition, the same forces that give rise to persistence may prevent workers from finding the best match for their individual skills, another form of misallocation.

In order to understand why specialization persists, we first study how it evolves. We construct flows of workers across states and sectors from US Census microdata covering 1860-2020. Critically, we observe the interaction of regional and sectoral mobility in a panel of workers constructed from the Census Full Count data, new to the study of economic geography. We document two novel facts that motivate our analysis. First, when a sector expands its employment share in a state, it hires relatively more interstate migrants than locals, i.e., the share of migrants in the sector increases. Second, migrants switch sectors at a higher rate than non-migrants. The first fact suggests that interstate migration facilitates changes in local and aggregate sectoral composition. The second indicates an important interaction between regional and sectoral mobility decisions at the individual level.

Motivated by these findings, we quantify the sources of persistence in regional specialization using a dynamic spatial model with frictional labor mobility. The economy is comprised of region-by-sector labor markets and evolves as a series of static equilibria, starting from the observed labor allocation in 1860. In each period, changes in exogenous natural advantage across labor markets—including amenities and productivity—induce workers to depart from their labor market origin, i.e., their state of birth and father's sector.[1] Mobility is costly, so workers will move only if their idiosyncratic skill in the destination justifies the cost.[2] The accumulation of skill enhances productivity through agglomeration, and this externality persists over time. Mobility frictions and agglomeration generate persistence on top of (temporary) natural advantage. We quantify the sources of persistence, and then use the model to evaluate the implications of persistence and specialization for aggregate output. We study two forms of misallocation: aggregate and idiosyncratic. Aggregate misallocation arises from the (mis)alignment of employment relative to productivity. Idiosyncratic misallocation reflects the costs that inhibit workers from reaching the labor market that provides the best match for their individual skills.

We develop a general equilibrium model of state-by-sector labor markets connected through the movement of workers and national markets for local output. Workers' mobility decisions characterize labor supply. Workers are endowed with a state and sector of origin and must choose a state and sector in which to live and work for their single period

[1] We say father's sector because we identify workers' sectors of origin in the historical Complete Count Census data, where data restrictions confine our attention to white males. See Section 2.1 for details.

[2] We consider idiosyncratic skill rather than preference to avail ourselves of additional tools for estimation. See Section 3. This assumption matches the finding that migrants and sector switchers earn more than non-movers.

of working life.[3,4] The utility derived from a particular destination depends on its productivity (inclusive of agglomeration), its amenity value, and the cost of getting there from the worker's state and sector of origin. The first two factors correspond to the fundamentals mentioned above. Note that mobility frictions apply to pairs of labor markets; they may include interstate and intersectoral costs, and the cost of switching sectors may differ for migrants and non-migrants, a novel feature of our model.[5] Fundamentals and mobility frictions are common to all workers from a given origin. However, not all workers move to the same place. Therefore, we assume that workers differ in terms of idiosyncratic skill for each prospective state-by-sector destination. Skills are drawn from a Fréchet distribution, giving rise to a standard discrete choice from each origin (McFadden, 1974). Although skills are drawn anew in each generation, labor mobility frictions generate persistence in the supply of skill in each state-sector labor market. The combination of cross-labor market mobility frictions and labor market-specific skills is the essential innovation in our study of regional specialization.

On the labor demand side, we introduce a two-stage production structure. This structure will give rise to an expression for labor market fundamentals containing a regional component, a sectoral component, and a labor market-specific component. In the first stage, representative firms in each state-sector hire efficiency units of labor on a competitive labor market, resulting in output proportional to local sectoral productivity. Productivity is enhanced by an agglomeration externality in terms of efficiency units that persists over time.[6] In the second stage, a national final goods producer purchases intermediates from each sector and all states to produce a numeraire final good. Intermediates are differentiated by sector but not by state, so sectoral intermediates are sold at a common national price. Therefore, from the standpoint of workers, the common appeal of each state-sector labor market is equal to the product of state-specific amenities, sector-specific prices, and local sectoral productivity.

The discrete choice model delivers a gravity equation for labor flows across state-sector pairs that we estimate using standard techniques. The combination of discrete choice and gravity estimation distinguishes the two sources of endogenous path dependence in our model: mobility frictions and agglomeration. The latter is embedded in the labor market funda-

[3]In static models, imperfect labor mobility is captured by idiosyncratic preference shocks. Tying workers to an origin allows us to estimate mobility costs. This approach is on display in several recent papers, including Bryan and Morten (2019), Caliendo et al. (2019), Tombe and Zhu (2019), Oliveira and Pereda (2020), Allen and Donaldson (2020), and Eckert and Peters (2018).

[4]We also work with a version in which individuals work for two periods and are immobile in the second period. We make the latter assumption to avoid modeling forward-looking behavior. Instead, our one-period model provides an upper bound on the importance of mobility frictions in sectoral persistence, while the two-period model provides a lower bound.

[5]It is equally true that the cost of migrating may differ for sector-switchers and sector-stayers. We adopt the interpretation in the main text as a convention.

[6]We extend the model of Allen and Donaldson (2020) (AD) to an economy with multiple sectors. While we make use of some of the machinery from AD, we do not share their focus on formal notions of path dependence. In a single-sector economy, AD derive conditions under which changes in fundamentals alter the (unique) long-run steady-state of the economy. A multi-sector economy resists this formalization: a characterization of long-run behavior requires convergence to an equilibrium not only in space, but also in terms of sectoral composition. A model of structural transformation might accommodate convergence in sectoral composition, but this lies beyond the scope of our study.

mentals, which are estimated in the form of origin-year and destination-year fixed effects. Mobility frictions are identified by orthogonality to the fixed effects. This procedure owes its simplicity to the separability of mobility frictions and labor market fundamentals that obtains in competitive labor and product markets.[7] We recover arbitrary non-parametric mobility frictions and specify parametric components that we manipulate in counterfactuals. Novel to this paper, we allow the cost of switching sectors to differ for migrants and non-migrants. We find that the cost of switching sectors is lower for migrants, in line with the motivating facts.

We recover switching costs from flows across state-sector pairs derived from the Restricted Complete Count Census Data. We link individuals across Census years by name and connect fathers to sons, implementing the methods of Abramitzky et al. (2012).[8] However, these data are available only up to 1940. In order to extend our analysis beyond 1940, we supplement the Restricted data with the Public Use Microdata Samples from 1860-2020. These data lack the sector of origin, but nonetheless report aggregated flows from states of origin to state-sector destinations. We disaggregate these flows under the assumption that the switching frictions estimated in 1880-1940 applied also to the 1960-2020 period.[9,10] We estimate the gravity equation in this extended dataset to recover labor market fundamentals and time-varying geographic mobility frictions up to 2020. We identify a change in the structure of migration costs between 1940 and 1980: while migration costs remain constant overall, the distance cost shrinks. This reflect a 40% increase in the average distance of interstate moves.

Finally, we use the structure of labor demand to decompose the estimated fundamentals into regional amenities, sectoral prices, and local sectoral productivity. From the dynamics of productivity and labor supply we identify the parameters of agglomeration and persistence. In order to distinguish agglomeration from exogenous productivity, we construct an instrument for local sectoral labor supply based on 1860 stocks of immigrants from different national origins. We observe that immigrants from different origins exhibit distinct patterns of sectoral specialization, measured in 1940. Building on the insight of Ottinger (2020), we argue that these patterns reflect latent skills that became more relevant over time as aggregate sectoral composition changed, but may have supported the development of sectoral specialization in the states where these immigrants settled. We impute immigrants' sectoral comparative advantage to states based on their 1860 locations and interpret this as a shock to sector-specific skill.

We implement a pair of counterfactual exercises to decompose the two sources of persistence in our model—mobility frictions and path dependence in productivity. In the first counterfactual, we simulate the economy forward from 1860 without mobility frictions, so that the persistence of regional specialization reflects only natural advantage and endogenous investment. We find that persistence declines by 32% at a twenty-year time horizon, but

[7] This restriction accommodates a variety of extended labor demand structures, discussed in Section 3. These extensions affect the model's behavior in counterfactuals, but have no bearing on the output of the gravity estimation.

[8] Section 2.1 explains the limitations that restrict our attention to US-born white men in the linked sample.

[9] We maintain a parsimonious structure of switching costs to remain agnostic as to the identity of the origin and destination sectors. This minimizes the bias that would arise from applying the same frictions to two time periods containing different trends in sectoral composition.

[10] The motivating facts support this extrapolation. See Section 2.2.

converges to the baseline level of persistence over longer periods. In the second counterfactual, we remove mobility frictions once again, and additionally impose that local sectoral productivity reflects the baseline investment process. This modification has little impact on persistence relative to the first counterfactual. Although workers reallocate relative to the baseline, both simulations—the baseline economy and the frictionless case with endogenous investment—are governed by the same investment process and hence exhibit a similar amount of persistence. We conclude that mobility frictions explain about one-third of persistence in specialization and the dynamics of composite productivity—the combination of exogenous shocks and endogenous investment—explain the rest.

How does persistence affect aggregate output? Our model is designed to study employment, but we can also use it to measure the effect of mobility frictions on GDP. This relationship is *a priori* ambiguous and may vary across contexts. Labor mobility frictions prevent workers from moving to locations with superior productivity or amenities. Persistence could dampen aggregate output if the economy fails to take advantage of productive opportunities in different locations. On the other hand, persistence could tie workers to productive locations when they might otherwise choose to live in high-amenity locations.

We conduct two counterfactuals to evaluate the implications for aggregate output. First, we use the model to measure the value of specialization in each year of our sample. We reallocate efficiency units across sectors within each state so that each state shares the national sectoral composition. GDP falls by 3% in 1880 and by less than 1% in later years. This reflects the decline in specialization documented by Crafts and Klein (2021). Second, we measure GDP in the frictionless economy. GDP increases by 31% on average. This substantial increase is due to improved allocative efficiency at the individual level rather than changes in states' sectoral specialization.

Which margin of mobility generates these gains? Migration costs constitute the bulk of mobility frictions, and we find that removing migration costs alone achieves nearly the same benefit as removing all frictions. However, in additional counterfactuals, we show that the benefits of free interstate migration are due substantially to sector-switching. This finding suggests that policies aimed at reducing mobility frictions could generate large aggregate gains even in the highly mobile US economy. In particular, if policy could reduce the cost of leaving a disadvantaged area, this would allow outbound workers to find better opportunities in different sectors.

Related literature. This paper relates to several strands of literature at the intersection of spatial and labor economics.

Our study contributes most directly to the empirical and theoretical analysis of path dependence. At least since Krugman (1991) formalized the possibility of multiple equilibria in a spatial economy, economists have studied the influence of various historical accidents on spatial outcomes, particularly the distribution of population (e.g., Davis and Weinstein (2002), Bleakley and Lin (2012); see Lin and Rauch (2020) for a review). The empirical literature on path dependence has touched on the role of both physical (Bleakley and Lin, 2012) and human (Simon and Nardinelli, 2002) capital. Turning to the analysis of regional industrial composition, Ottinger (2020) measures the comparative advantage of immigrants in narrow manufacturing industries and shows that stocks of immigrants from high-skill origins predict industry employment growth in US counties. Allen and Donaldson (2020) provide a structural treatment of Bleakley and Lin (2012); we do so for Ottinger (2020).

Our dynamic structural analysis of specialization speaks to the literature on the reasons for industry location. These studies attribute specialization to natural advantage, interindustry linkages, and agglomeration. Ellison and Glaeser (1997) and Ellison et al. (2010) provide a seminal treatment of concentration in US manufacturing; Kim (1995) and Crafts and Klein (2021) its dynamics. Other papers study trends in specialization in terms of tasks (Michaels et al., 2018), occupations (Gervais et al., 2020), and business functions (Duranton and Puga, 2005). Our model emphasizes the role of agglomeration in sectoral specialization, motivated by an extensive empirical literature (Greenstone et al., 2010, Kline and Moretti, 2014, Hanlon and Miscio, 2017, Helm, 2019). We contribute novel estimates of agglomeration and persistence parameters for the United States during the period 1880-2020. Our structural approach strips out the influence of mobility frictions in our estimates, and we introduce an instrumental variable building on Ottinger (2020).

Recent studies of regional specialization use spatial models in the style of Redding and Rossi-Hansberg (2017) to capture the causes and consequences of industry collocation. Caliendo et al. (2018) study the implications of interindustry linkages and collocation for the propagation of local industrial shocks. Rossi-Hansberg et al. (2019) study agglomeration in skilled occupations and its implications for efficiency. We join these papers in taking a flexible approach to industry location, introducing some endogenous economic mechanisms to explain the location of industries and letting structural residuals take care of the rest. However, the aforementioned static models contain only a limited notion of mobility frictions that does not recognize the role of initial conditions. This is not to say that these papers ignore initial conditions altogether; the "persistence" of economic geography is captured in the structural residuals. We identify the component of those residuals that can be attributed to historical factors, namely stocks of skill and capital, which is possible only in a dynamic framework.

Three recent papers take a dynamic approach similar to our own. Eckert and Peters (2018) study structural transformation from agriculture to manufacturing in a spatial economy. Berkes et al. (2020) analyze the specialization of US cities in different patent classes, finding persistence in this form of specialization as well. Pellegrina and Sotelo (2021) document the contribution of migrants to the evolution of internal comparative advantage in Brazil. We emphasize labor mobility frictions across regions and sectors and their implications for persistence. None of these papers studies the cost of switching sectors. Pellegrina and Sotelo (2021) share our focus on the contribution of interstate migrants to local sectoral specialization, but in a very different setting. In Brazil, westward migration was key to realizing that region's specialization in new export crops. By contrast, in the US, Eckert and Peters (2018) show that changes in sectoral composition were orthogonal to changes in regional population. Hence, we hypothesize that migrants contribute to evolving comparative advantage through gross flows rather than net flows.

We contribute to a growing literature using discrete choice models to estimate mobility frictions. Our central contribution is to unite the estimation of switching and migration costs. Few papers have studied mobility along geographic and sectoral margins simultaneously, Caliendo et al. (2019) being the most notable example.[11] However, they estimate the two sets

[11] Eckert and Peters (2018) also model movement along both margins but omit sector switching costs.

of frictions in separate datasets. Our data allow us to observe the sector-switching behavior of migrants and non-migrants and thus estimate the interaction of geographic and sectoral mobility costs. We provide novel estimates of time-varying migration costs for the US, extending the approach of Bryan and Morten (2019), Tombe and Zhu (2019), and Oliveira and Pereda (2020), which study a single time period, study the evolution of the economy over a long period of time. Allen and Donaldson (2020) perform a similar exercise. We build on their analysis with a higher frequency (twenty years instead of fifty) and multiple sectors. We also add a fixed cost of migration in addition to a distance elasticity, which changes the results.

Our analysis also speaks to the literature on selective mobility and misallocation in the labor market. Previous work has studied misallocation across regions (Bryan and Morten, 2019) and occupations (Hsieh et al., 2019). Although it appears that these two forms of mismatch are closely connected (Young, 2013, Gollin et al., 2014), research has thus far assumed that regional and occupational/sectoral mobility frictions are independent. We study their interaction, finding that the cost of switching sectors is lower for migrants than for non-migrants. If geographic and sectoral mobility is motivated by the desire to find a better match for one's idiosyncratic skills, then workers might improve their match quality by moving not along one margin but both. A thorough analysis of this interaction and its consequences for misallocation is highlighted as an important direction for future research.

In the remainder of the paper, we present the data and motivating facts in Section 2, the model in Section 3, estimation and results in Section 4, and counterfactual results in Section 5. Section 6 concludes.

2 Data and Facts

2.1 Data

We describe the evolution of employment in the US from 1860 to 2020 at twenty-year intervals. We work with two datasets from the US Census: the Public Use Microdata Samples (Ruggles et al., 2021b) and the Restricted Complete Count Historical Census (Ruggles et al., 2021a). This section explains the key features of these data. Appendix A provides additional details.

The Public Use data contain cross-sections of workers and record their state of residence, sector of employment, and state of birth. The Restricted data, available up to 1940, allow us to link fathers and sons and thereby identify a worker's sector of origin as well. We group workers into four broad sectors: agriculture, manufacturing, services, and a residual sector.[12] Finer delineations introduce many zeros in early decades. Throughout, we denote states by r and sectors by s. Workers move across state-sector pairs, $i = (r_i, s_i)$ or $j = (r_j, s_j)$. $L_{i,j,t}$ denotes the number of workers residing in j at time t who came from i. The Restricted data record fully detailed flows across state-sector pairs. The Public Use data record aggregated flows, $L_{r_i,j,t}$, omitting the sector of origin.

[12] Appendix Table A2 presents the sectors and their component industries.

We study lifetime mobility of workers across states and sectors, i.e., relative to the previous generation. We restrict attention to workers aged 20-39 so that there is no overlap between cohorts of workers in each twenty-year period. Although the Restricted data also contain the mobility decisions of adult workers, we focus on the lifetime mobility of young workers for three reasons. First, lifetime mobility is available in both the Restricted and Public Use data. Second, lifetime mobility is conceptually consistent with our focus on adjustment to changes in national sectoral technology, which play out over long periods of time.[13] Third, lifetime mobility supports a simple model of myopic workers, which we favor for its transparency. It is important to note that young workers are more mobile than older workers across both regions (Greenwood, 1997) and sectors (Adão et al., 2020) (and see Table A3). By ascribing their behavior to the entire workforce, we overstate mobility and hence the importance of labor mobility frictions in persistence. To address this concern, we show that our results are robust to the choice of age group. We replicate the motivating facts and the structural results using a sample of workers aged 40-59. In addition, we will replicate our structural results in a model where workers live for two periods and are immobile in the second period.

We make additional adjustments to harmonize the two datasets, since we will use them in conjunction to quantify the model.[14] We adjust the Restricted data so that $\sum_s L_{(r_i,s),j,t}$, calculated in the Restricted data, is equal to $L_{r_i,j,t}$, observed in the Public Use data. We also impose balance restrictions needed to identify the model. We restrict attention to state-years in which all four sectors have nonzero employment. Within these, we study a balanced panel of employment flows in terms of destinations (where people are working) and origins (where people are born). The final dataset contains 37 states and roughly 80% of the workforce.

Before proceeding to the motivating facts, we document the persistence of specialization in Table 1. Specialization declines over time (Crafts and Klein, 2021), so we calculate persistence based on ranks rather than levels of employment. We calculate the location quotient (LQ) of each state-sector, equal to the sector's share of employment in the state divided by the state's share of national population:

$$LQ_{j,t} = \frac{L_{j,t}}{L_{r,t}} \Big/ \frac{L_{s,t}}{L_t}$$

We then rank state-sectors by LQ within each sector and run the following regression:

$$\log \mathrm{RankLQ}_{j,t} = \alpha + \rho_h \log \mathrm{RankLQ}_{j,t-h} + \delta_{r,t} + \gamma_{s,t} + e_{j,t}$$

ρ_h is positive and statistically significant out to $h = 120$. Appendix Table A1 shows the same regression but with $\log L_{j,t}$ in place of $\mathrm{RankLQ}_{j,t}$. The results are similar.

[13] Adult mobility is appropriate in studies of adjustment to shocks over shorter time horizons, e.g., Caliendo et al. (2019).

[14] Full details of harmonization can be found in Appendix A.3.

Table 1: Persistence of Regional Specialization

	$h = 20$	$h = 40$	$h = 60$	$h = 80$	$h = 100$	$h = 120$	$h = 140$	$h = 160$
Persistence	0.869***	0.705***	0.550***	0.401***	0.261***	0.131***	0.0442	0.0384
	(0.0139)	(0.0185)	(0.0225)	(0.0250)	(0.0264)	(0.0278)	(0.0309)	(0.0537)
N	1182	1034	886	738	590	442	294	146

Data source: IPUMS 1860-2000, 2014-2018 ACS. Notes: Rank-rank elasticity of location quotient (LQ). We compute LQ for each state-sector, rank these within each sector, and regress log rank on its lag at horizon h with state-year and sector-year fixed effects. Each column of the table represents a different regression. Each coefficient averages across the years in which that horizon is observed.

2.2 Motivating Facts

In order to understand the persistence of states' sectoral specialization, we study how it changes, leveraging our data on employment flows across state-sector pairs. The following facts demonstrate an interaction between interstate migration and sector switching. Overall, it appears that interstate migrants contribute disproportionately to changes in local sectoral composition. Equivalently, sector-switchers contribute disproportionately to interstate migration. This suggests that costs of interstate migration and sector-switching could contribute to persistence.

2.2.1 Migrants Contribute Disproportionately to Local Sectoral Growth

When a sector expands its employment share in a state, it hires more migrants than non-migrants. We illustrate this fact with a regression. Let $M_{j,t} = \sum_{m:r_m \neq r_j} L_{m,j,t}$ denote the stock of migrants working in state-sector j at time t. We run the following regression:

$$\frac{M_{j,t}}{L_{j,t}} - \frac{M_{j,t-1}}{L_{j,t-1}} = \beta_0 + \beta_1 \Big(\frac{L_{j,t}}{L_{r_j,t}} - \frac{L_{j,t-1}}{L_{r_j,t-1}} \Big) + \delta_{r_j,t} + \gamma_{s_j,t} + v_{j,t} \tag{1}$$

where $\delta_{r_j,t}$ and $\gamma_{s_j,t}$ represent state-year and sector-year fixed effects and $v_{j,t}$ is an error term. We estimate $\beta_1 > 0$, as shown in Table 2. This illustrates the importance of interstate migrants to the evolution of regional comparative advantage.

This result reflects a difference in behavior between migrants and non-migrants, rather than selective migration into particular states. $\beta_1 > 0$ is not driven by in-migration to specific states, sectors, or labor markets. In Appendix Table A13, we repeat the regression while restricting the sample to growing or shrinking states, growing or shrinking sectors, and growing or shrinking labor markets. We estimate a positive coefficient in each case. It is not the case that workers move across states in order to join expanding sectors or vice versa. This echoes the finding of Eckert and Peters (2018): although the US experienced substantial changes in the allocation of workers across both states and sectors (see Table A11), these shifts were largely orthogonal to each other. Given this context, the regression indicates a difference between migrants and non-migrants rather than a difference between states that were receiving more or less migrants at a given point in time. Indeed, all states experienced gross flows far in excess of net flows (see Table A12). The contribution of interstate migrants to changing specialization is realized through gross rather than net migration.

In light of this historical context, this regression serves as a test of differential switching behavior on the part of migrants as compared to non-migrants beyond 1940, when this behavior cannot be directly observed. Appendix B studies a two-region, two-sector model with mechanical movement across labor markets. Workers move back and forth across states independent of sectoral demand; and back and forth across sectors independent of interstate flows, in keeping with the historical patterns discussed above. We then consider a shock to the economy that increases the employment share of one sector in both states. We show that $\beta_1 > 0$ obtains only when migrants switch sectors at a higher rate than non-migrants. We estimate equation (1) separately for 1880-1940 and 1960-2020 and find $\beta_1 > 0$ in both periods. Consider columns (2) and (3) of Table 2, migrants' sector-switching advantage is evident in the early period and somewhat attenuated in the later period. The next fact studies migrants' sector-switching behavior directly in the Restricted data.

Table 2: Migrant fraction regression

	(1)	(2) 1880-1940	(3) 1960-2020
$\Delta \frac{L_{j,t}}{L_{r,t}}$	0.17^{***}	0.18^{***}	0.13
	(0.05)	(0.06)	(0.09)
N	1184.00	592.00	592.00
R2	0.82	0.81	0.72
F	11.42	8.78	2.17

Notes: This table shows the coefficient from a regression of the change in the migrant fraction (share of migrants in state-sector j) on the change in the sector's employment share in the state. Each column is a separate regression. The first column uses changes from 1880-2020; the others restrict to the time period indicated in the title. Standard errors are clustered by state-year and sector-year and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

2.2.2 Migrants switch sectors at a higher rate than non-migrants

We compare the sector-switching behavior of migrants and non-migrants directly in the Restricted data. On average from 1880-1940, non-migrants switch sectors at a rate of 49%, while migrants switch at a rate of 60%—23% more, on average. Overall, migrants account for 25% of the workforce and 29% of sector-switchers. The regression above suggests that migrants' sector-switching advantage holds true beyond 1940 as well, when migrants' sector-switching behavior is not directly observed.

Migrants' sector-switching advantage is observed directly from 1880-1940, a period of secular decline in the employment share of agriculture and a related movement from rural to urban areas. These trends could drive migrants' sector-switching advantage if workers who move from farm to city—which entails switching sectors—are also more likely to move across states. We reproduce the statistics in the previous paragraph, this time excluding agriculture from the sample of flows. The difference in switching rates is somewhat muted when agriculture is excluded. The average switching rates in this sample are 49% for non-migrants and 55% for migrants. Migrants' switching rate exceeds that of non-migrants by 16% on average. Migrants account for 25% of the non-agricultural workforce and 26% of sector-switchers among non-agricultural sectors. We conclude that the rural-urban movement implied by the decline of agriculture contributes to migrants' estimated sector-switching advantage, but migrants still maintain their switching advantage among non-agricultural sectors.

3 Model

This section develops a model to evaluate the interaction between migration and sectoral reallocation in general equilibrium. The economy is comprised of regions $r \in \mathcal{R}$ and sectors $s \in S$ populated by an exogenous measure of workers $L_t = 1$ in every period. A region-sector pair comprises a labor market $i = (r_i, s_i)$; there are $N =\mid \mathcal{R} \mid \cdot \mid \mathcal{S} \mid$ labor markets.

Time is discrete. We consider an overlapping generations framework in which each period spawns a fresh cohort of workers who live for one period.[15] Workers choose the destination that maximizes their utility. The choices of each generation determine the initial allocation in the next period.

Section 3.1 specifies three components of workers' utility: labor market fundamentals (amenities and productivity), mobility frictions, and idiosyncratic skill shocks. Mobility frictions constitute one source of persistence. Section 3.2 describes the labor demand side of the economy. We introduce persistence in productivity arising as an agglomeration externality.

3.1 Preferences and Location Choice

Workers, indexed by n, are endowed with an origin i and choose the destination j that maximizes their utility, given by,

$$U_{i,j,t}(n) = V_{j,t}\mu_{i,j,t}z_{i,j,t}(n) \tag{2}$$

where $V_{j,t}$ is a labor market fundamental, $\mu_{i,j,t}$ is an iceberg mobility friction, and $z_{i,j,t}(n)$ is an idiosyncratic shock representing the efficiency units supplied by worker n to destination j. We discuss each component in turn.

$V_{j,t}$ summarizes the appeal of destination j relative to all others (ignoring $\mu_{i,j,t}$ and $z_{i,j,t}(n)$ for the moment). In particular, destinations are differentiated by a regional amenity $B_{r,t}$, common to all labor markets within a state; and a local sectoral skill price $w_{j,t}$ paid per efficiency unit. Labor market earnings are the only source of income and spent entirely on the consumption of a numeraire good. We define the labor market fundamental,

$$V_{j,t} = B_{r,t}w_{j,t}. \tag{3}$$

We assume a numeraire consumption good and exogenous amenities for simplicity. However, under the assumption that workers are price-takers, $V_{j,t}$ can embed various extensions, including trade costs, nontradable goods, land (as housing and/or in production), and income from firms' profits. $V_{j,t}$ still summarizes workers' mean utility per efficiency unit of working in labor market j.

$\mu_{i,j,t}$ discounts the utility received at destination j when moving from origin i. Frictions may be associated with moving across states, across sectors, and across state-sector

[15] We elide forward-looking behavior to maintain tractability and focus on mobility frictions.

pairs. These could include pecuniary costs as well as opportunity costs incurred by leaving behind one's connections in a state, sector, or state-sector pair (Greenwood, 1997). To build intuition, consider the potential connection between mobility frictions and migrants' sector-switching advantage, documented in Section 2.2. The fact that migrants switch sectors at a higher rate than non-migrants could indicate that the cost of switching sectors is lower for migrants compared to non-migrants. Opportunity costs could explain this pattern. Migrants who do not switch sectors sacrifice state- and pair-specific connections. Sector-switchers who do not migrate lose sector- and pair-specific connections. For a worker who has already migrated, switching sectors cuts off only sector-specific connections; they gave up their pair-specific advantage when they migrated.

Mobility frictions are assumed to have parametric and nonparametric components:

$$\mu_{i,j,t} = \mu_{i,j,t}^{Par}\epsilon_{i,j,t} \tag{4}$$

both of which will be estimated to fit observed labor flows. $\mu_{i,j,t}^{\text{Par}}$ are those mobility frictions associated with observables. We defer the full specification to estimation in Section 4. The nonparametric component $\epsilon_{i,j,t}$ represents a cost or preference shifter that applies equally to all agents moving along a particular path. For example, $\epsilon_{i,j,t}$ could represent an information treatment, such as a local newspaper article boosting a particular destination. Critically, $\epsilon_{i,j,t}$ is known to agents prior to making their migration decision, but unobservable to the econometrician. $\epsilon_{i,j,t}$ represents the modeler's uncertainty about migration decisions relative to the mechanical predictions of our gravity model.

Workers are homogeneous at birth and $V_{j,t}$ and $\mu_{i,j,t}$ affect all workers at a given origin equally. However, in the data, not all workers from i choose the same j. Therefore, we assume that workers receive idiosyncratic shocks $z_{i,j,t}(n)$ that represent their efficiency units in labor market j. We assume that $z_{j,t}(n)$ is drawn i.i.d. from a Fréchet distribution:

$$F(z_{1,t}, z_{2,t}, ..., z_{N,t}) = \exp\Big(-\sum_j z_{i,j,t}^{-\nu}\Big) \tag{5}$$

Workers' skills may have state-, sector-, and pair-specific components, but we focus exclusively on the latter. This facilitates our analysis of persistence. Under independent draws, the probability that a worker from i chooses j is given by,

$$\lambda_{i,j,t} = \frac{(V_{j,t}\mu_{i,j,t})^{\nu}}{\sum_k (V_{j,t}\mu_{i,j,t})^{\nu}}. \tag{6}$$

ν has three interpretations. From equation (5), ν is inversely related to the dispersion of z. As ν increases, a worker's vector of z draws becomes more compressed. From equation (6), ν can be interpreted as a migration elasticity. As ν increases, smaller values of $V_{j,t}\mu_{i,j,t}$ are needed to generate the same choice probability. Workers become more sensitive to differences between destinations because they are more sensitive to deviations in their own skill draws. Note that changing the value of ν has no effect on the size of $V_{j,t}$ and $\mu_{i,j,t}$ relative to each other.[16]

[16] We return to this discussion in Section 4.

Under the assumption that $z_{i,j,t}(n)$ represents a workers' skill, rather than taste, ν also governs the strength of selection in our model. The expected value of $z_{i,j,t}(n)$ for a worker who chooses to move from i to j is given by $\bar{\Gamma}\lambda_{i,j,t}^{\frac{-1}{\nu}}$ where $\bar{\Gamma} = \Gamma(1 - \frac{1}{\nu})$. Therefore, the average wage paid to workers who move along this path is equal to,

$$\mathbb{E}[\text{wage}_{i,j,t}(n) \mid i \to j] = \bar{\Gamma}\lambda_{i,j,t}^{\frac{-1}{\nu}} w_{j,t} \tag{7}$$

where $w_{j,t}$ is the skill price in labor market j. The model predicts that workers moving along paths with lower probability will earn higher wages on average. This is intuitive: higher cost or lower return moves must be justified by higher idiosyncratic skills. As ν gets larger, the slope of the relationship between wages and $\lambda_{i,j,t}$ gets flatter, i.e., selection gets weaker.

The Fréchet assumption also carries a useful implication for the variance of wages along a given path that we exploit in estimation.

$$\frac{\text{Var}\big(\text{wage}_{i,j,t}(n)\big)}{\mathbb{E}[\text{wage}_{i,j,t}(n)]^2} = \frac{\Gamma\big(1 - \frac{2}{\nu}\big)}{\Gamma\big(1 - \frac{1}{\nu}\big)^2} \tag{8}$$

As ν increases and $z_{i,j,t}$ becomes less dispersed, the variance of wages shrinks.

Finally, labor supply in each labor market is characterized by the total flow of efficiency units:

$$S_{j,t} = \bar{\Gamma} \sum_i \lambda_{i,j,t}^{1-\frac{1}{\nu}} L_{i,t-1} \tag{9}$$

which is the product of the flow of workers $\lambda_{i,j,t}L_{i,t-1}$ and their average skill $\bar{\Gamma}\lambda_{i,j,t}^{\frac{-1}{\nu}}$ summed over all region-sector origins.

Discussion: These are the essential elements of our model of labor supply. We make two assumptions that merit further discussion. First, we assume that workers live only one period, which forecloses forward-looking decisions. This is not such a great departure from the literature: many studies of migration make the same assumption (Allen and Donaldson, 2020, Bryan and Morten, 2019, Pellegrina and Sotelo, 2021) and the endogenous growth studies building on Desmet and Rossi-Hansberg (2014) make assumptions so that forward-looking agents end up making static decisions. With myopic decisions, workers may be more willing to migrate in response to a given difference in fundamentals. As a result, we may overestimate labor mobility frictions relative to a dynamic model (e.g., Bayer et al., 2016). Studies of dynamic decisions uniformly focus on short-run adjustment to shocks within a worker's lifetime.[17] The twenty-year time horizon of our data discounts the importance of dynamic considerations for workers in our framework.

The second assumption is that skill shocks are i.i.d. across generations. One might wish to account for the possible hereditary transmission of skills. This concept appears in our model in the form of cross-sector mobility frictions. Further refinement requires additional types of workers within each sector. At the extreme, one might allow the distribution of skill draws at each state-sector of origin to evolve endogenously. This raises a number of

[17] See Artu et al. (2010) and Dix-Carneiro (2014) for sectoral mobility and Kennan and Walker (2011) for migration.

questions. How does a father's skill in one sector translate to the child's skill in another? How does the stock of skill contribute to the skills of children, either in the same sector or other sectors? These are difficult questions to address without a panel of wages. Our data contain only a cross-section in 1940. In that cross-section, we find that sector of origin has little bearing on wages. These questions of intergenerational skill transmission is left to future research.

3.2 Production

Labor demand arises from a two-stage production structure. In the first stage, representative firms in each region-sector produce intermediates. Sectoral intermediates are homogeneous in the sense that they are not differentiated by location. In the second stage, a national final goods producer combines the sectoral intermediates into a consumption good. Both intermediates and final goods are traded without cost. Therefore, sectoral intermediates from any state are sold at a common sectoral price. The consumption good is treated as the numeraire.

Intermediate goods producers hire labor in a competitive market. Productivity benefits from agglomeration spillovers external to the firm that persists over time. This proxies for physical capital and introduces another form of persistence in the model.

3.2.1 Intermediate Goods Production

Each state-sector contains a representative firm using embodied skill $S_{j,t}$ to produce output:

$$X_{j,t} = P_{s,t} A_{j,t} S_{j,t} \tag{10}$$

where skill S is given by equation (9). The firm hires efficiency units at market wage $w_{j,t}$. It's first-order condition is given by,

$$w_{j,t} = P_{s,t} A_{j,t} \tag{11}$$

Agglomeration: Local sectoral productivity benefits from agglomeration spillovers, which have dynamic effects through path dependence. In particular, we assume that,

$$A_{j,t} = \bar{A}_{j,t} S_{j,t}^{\alpha_1} A_{j,t-1}^{\alpha_2} \tag{12}$$

$\bar{A}_{j,t}$ is the exogenous component of productivity. It represents the geographic suitability of state r_j for sector s_j at time t. As sectoral technology changes over time, the characteristics of a suitable location may change as well. Historical accidents may also contribute to $\bar{A}_{j,t}$.

Productivity is further augmented by agglomeration and persistence. Agglomeration is an externality in the model: atomistic workers and firms do not recognize their contribution to productivity. The accumulation of skill in a labor market results in spillovers that incentivize investment in durable, sector-specific capital.[18] This might take the form of plants and equipment or public infrastructure. Lacking data on capital investment, we are confined to this reduced form approach and we elide forward-looking behavior for simplicity.

[18] Appendix F presents a formal model of endogenous investment that is isomorphic to equation (12).

3.2.2 Final Goods Production

Intermediates are aggregated into a final good at the national level. Within each industry, output from different locations are perfect substitutes so that $\bar{X}_{s,t} = \sum_{j:s_j=s} X_{j,t}$. The final goods production function is,

$$Y_t = Z_t\Big(\sum_s \eta_s^{\frac{1}{\sigma}} \Big\{\bar{X}_{s,t}\Big\}^{\frac{\sigma-1}{\sigma}}\Big)^{\frac{\sigma}{\sigma-1}} \tag{13}$$

where η_s is a demand shifter. This production function maps the sectoral composition of aggregate employment in the data to the relative productivity of the intermediate producing sectors in the model. We will assume that $\sigma > 1$, so that inputs are substitutes. Therefore, a sector that exhibits a relatively higher rate of productivity growth expands it share of national employment (Herrendorf et al., 2013).

Standard CES results imply,

$$P_{s,t}X_{s,t} = P_{s,t}^{1-\sigma} P_t^{\sigma-1} Y_t \tag{14}$$

where

$$P_t = \Big(\sum_k P_{s,t}^{1-\sigma}\Big)^{\frac{1}{1-\sigma}}$$

Discussion: Our production specification serves two purposes. First, the two-stage structure and the assumption of free trade isolate the industry-specific component of workers' employment decision. Shifts in national employment composition are driven by changes in prices $P_{s,t}$. During calibration below, prices will find their reflection in local sectoral productivity, so that $A_{j,t}$ captures comparative advantage.

Second, the two-stage structure distinguishes the determinants of aggregate and local productivity. Ours is a model of exogenous aggregate growth, but aggregate and local output may change in counterfactuals. During calibration, we will assume exogenous GDP growth at a 2% annual rate. This will serve to scale the exogenous fundamentals, $\bar{A}_{j,t}$. These are fundamental productivity shocks, "historical accidents" whose provenance lies outside the model, and we hold them fixed in counterfactual analysis. Changes in other fundamentals, such as mobility frictions, will cause labor to deviate from its original allocation, causing changes in the endogenous component of local sectoral productivity and aggregate output. Our model measures the extent to which the spatial allocation of economic activity can generate deviations from trend growth. But these deviations cannot fundamentally alter the growth path.

Stripping out the aggregate trend in local productivity facilitates the analysis of agglomeration and path dependence as they pertain to local sectoral specialization. These processes, captured by the parameters α_1 and α_2, represent the role of physical capital—as opposed to workers' skills—in the persistence of specialization. Under the assumption of exogenous growth in aggregate productivity, $A_{j,t}$ represents comparative rather than absolute advantage. Therefore, path dependence applies to comparative advantage rather than productivity growth per se, as in the endogenous growth spatial frameworks of Desmet and Rossi-Hansberg (2014) and Desmet et al. (2018). Spatial reallocation in a given sector may increase output by improving the alignment of production and productivity. However, it will not alter the growth path of that sector.

Firms, like workers, are assumed to be myopic. This assumption affects the decomposition of local sectoral productivity into its exogenous and endogenous components. In the calibration procedure described in Section 4, $A_{j,t}$ is derived without reference to the problem of the intermediate goods producer—another convenient separability provided by the two-stage production structure. We quantify the endogenous component of productivity through the parsimonious and easily estimable equation (12), which is the centerpiece of the intermediate goods producer's problem. Investment in our model is simply a device to motivate a particular statistical model of local sectoral productivity.

3.3 Equilibrium

Given parameters, geography $\{B_{r_j,t}, A_{j,t}\}$, total population L_t, and initial labor allocation $L_{j,1860}$, an equilibrium is a sequence of prices $\{P_{s,t}\}$ and labor allocations $\{L_{j,t}\}$ such that,

1. The market for efficiency units clears, per equation (9).
2. The market for intermediate goods clears:

$$\bar{X}_{s,t} = P_{s,t}^{-\sigma} P_t^{\sigma-1} Y_t = \sum_j A_{j,t} S_{j,t}$$

3. Local output equals payments to labor:

$$P_{s,t} X_{j,t} = w_{j,t} S_{j,t} \tag{15}$$

4. Aggregate output equals payments to labor:

$$Y_t = \sum_j w_{j,t} S_{j,t} \tag{16}$$

3.4 An illustration of the labor market

This section illustrates the labor market as described by our model. We begin by presenting the labor supply elasticity and relating it to the facts presented in Section 2.2. We then present an illustrative model and discuss the implications of selection in our framework.

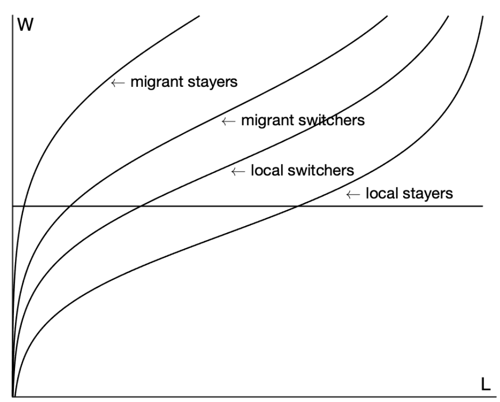
Labor supply elasticity: Our model parses the returns to and costs of migrating and switching sectors. All workers face a common wage per efficiency unit—a perfectly elastic labor demand curve represented by $w_{j,t}$—regardless of origin. An increase in this skill price enhances the appeal of destination j with elasticity ν, also common to all workers regardless of origin. However, the elasticity of labor supply from origin i to destination j varies across origins due to labor mobility frictions. In our model, the elasticity of labor supply from origin i with respect to an increase in the skill price $w_{j,t}$ is given by,

$$\text{Elasticity}_{i,j,t} = \nu \frac{L_{i,j,t}}{L_{j,t}} (1 - \lambda_{i,j,t})$$

The elasticity is decreasing in the choice probability. Higher mobility costs imply a lower choice probability and thus a higher elasticity of labor supply, in line with the finding of Section 2.2 that expanding sectors hire more migrants than non-migrants. The aggregate labor supply elasticity to labor market j is equal to the average of the origin-specific elasticities, weighted by the employment share of each origin in the destination.

Illustrative model: Consider an economy with two regions and two sectors, for a total of four labor markets. All labor markets are initially symmetric in terms of fundamentals and employment. Figure 1 illustrates the labor market in region 1, sector 1. Labor demand is given by the horizontal line. There are four sources of labor supply to this market, hence four labor supply curves, which come from the discrete choice structure of the workers' problem. The four curves differ because of labor mobility frictions. The workers facing the lowest mobility cost comprise the largest share of the local workforce, and the corresponding labor supply curve intersects labor demand farther to the right. Local stayers face no cost. Local switchers face a switching cost. Migrants face even higher mobility frictions. The aggregate labor supply curve to this destination (not shown) is a weighted average of these four curves.

Figure 1: Labor market in 2-by-2 model



Notes: Labor supply curves to a given labor market in a model of two locations and two sectors. Each curve is derived from the model-implied choice probability, equation (6).

Importantly, in this calibration, migrant stayers face a higher cost than migrant switchers. This is in line with the fact that migrants switch industries at a higher rate than non-migrants, though it is not strictly necessary. That the cost of switching is relatively smaller for migrants than non-migrants is sufficient to produce this result.

Workers are less likely to choose a destination that entails higher mobility costs. In our model where the Fréchet draw represents skill, workers will be positively selected in proportion to migration costs (see equation (7)). Workers who make costly moves must be relatively more skilled at their destination. In the model, differences in wages of, say, migrants relative to non-migrants, for example, are attributed entirely to selection. There is no role for demand-side factors.

This model implies a particular narrative about the migrant switchers and non-switchers. Consider two migrants moving from New York to California. Both migrants will work in manufacturing. Worker a is the son of a farmer, and worker b is the son of a factory worker. Worker a—the switcher—is negatively selected relative to worker b, in the sense of having lower expected skill. Intuitively, worker a works on the production line, whereas worker b, who has a manufacturing pedigree, is more likely to work as a manager.

This discussion clarifies the role of the model's parameters in describing the labor market. The Fréchet elasticity, ν, governs the strength of selection; and the mobility frictions control the elasticity of labor supply.

4 Estimation and calibration

This section describes the estimation of model parameters and economic fundamentals. We proceed in three steps. First, we distinguish mobility frictions $\mu_{i,j,t}$ from fundamentals $V_{j,t}$ by estimating the workers' discrete choice. Second, we decompose the fundamentals into their components, $B_{r,t}$, $P_{s,t}$, and $A_{j,t}$. Finally, we use the resulting values of $A_{j,t}$ to estimate the elasticities of agglomeration and path dependence, α_1 and α_2. All results are presented in Section 4.5.

Two parameters are calibrated externally. We set the Fréchet elasticity $\nu = 3.5$ and the final goods elasticity of substitution $\sigma = 4$. We discuss the Fréchet elasticity following the estimation of the worker's problem.

4.1 Labor mobility frictions and local sectoral fundamentals

In this section, we estimate local sectoral fundamentals $V_{j,t}$ and labor mobility frictions across state-sector pairs $\mu_{i,j,t}$. We infer these objects from a gravity equation based on the model-implied choice probability, equation (6). We begin in 4.1.1 by applying the gravity equation to flows across state-sector pairs observed in the Restricted data, 1880-1940, using standard techniques (Guimaraes et al., 2003, Silva and Tenreyro, 2006, Fally, 2015). This delivers local sectoral fundamentals and mobility frictions up to 1940. We discuss the parametric specification of $\mu_{i,j,t}^{\text{Par}}$ and the identification of its parameters.

In 4.1.2, we extend our analysis beyond 1940 using the Public Use data. The Public Use data report flows from region of origin to state-sector destinations, but omit the sector of origin. Although the Public Use data cannot speak to *sector-related* frictions, the aggregated flows contain information about local sectoral fundamentals and *geographic* mobility frictions from 1880-2020. However, we cannot immediately apply the estimating equation implied by the model; estimation in the raw Public Use data is feasible only by ignoring sector-related frictions, which could bias our estimates of geographic mobility frictions. Our solution to this problem is to disaggregate the Public Use data *as if* the sector-related frictions estimated in the 1880-1940 data apply also to the post-1940 period. We can then infer fundamentals and frictions from the disaggregated data using the model-implied gravity equation.

4.1.1 Gravity Estimation in the Restricted Data

This section describes the estimation of the worker's discrete choice problem in the Restricted data (though we will use this equation again in 4.1.2). The model-implied choice probability, equation (6), motivates a gravity equation for labor flows across state-sector pairs. We estimate this equation by Poisson Pseudo-Maximum Likelihood (PPML) to recover the parameters of the worker's problem. First, we develop a statistical representation of the model. Define,

$$\omega_{j,t} = \nu \log V_{j,t} \tag{17}$$

$$\phi_{i,t} = -\log \sum_k (V_{k,t}\mu^{\text{Par}}_{i,k,t})^{\nu} \tag{18}$$

Then the choice probability can be written,

$$\lambda_{i,j,t} = \exp\{\omega_{j,t} + \nu \log \mu^{\text{Par}}_{i,j,t} + \phi_{i,t}\} = \exp\{D'_{i,j,t}\delta\}, \tag{19}$$

and the PPML moment condition is,

$$\mathbb{E}\Big[\big(\lambda_{i,j,t} - \exp\{D'_{i,j,t}\delta\}\big)D_{i,j,t}\Big] = 0 \tag{20}$$

where $D_{i,j,t}$ represents the vector $\omega_{j,t}, \mu_{i,j,t}, \phi_{i,t}$. PPML returns model-consistent values of $V_{j,t}$ as defined in equations (17) and (18). $\mu^{\text{Par}}_{i,j,t}$ and $\epsilon_{i,j,t}$ are identified by orthogonality to the origin and destination fixed effects. The nonparametric friction is obtained as a multiplicative residual: $\epsilon_{i,j,t} = \lambda_{i,j,t}/\hat{\lambda}_{i,j,t}$ where $\lambda_{i,j,t}$ is the empirical choice probability and $\hat{\lambda}_{i,j,t}$ is the prediction obtained by minimizing equation (20). It is easy to verify that this residual is consistent with the model.[19]

The parameters of $\mu^{\text{Par}}_{i,j,t}$ are identified by orthogonality of the regressors (specified below) to $\epsilon_{i,j,t}$, conditional on the origin-time and destination-time fixed effects. Precise identification of $\mu^{\text{Par}}_{i,j,t}$ relative to $\epsilon_{i,j,t}$ is not critical to our analysis.[20] Parametric frictions $\mu^{\text{Par}}_{i,j,t}$ will help connect the model and the data and allow us to assess the aggregate implications of mobility frictions across regions, sectors, and their interaction. We do not aim to study the specific channels underlying mobility frictions. Rather, our estimates potentially capture various explanations that apply along different margins. We discuss some of these channels below, after specifying the components of $\mu^{\text{Par}}_{i,j,t}$.

Assumption 1. *Parametric mobility frictions contain a time-varying inter-regional component and a constant sector-related component:*

$$\mu^{Par}_{i,j,t} = \bar{\mu}_{r_i,r_j,t}\tilde{\mu}_{i,j}$$

In particular, we specify:

$$\log \bar{\mu}_{r_i,r_j,t} = \delta_{0,t}\mathbb{1}_{r_i \neq r_j} + \delta_{1,t} \log Dist_{r_i,r_j} \tag{21}$$

$$\log \tilde{\mu}_{i,j} = \theta_0 \mathbb{1}_{s_i \neq s_j} + \theta_1 \mathbb{1}_{s_i \neq s_j}\mathbb{1}_{r_i \neq r_j} \tag{22}$$

[19] To do this, insert the definition of $\epsilon_{i,j,t}$ in equation (6). Replace $\hat{\lambda}_{i,j,t} = \exp\{\hat{\omega}_{j,t} + \log\hat{\mu}^{\text{Par}}_{i,j,t} + \hat{\phi}_{i,t}\} = \hat{V}_{j,t}\hat{\mu}^{\text{Par}}_{i,j,t}/\sum_k \hat{V}_{k,t}\hat{\mu}^{\text{Par}}_{i,k,t}$. The estimated (hat) values cancel in the numerator and denominator to return $\lambda_{i,j,t}$. See Fally (2015) for an analysis of PPML in the context of trade flows.

[20] In principle, it is possible to estimate the composite mobility friction, $\mu_{i,j,t}$, in a single step by removing $\mu^{\text{Par}}_{i,j,t}$ from equation (19). One can then project $\mu_{i,j,t}$ on observables. We combine these steps.

Migrants face a fixed cost to leave their state of birth ($\delta_{0,t}$) and a cost proportional to distance ($\delta_{1,t}$). These costs are allowed to vary over time because we have sufficient data to identify them in each twenty-year period from 1880 to 2020 (see Section 4.1.2). These are standard components of migration costs, theorized to include both pecuniary and non-pecuniary (opportunity cost) components (Greenwood, 1997). The time subscript accounts for changing transportation and communication technology over our sample period. Changes in the incentives to migrate, as emphasized by Kaplan and Schulhofer-Wohl (2017), are captured in $V_{j,t}$. Recent studies highlight the potential importance of home ties in migration decisions (Zabek, 2019). The strength of home ties might change endogenously: those with the weakest ties would be the first to leave a declining labor market (Coate and Mangum, 2018). This effect applies to specific labor markets rather than labor market pairs. If workers in a labor market i that has recently lost workers are less mobile as a result of home ties, this will be captured by a higher value of $\epsilon_{i,i,t}$ relative to $\epsilon_{i,j,t}$ for $i \neq j$. This effect is unlikely to bias the estimates of $\mu^{\text{Par}}_{i,j,t}$.

Sector-switchers face a fixed cost (θ_0) that may be different for migrants (θ_1). These costs are assumed constant over time so that we can impose these frictions in the post-1940 data in Section 4.1.2. This specification remains agnostic as to the identity of particular sectors. Sector pair-specific costs largely reflect the movement out of agriculture in the pre-1940 data (see Tables A7 and A8). This secular trend threatens to pollute our focus on migrants' sector-switching advantage. This concern also applies to notions of skill distance between sectors. We leave these extensions for future work.

4.1.2 Gravity Estimation in the Public Use Data

To this point, we have discussed the estimation of fundamentals and mobility frictions in the Restricted data, up to 1940. In principle, information about these objects beyond 1940 can be found in the Public Use data. However, these data report labor flows at the r_i, j, t level, omitting the sector of origin (we write r rather than r_i when possible without confusion). Since we wish to account for sector-related mobility frictions in estimation, this omission presents a challenge.[21]

Our solution to this problem is to disaggregate the r, j, t flows observed in the Public Use data *as if* the sector-switching costs recovered in the Restricted data had applied in the post-1940 period. $\tilde{\mu}_{i,j}$ provides one component of the sector-related switching cost. The other is embedded in $\epsilon_{i,j,t}$. As with the parametric frictions, we will specify the component of the nonparametric frictions that can be inferred only from the Restricted data and obtain the remaining component from the Public Use data.

Assumption 2. *Nonparametric frictions are assumed to have two components:*

$$\epsilon_{i,j,t} = \bar{\epsilon}_{r,j,t}\tilde{\epsilon}_{i,j,t}$$

where

$$\sum_{i:r_i=r} \log \tilde{\epsilon}_{i,j,t} = 0 \tag{23}$$

[21] Estimation by PPML at the r, j, t level is feasible only by ignoring sector-related frictions.

$\bar{\epsilon}_{r,j,t}$ corresponds to the level of detail observed in the Public Use data. $\tilde{\epsilon}_{i,j,t}$ is observed only in the Restricted data, and must be extrapolated beyond 1940. Our first task is to separate $\tilde{\epsilon}_{i,j,t}$ from the values of $\epsilon_{i,j,t}$ recovered from the Restricted data. We will then be able to disaggregate the Public Use data based on sector-related mobility frictions $\tilde{\mu}_{i,j}\tilde{\epsilon}_{i,j,t}$. Finally, we will be able to infer $\bar{\epsilon}_{r,j,t}$ (as well as $\bar{\mu}_{r_i,r_j,t}$) from 1880-2020 using the disaggregated data.

Equation (23) in assumption 2 implies that we can recover $\log \tilde{\epsilon}_{i,j,t}$ as the residual from a regression of $\log \epsilon_{i,j,t}$ on r, j, t dummies. We do this in the restricted data, recovering $\tilde{\epsilon}_{i,j,t}$ for 1880-1940.

The next step is to extrapolate values of $\tilde{\epsilon}_{i,j,t}$ beyond 1940. We assume that $\tilde{\epsilon}_{i,j,t}$ follows an AR(1) process.[22]

Assumption 3. $\log \tilde{\epsilon}_{i,j,t}$ *follows an AR(1) process with normal innovations,*

$$\log \tilde{\epsilon}_{i,j,t} = \tilde{\rho} \log \tilde{\epsilon}_{i,j,t-1} + \eta_{i,j,t} \tag{24}$$

where $\eta_{i,j,t} \sim \mathcal{N}(0, \tilde{\sigma}^2)$

We estimate equation (24) using the residuals from (23) and then extrapolate $\tilde{\epsilon}_{i,j,t}$ forward from its 1940 values. Note that the regressions above use only the non-zero values of $\epsilon_{i,j,t}$; $\log \tilde{\epsilon}_{i,j,t}$ is undefined whenever $\lambda_{i,j,t} = 0$.[23] Therefore, it is necessary to fill in missing values of $\log \tilde{\epsilon}_{i,j,1940}$ from which to extrapolate. We use the empirical distribution of $\log \tilde{\epsilon}_{i,j,1940}$ for $\tilde{\epsilon}_{i,j,t} > 0$ to fill in the missing values, and then extrapolate forward. We have now identified sector-related mobility frictions for all state-sector pairs from 1960-2020.

Disaggregation of Public Use data. Our next task is to disaggregate the Public Use data flows in accordance with the estimated intersectoral mobility frictions. After 1940, we observe $L_{r,j,t}$, and wish to impute $\hat{L}_{i,j,t}$ such that (i) the imputed values sum up to the aggregate flow, $L_{r,j,t} = \sum_{i:r_i=r} \hat{L}_{i,j,t}$; and (ii) outflows from each state-sector labor market add up to total employment observed in the previous period, $L_{i,t-1} = \sum_j \hat{L}_{i,j,t}$.[24] Imputation is achieved by using the model to isolate the relationship between sector-related frictions and the size of $L_{i,j,t}$ flows within r, j, t cells, relative to the flow of sector-stayers, $L_{(r_i,s_j),j,t}$. The imputation equation will satisfy (i) automatically. The scale of $L_{(r_i,s_j),j,t}$ is pinned down by imposing (ii).

We now derive the relationship between sector-related frictions and sectoral outflows. Let $\sigma_{i,j,t} = \tilde{\mu}_{i,j}\tilde{\epsilon}_{i,j,t}$. We isolate $\sigma_{i,j,t}$ in the model-implied choice probability, (6), by two normalizations. First, we divide by the flow of sector-stayers to the same destination to

[22] Our results are robust to alternative assumptions, including assuming that $\tilde{\epsilon}_{i,j,t} = 1$ for all i, j, t. This is not surprising: The R-squared from the regression of $\log \epsilon_{i,j,t}$ on r, j, t dummies is 0.79. $\bar{\epsilon}_{r_i,j,t}$ explains the bulk of the variation in unobservable mobility frictions.

[23] Positive values of $\epsilon_{i,j,t}$ correspond to state-sector pairs with positive flows. $\lambda_{r,j,t} = 0$ requires $\bar{\epsilon}_{r,j,t} = 0$; $\tilde{\epsilon}_{i,j,t}$ cannot be identified as is set to zero. $\lambda_{i,j,t} = 0$ requires $\tilde{\epsilon}_{i,j,t} = 0$ when $\lambda_{r_i,j,t} > 0$.

[24] Note that $L_{i,t-1}$ is observed in the panel of state-sector employment totals.

eliminates the influence of the $V_{j,t}$ and $\bar{\mu}_{r_i,r_j,t}$:

$$\begin{aligned}\frac{L_{i,j,t}}{L_{(r_i,s_j),j,t}} &= \frac{\lambda_{i,j,t}L_{i,t-1}}{\lambda_{(r_i,s_j),j,t}L_{(r_i,s_j),t-1}} \\ &= \frac{V_{j,t}\bar{\mu}_{r_i,r_j,t}\bar{\epsilon}_{r_i,j,t}\sigma_{i,j,t}\big/\Phi_{i,t}}{V_{j,t}\bar{\mu}_{r_i,r_j,t}\bar{\epsilon}_{r_i,j,t}\sigma_{(r_i,s_j),j,t}\big/\Phi_{(r_i,s_j),t}} \\ &= \frac{\sigma_{i,j,t}L_{i,t-1}\big/\Phi_{i,t}}{\sigma_{(r_i,s_j),j,t}L_{(r_i,s_j),t-1}\big/\Phi_{(r_i,s_j),t}}\end{aligned}$$

This normalization eliminates destination-specific factors and the portion of bilateral frictions pertaining to r_i, j, t. A second normalization is needed to eliminate two origin-specific factors: last period's employment $(L_{i,t-1}/L_{(r_i,s_j),t-1})$ and current outside options $(\Phi_{i,t}/\Phi_{(r_i,s_j),t})$.

We repeat the first normalization for the same pair of origins, but now with a different destination. We choose a reference destination, $j'_{i,t} \neq j$, such that $\sigma_{i,j'_i,t} > 0$ and $\sigma_{(r_i,s_j),j'_i,t} > 0$. We obtain,

$$\frac{L_{i,j',t}}{L_{(r_i,s_j),j',t}} = \frac{\lambda_{i,j',t}L_{i,t-1}}{\lambda_{(r_i,s_j),j',t}L_{(r_i,s_j),t-1}} = \frac{\sigma_{i,j',t}L_{i,t-1}\big/\Phi_{i,t}}{\sigma_{(r_i,s_j),j',t}L_{(r_i,s_j),t-1}\big/\Phi_{(r_i,s_j),t}}$$

We now have in hand another ratio of sector-related frictions polluted by the same origin-specific factors as the previous expression.

Taking the ratio of the previous two equations yields,

$$\frac{L_{i,j,t}}{L_{(r_i,s_j),j,t}}\Bigg/\frac{L_{i,j',t}}{L_{(r_i,s_j),j',t}} = \frac{\sigma_{i,j,t}}{\sigma_{(r_i,s_j),j,t}}\Bigg/\frac{\sigma_{i,j',t}}{\sigma_{(r_i,s_j),j',t}} := \tilde{\sigma}_{i,j,t} \tag{25}$$

which we rearrange to obtain,

$$\hat{L}_{i,j,t} = \tilde{\sigma}_{i,j,t}\frac{\hat{L}_{i,j',t}}{\hat{L}_{(r_i,s_j),j',t}}\hat{L}_{(r_i,s_j),j,t} \tag{26}$$

This equation defines flows by sector of origin for sector switchers relative to sector stayers $(L_{(r_i,s_j),j,t})$. When $s_i = s_j$, (26) reduces to an identity. Moreover, equation (26) satisfied automatically satisfies condition (i) mentioned above.[25]

To pin down the flow of sector stayers between r and j, we impose that the outflows from state-sector i at time t add up to total employment in that labor market at $t-1$:

$$\sum_j \hat{L}_{i,j,t} := \hat{L}_{i,t-1} = L_{i,t-1} \tag{27}$$

where $\hat{L}_{i,t-1}$ is the sum of the imputed values and $L_{i,t-1}$ is the quantity observed in the data. We will calculate the imputed values through an algorithm described below; in any given step, the sum of the imputed values may or may not match the data, $L_{i,t-1}$.

Equations (26) and (27) together define a fixed point system for $L_{i,j,t}$ that we use to impute sectors of origin from regional flows. We solve this system as follows:

[25]Equation (26) is just a rearrangement of equation (25), so replacing $\tilde{\sigma}_{i,j,t}$ in equation (26) using its definition in equation (25) returns $\hat{L}_{i,j,t}$. Summing over $\{i : r_i = r\}$ satisfies constraint (i).

1. Initialize $\hat{L}_{i,j,t}$. We guess, $\hat{L}^{(0)}_{i,j,t} = \frac{\sigma_{i,j,t}}{\sum_{k:r_k=r_i}\sigma_{k,j,t}}$.
2. Substitute $\hat{L}^{(0)}_{i,j,t}$ in the left-hand side of equation (26) to obtain $\hat{L}^{(1)}_{i,j,t}$.
3. Use $\hat{L}^{(1)}_{i,j,t}$ to construct $\hat{L}_{i,t-1}$.
4. Construct $\hat{L}_{i,t-1} = \sum_j \hat{L}_{i,j,t}$ using equation (27).
5. Multiply $\hat{L}^{(1)}_{i,j,t}$ by $L_{i,t-1}/\hat{L}_{i,t-1}$.
6. Compute $\sum_{i,j,t}(\hat{L}^{(1)}_{i,j,t} - \hat{L}^{(0)}_{i,j,t})^2$.
7. Repeat from Step 2 until $\hat{L}^{(n+1)}_{i,j,t} \approx \hat{L}^{(n)}_{i,j,t}$.

These steps return a panel of disaggregated flows between state-sector pairs from 1960-2020. We append the Restricted data and repeat the gravity estimation. This returns $\bar{\mu}_{r_i,r_j,t}$ and $V_{j,t}$ for the entire sample period, 1880-2020.

Validation: We validate the disaggregation procedure in two ways. First, we apply the procedure to an artificial dataset in which all flows are positive. This avoids difficulties that arise in the presence of zeros.[26] We regress $\log L_{i,j,t}$ on $\log \hat{L}_{i,j,t}$ along with r, j, t fixed effects. We obtain a within R-squared of 1. When we apply the disaggregation procedure to the pre-1940 data and run the same regression, we obtain a within R-squared of .73, due to the errors inflicted by missing values of $\hat{L}_{i,j,t}$.

Second, we validate the disaggregation procedure using those estimands that are recovered twice, i.e., in Section 4.1.1 as well as Section 4.1.2. In Section 4.1.1, we use the Restricted data to estimate fundamentals and mobility frictions from 1880-1940. In Section 4.1.2, we use the disaggregated Public Use data to estimate these objects from 1880-2020. The two sets of estimates overlap in 1880-1940. In principle, the overlapping estimates should match exactly. We find that the estimated values of $V_{j,t}$ are identical. The correlation across the two estimation steps for the nonparametric components of the mobility frictions, $\bar{\epsilon}_{r_i,j,t}$ and $\tilde{\epsilon}_{i,j,t}$, are .81 and .85, respectively. This makes sense: Due to the adjustments made in Section 2.1, the Restricted and Public Use data contain identical values of $L_{j,t}$, which drive the identification of $V_{j,t}$. Some error arises in $\epsilon_{i,j,t}$ from the imputation of $\hat{L}_{i,j,t}$ in the pre-1940 data, as discussed above. Comparison of parametric mobility frictions, $\bar{\mu}_{r_i,r_j,t}$ and $\tilde{\mu}_{i,j}$, is left to Section 4.5.

4.2 The Fréchet elasticity

Note that in equation (19), the Fréchet elasticity ν is not separately identified from the systematic components of the worker's utility. We set $\nu = 3.5$, a central value relative to estimates in the literature.[27] We adjust the estimates of $V_{j,t}$ and $\mu_{i,j,t}$ in accordance with this parameter value.

[26] The Restricted data are the natural candidate for validation. However, in some cases, a valid reference sector with positive flows cannot be found. Validation in a dataset of positive flows is consistent with our application. We will impute only positive flows in the post-1940 data, since our extrapolation procedure imposes $\tilde{\epsilon}_{i,j,t} > 0$ for $t > 1940$.

[27] The literature offers several estimation methods that we tested in our data. Bryan and Morten (2019) obtain ν by regressing log wages on log choice probabilities, motivated by equation (7). They find $\nu = 2.69$ for the US. Tombe and Zhu (2019) treat the extreme value term as a preference shock. Their estimating

ν appears to be an important parameter because it scales the product $V_{j,t}\mu_{i,j,t}$. These parameters are not separately identified in a migration gravity regression, so it is important to choose a value of ν, but that value does not affect our analysis of the persistence of regional specialization (Section 5.5). Although ν serves to scale $V_{j,t}$ and $\mu_{i,j,t}$, it does not affect the ranking of either set of parameters, nor their relative contribution to dynamics. In particular, it will not affect the estimation of sector-switching costs for migrants relative to non-migrants. Rather, ν determines the scope of comparative advantage by scaling the distribution of fundamentals; as well as the strength of selection (equation (7)). Through these two channels, ν determines the magnitude of changes in output in counterfactual scenarios. We follow our main results with a sensitivity analysis to $\nu = 2$ and $\nu = 6$. Our findings are robust to this value.

4.3 Decomposition of the composite fundamental

Gravity estimation distinguishes labor supply and labor demand, where the latter is summarized by $V_{j,t}$. The next step is to decompose this object using the structure of labor demand. Combining equations (3) and (11), we obtain,

$$V_{j,t} = B_{r,t} P_{s,t} A_{j,t} = B_{r,t} w_{j,t} \tag{28}$$

In this endeavor, we also have at our disposal the supply of efficiency units to each labor market, $S_{j,t}$. This comes from equation (9) and the values of $\lambda_{i,j,t}$ constructed in Section 4.1. First, we distinguish $w_{j,t}$ and $B_{r,t}$. Subsequently, we distinguish $P_{s,t}$ and $A_{j,t}$.

To distinguish $w_{j,t}$ and $B_{r,t}$, we introduce additional data on manufacturing output by state from 1880-1920 and manufacturing wage bill from 1940-2020. According to the model, the wage bill and total output are equal. We observe both in 1940 and confirm that they are highly correlated across states. Data for one sector is sufficient to identify the location-specific amenity. Since the fundamentals are scale-free, we normalize the output and wage bill in each period, and recover,

$$B_{r,t} = \frac{V_{(r,\text{Mfg}),t} S_{r,\text{Mfg},t}}{\Upsilon_{(r,\text{Mfg}),t}}$$

where $\Upsilon_{(r,\text{Mfg}),t}$ represents data on output for $t \in \{1880, 1900, 1920\}$ and the wage bill for $t \in \{1940, 1960, ..., 2020\}$.

This procedure so far does not pin down the scale of $w_{j,t}$. The resulting values of $w_{j,t}$ would imply constant aggregate output. In order to obtain more sensible values, we make the following assumption:

Assumption 4. *Aggregate output—Y_t in equation (13)—grows at a constant annual rate of 2%; and $Y_{1880} = 1$.*

equation regresses log choice probabilities on log wages. They find $\nu = 1.5$ for China. Both methods rely on shift-share IVs—for the choice probability in the first case and for destination wages in the second case. Applying these methods in our data, we estimated $\nu = 0.8$ using the approach of Tombe and Zhu (2019); and $\nu = 5$ using the approach of Bryan and Morten (2019). We chose to calibrate ν for ease of exposition. We report a sensitivity analysis following our counterfactual results.

Assumption 4 implies a sequence of $\{Y_t\}_{2020}^{t=1880}$ that pins down the scale of $w_{j,t}$, which is achieved by a national rescaling of $B_{r,t}$. Given initial values of $B_{r,t}$, we obtain $w_{j,t} = V_{j,t}/B_{r,t}$. Then we construct aggregate output using equation (16). If Y_t is greater (less than) $(1.02)^{t-1880}$, then we scale the amenities up (down) by a small amount. We repeat this process until we match the entire sequence of Y_t.

Finally, we use the final goods producer's FOC, equation (14), to decompose output per efficiency unit $w_{j,t}$ into a sector-specific component $P_{s,t}$ and a local sectoral component $A_{j,t}$. $P_{s,t}$ is identified as the market-clearing price given sectoral output $X_{s,t}$ and aggregate output Y_t. $A_{j,t} = w_{j,t}/P_{s,t}$ thus captures regional variation in sectoral productivity.

4.4 Endogenous and exogenous productivity

Next, we use the calibrated composite productivities $A_{j,t}$ and the structure of intermediate goods production to estimate the investment parameters, α_1 and α_2. To do this, we take logs of equation (12) to obtain,

$$\log A_{j,t} = \alpha_1 \log S_{j,t} + \alpha_2 \log A_{jt-1} + \log \bar{A}_{j,t} \tag{29}$$

This regression suffers from endogeneity. $\log \bar{A}_{j,t}$ enters the worker's choice and is therefore correlated with the supply of skill, $\log S_{j,t}$. We construct an instrument to address this issue.

Following Card (2001) and Ottinger (2020), our instrument uses stocks of foreign immigrants from different origins in 1860 as a shifter of the skill content of a state's workforce. Workers from different origins show comparative advantage in different sectors in 1940, in terms of their national propensities to work in each sector. We treat this as a measure of sector-specific skill. Our instrument sums the stock of sector-specific skill in each state-sector implied by the number of immigrants from each origin in 1860 and the skills of that immigrant group observed in 1940.

Formally, we define immigrants' skills based on revealed comparative advantage (RCA) (Balassa, 1965):

$$S_s^m = \frac{L_{s,1940}^m}{L_{1940}^m} \Big/ \frac{L_{s,1940}}{L_{1940}}$$

where m denotes foreign countries of origin. S_k^m applies the logic of the location quotient to origins and sectors. Our set of origins rounds the IPUMS birthplace codes to the tens place, denoting groups of countries such as Western Europe and East Asia. We restrict to origins assigned to at least 1000 workers in 1920, resulting in 26 unique origins. Appendix Table A14 reports S_s^m averaged for each continent. Workers from different origins specialize in different sectors. We impute immigrants' skills to US states using the stock of migrants from each origin as of 1860:

$$S_{j,1860}^{IV} = \sum_m L_{r_j,1860}^m \times S_{s_j}^m \tag{30}$$

In the first stage regression, we interact our IV with time fixed effects:

$$\log L_{j,t} = \alpha + \sum_{\tau=1900}^{2020} \beta_\tau \log S_{j,1860}^{\text{IV}} \times \mathbb{1}_{t=\tau} + \mu_{j,t} \tag{31}$$

Note that the first stage omits 1880. We do this because for 1880, we initialize with the assumption that $A_{j,1860} \propto A_{j,1880}$. This assumption, though hopefully innocuous for simulation, could be problematic for estimation. In both equations, we assume that $\log \bar{A}_{j,t} = \gamma_{s_j,t} + \log \tilde{A}_{j,t}$. We omit a state-by-year effect to avoid collinearity with the instrument.

This is a shift-share instrument where immigrant stocks comprise the shares and their latent skills comprise the shocks. Exogeneity in just one of these components is sufficient for a valid instrument.[28] We argue both. Exclusion based on shares comes from timing: immigrant stocks in 1860 predate employment outcomes in 1880. Therefore, the presence of immigrants from a certain group affects labor market productivity in 1880, 1900, etc. only through its influence on follow-on migration and the skills of that immigrant group.

We argue exogeneity of immigrants' latent skills also in terms of timing. Latent skills are revealed under 1940 sectoral technology. The idea is that these skills were not relevant given the national sectoral composition in 1860, but became relevant in later decades.[29]

There are two threats to this argument. First is reverse causality: immigrants' skills as of 1940 may reflect skill acquired in the United States, rather than skills brought from their home country. We address this issue with an ad hoc test. We calculate immigrants' sectoral employment shares in 1940 as implied by their states of residence in 1860. We find that the predicted value is uncorrelated with the actual employment counts.[30]

The second concern is that immigrants' sectors may reflect discrimination. While we cannot rule this out entirely, we find that our IV results are robust to defining S_s^m in years other than 1940 (see Table A20). This gives us some comfort, assuming that labor market discrimination has diminished over time (Hsieh et al., 2019).

4.5 Estimation results

We present the results in the order they are obtained. We first present mobility frictions and fundamentals. We then discuss and test the assumptions underlying the Fréchet elasticity. Finally, we present the estimated agglomeration and path dependence elasticities.

Mobility frictions: The parametric mobility frictions capture the cost of switching for migrants and non-migrants and the cost of interstate migration. Since the migration costs are estimated for eight twenty-year periods, the full regression table can be found in Appendix Table A15.

[28]For shares, see Goldsmith-Pinkham et al. (2020). For shocks, see Borusyak et al. (2020).

[29]Ottinger (2020) provides support for this idea. He measures skills in narrow manufacturing industries based on the prominence of origin countries in US imports in 1909. Many of these industries had little or no presence in the labor market at the time that immigrants arrived, and yet their presence predicts subsequent industry growth in their county of residence.

[30]Formally, we predict the number of workers from origin m working in sector s by, $\hat{L}^m_{s,1940} = \sum_{j:s_j=s} L^m_{r_j,1860} \frac{L_{r_j,s_j,1860}}{L_{r_j,1860}}$. We then regress,

$$\log L^m_{s,1940} = \beta_0 + \beta_1 \log \hat{L}^m_{s,1940} + \gamma^m + \delta_s + u^m_s$$

A positive coefficient β_1 would threaten our exclusion restriction. We estimate $\beta_1 = -0.03$ (0.24), which suggests that workers' skills in 1940 do not reflect skills acquired from their residences in the United States.

The estimated cost of switching sectors is lower for migrants than non-migrants: $\theta_0 < 0$ and $\theta_1 > 0$, as shown in Table 3. The table shows two sets of results, corresponding to the two gravity regressions estimated in Section 4.1.1 and 4.1.2. Similar values supports the validity of the disaggregation procedure. We use the latter set of results to quantify the structural model.

Table 3: Estimated switching cost parameters

	θ_0	θ_1
Restricted	-0.93 (0.02)	0.34 (0.03)
Public Use	-0.82 (0.01)	0.36 (0.02)

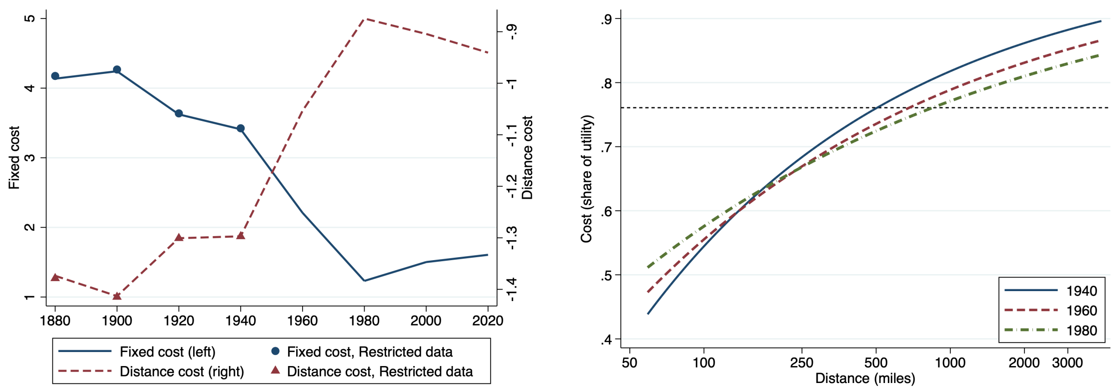
Notes: Estimated coefficients on indicators for switching sector (θ_0) and switching sector while migrating (θ_1). Each row shows a different regression. The first row uses the Restricted data (Section 4.1.1) and the second uses the disaggregated Public Use data (Section 4.1.2). Standard errors in parentheses. See Table A15 for full results.

Figure 2a plots the time-varying geographic mobility frictions. Again, estimates from the Restricted data (points) and the disaggregated Public Use data (lines) are nearly identical. We estimate a positive fixed cost of migration and a negative distance elasticity. Migration is costly on net: the minimum utility cost of migration is 40%.

Geographic mobility frictions exhibit a substantial change between 1940 and 1980. The fixed cost $\delta_{0,t}$ became less positive while the distance elasticity became less negative. This reflects a change in the pattern of migration flows. While the utility cost of migration remained roughly constant, the average distance of migration increased by 40% between these two time periods. Figure 2b plots the cost of migration as a function of distance for 1940, 1960, and 1980. Longer moves demand a larger fraction of utility. Over time, the cost of long moves declined, while short moves became relatively more costly.

Figure 2: Migration frictions and utility cost

(a) Regression estimates of mobility frictions (b) Migration cost vs. distance



Notes: Panel 2a presents estimated migration cost parameters. The fixed cost $\delta_{0,t}$ is plotted against the left axis and the distance elasticity $\delta_{1,t}$ is plotted against the right axis. Points represent the estimates from the Restricted data (Section 4.1.1) and lines represent the estimates from the disaggregated Public Use data (Section 4.1.2). Standard errors not shown; see Table A15 for full results. Figure 2b plots the utility cost of migration as a function of distance for various years, $1 - \exp\{\delta_{0,t} + \delta_{1,t} \log \text{Dist}_{r_i,r_j}\}$.

The magnitude of the estimated frictions is best understood in terms of utility. On average during our sample period, migrants paid 73% of their utility to move. Switchers paid just 21%. Migrants enjoyed slightly smaller switching costs, equal to 15% of their utility, on average.

Earnings of migrants and sector-switchers: Owing to the structure of idiosyncratic shocks, labor mobility is selective, as shown in equation (7).[31] Moves to low-appeal or high-cost labor markets are justified by high idiosyncratic productivity in that destination. We find that both migration and sector-switching are costly and that the cost of switching sectors is lower for migrants. Therefore, the model predicts that migrants and switchers would earn higher wages compared to non-movers from their origin and from their destination; but the size of the gap would be slightly smaller for migrant switchers.

We test this prediction using data on wages observed in the 1940 Census. We run the following regression:

$$\log(\overline{\text{wage}}_{i,j,1940}) = \delta \mathbb{1}_{r_i \neq r_j} + \theta_0 \mathbb{1}_{s_i \neq s_j} + \theta_1 \mathbb{1}_{r_i \neq r_j} \mathbb{1}_{s_i \neq s_j} + \gamma_{j,1940} + \phi_{i,1940} + v_{i,j,1940}$$

where $\overline{\text{wage}}_{i,j,1940}$ is the average labor market income of workers from i observed in j. In the model, we calculate this object using equation (7). The regressors correspond to indicators for migration, sector switching, and their interaction. We compare within origins and destinations by including the corresponding fixed effects.

Table 4 shows the results. By construction, the model predicts $\delta > 0$, $\theta_0 > 0$, and $\theta_1 < 0$, shown in Column (1). The data match the model's prediction in terms of sign. The model

[31]The notion of selective migration is not new (Sjaastad, 1962). Our model and empirical test builds on the logic of Bryan and Morten (2019), extending the analysis of internal migration to a setting with multiple sectors.

overstates the strength of selection for interstate migrants, but the other two coefficients are similar in value. Overall, this test supports the assumption that mobility is motivated by both state- and sector-specific skill.

Table 4: Selective mobility

	(1) Model	(2) Data
Migrate	1.536***	0.273***
	(0.00826)	(0.0205)
Switch	0.301***	0.143***
	(0.00146)	(0.0236)
Migrate & switch	-0.145***	-0.106***
	(0.00152)	(0.0173)
N	101723	11780
R2	0.900	0.458

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: Regression of log average wage by i, j cell on indicators for migration, sector-switching, and their interaction. Standard errors are clustered two ways by i and j and are shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

On the other hand, differences in the magnitudes of the estimated coefficients across the two columns may indicate misspecification. This may have implications for the counterfactual analysis of GDP, discussed in Section 5.4. We assume that workers draw skills for each state-sector pair independently. A richer model would allow for separate draws of state- and sector-specific skill with a nested or generalized correlation structure (Lind and Ramondo, 2021). Table (4) provides moments to calibrate such a model. We highlight this as a direction for future research.

We also assume that ν is constant across demographic groups and over time. We test this assumption in Appendix **??**. There is some variation in ν, but not enough to materially change our results.

Fundamentals: Local sectoral fundamentals allow the model to exactly match observed local sectoral employment outcomes. The labor demand model defines wages in each state-sector as a component of the composite fundamental. We validate the decomposition by testing the correlation between model-implied wages and their counterpart in the data, available from 1940 onward. In particular, we run the following regression:

$$\log w_{j,t}^{\text{model}} + \alpha + \sum_{k} \mathbb{1}_{s=s_j} \beta_s \log w_{j,t}^{\text{data}} + \delta_{r_j,t} + \gamma_{s_j,t} + u_{j,t}$$

Table 5 shows the results with varying sets of fixed effects. The choice of fixed effects is important to the results. The unconditional correlation, shown in column (1), is positive and statistically significant. Column (2) adds state-year fixed effects. These coefficients represent the correlation between observed and model-implied wages across sectors within each state. When we add sector-year fixed effects in column (3), the correlation breaks down. Results are even weaker with two-way fixed effect in column (4).

Taken together, these results are encouraging. Differences across states are smaller than differences across sectors. Column (2) produces strong results comparing sectors within states. Columns (3) and (4) have less variation to work with. This aligns with the observation that geographic differences in sectoral specialization are small.

Table 5: Correlation of log wages in the model and the data

	(1)	(2)	(3)	(4)
Agriculture	0.916***	0.344***	0.377***	0.288***
	(0.0265)	(0.0545)	(0.0332)	(0.0413)
Manufacturing	1.096***	0.542***	0.236***	0.197***
	(0.0259)	(0.0529)	(0.0359)	(0.0456)
Services	1.040***	0.484***	0.166***	0.140***
	(0.0262)	(0.0531)	(0.0369)	(0.0467)
Other	1.057***	0.490***	0.135***	0.0821
	(0.0266)	(0.0541)	(0.0399)	(0.0510)
N	740	740	740	740
F	503.6	187.2	51.03	16.62
R2	0.733	0.960	0.960	0.984
Within R2		0.573	0.218	0.104
FE	None	State-year	Sector-year	State- and sector-year

Notes: Coefficients from a regression of log wages observed in the data on wages implies by the model $X_{j,t}/S_{j,t}$ interacted with sector dummies, and controlling for fixed effects indicated in the bottom row. Standard errors in parentheses * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Next, we use the decomposed fundamentals to illustrate the force driving sectoral reallocation in our model. Since sectoral intermediates are gross substitutes in the production of final goods, an expansion in a sector's employment share requires an improvement in productivity relative to other sectors, and relative to the aggregate annual growth rate of 2%. Growth exceeding 2% lowers the price of the sectoral intermediate and increases demand for inputs from that sector. Table 6 illustrates this mechanism in the long run. For each sector, the first column reports the change in employment share between 1880 and 2020, and the second column reports the annualized growth rate in the average value of $A_{j,t}$. Agriculture experienced the deepest loss in employment and has the lowest growth rate. The service sector lies at the other end of the spectrum on both accounts.

Table 6: Productivity Growth and Employment Composition

	Employment growth (p.p.)	Productivity growth
Agriculture	-50	0.82
Manufacturing	-3	1.74
Services	36	2.21
Other	17	2.03

Notes: Employment growth is the change in the sector's national employment from 1880-2020. Productivity growth is the annualized rate of output growth for that sector from 1880-2020. The model is calibrated so that GDP grows at a 2% annual rate, so we expect that productivity growth in expanding sectors is not only higher than in other sectors but also exceeds 2%.

Agglomeration and persistence: IV estimates of α_1 and α_2 are presented in Table 7. Our results imply $\alpha_1 = 0.07$ and $\alpha_2 = 0.75$. Our estimate of α_1 lies on the upper range of estimates in the literature reviewed by Combes and Gobillon (2015). The estimated value of α_2 is equivalent to an annual depreciation rate of 1.4%. Practically speaking, a doubling of efficiency units in a labor market increases productivity by 8% on impact and by 6% after twenty years.

First stage estimates can be found in Appendix Table A19. Table A20 shows 2SLS estimates when S_s^m is defined in different years.

Table 7: Persistence: 2SLS

	(1) OLS	(2) IV
Efficiency Units	0.103***	0.0709**
	(0.0169)	(0.0264)
Lag Productivity	0.463***	0.746***
	(0.0889)	(0.0613)
N	1184	1184
CDW F		27.53

Notes: 2SLS estimates of composite productivity on efficiency units and lag productivity (equation (29)). Log efficiency units is instruments with log embodied skill (equation (30)). Controls include log $A_{j,t-1}$ and sector-year fixed effects. Standard errors are clustered by sector-year and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

4.6 Equilibrium solution method

In counterfactual analysis, we compare results with and without endogenous investment. With exogenous productivity (no investment), an equilibrium is a set of prices $\{P_{s,t}\}$ and employment outcomes $\{L_{j,t}\}$. To solve for an equilibrium, given parameters, fundamentals, and initial conditions, we guess a vector of prices, which implies employment outcomes through the workers' discrete choice, equation (6), and then iterate on prices to solve the final goods producer's FOC, equation (14).

With endogenous investment, the skill price $w_{j,t}$ is also an equilibrium object. Solving for equilibrium now requires an additional loop. As before, we can solve for employment outcomes given prices $\{P_{s,t}\}$ and $\{w_{j,t}\}$. The inner loop of our procedure takes $\{P_{s,t}\}$ as given and solves for $w_{j,t}$ to satisfy the intermediate goods producers' FOCs and clear the market for efficiency units. A fixed point system for $w_{j,t}$ comes from substituting equation (36) into equation (34), and then replacing $S_{j,t}$ with its equilibrium value. The fixed point equation is homogeneous of degree one, so the equation has a unique fixed point regardless of parameter values.

5 Counterfactual exercises

Counterfactual analysis aims to address three questions. (1) How influential is history for present-day specialization? (2) What accounts for persistence in specialization? (3) What are the implications of specialization and persistence for aggregate output?

5.1 The origins of regional specialization

To explain the origins of regional specialization in terms of our model, we study the impact of historical shocks, $\bar{A}_{j,t}$, on present-day specialization. In particular, we randomize $\bar{A}_{j,t}$ across states within each sector for a particular base year, and then run the model forward. Randomization is applied only in the base year of a given simulation. After that, $\bar{A}_{j,t}$'s take on their estimated values. We perform this exercise for each year of our sample up to 2000. We calculate location quotients in the simulated data and compute the correlation of these values to the actual location quotients in 2020. Table 8 shows the results. A lower value of the correlation means that the particular historical shocks are more influential to present-day specialization. By this metric, all decades of our sample bear on the present allocation. 2000 is the most influential. But the runners-up lie in the earliest years, 1880 and 1900. These decades appear to be modestly more influential than the intervening years from 1920 to 1960. This might correspond to substantial changes in sectoral specialization that took place during the Second Industrial Revolution.

Table 8: Correlation of present-day specialization to counterfactual values with randomized historical shocks

$t_0 =$	1880	1900	1920	1940	1960	1980	2000
$\text{Corr}(L_{j,2020}^{\text{Sim},t_0}, L_{j,2020})$	0.08	0.07	0.10	0.12	0.15	0.08	0.06

5.2 Decomposing persistence

We use the quantified model to study the role of labor mobility frictions and endogenous investment in the persistence of regional specialization. Absent these forces, exogenous geography and historical accidents, $\bar{A}_{j,t}$, define the benchmark level of persistence in the economy. As a first look persistence, consider the persistence of location quotients, composite productivity $A_{j,t}$, and fundamental productivity $\bar{A}_{j,t}$. These values are 0.76, 0.45, and -0.08, respectively. This already suggests that composite productivity explains a large portion of persistence, coming from endogenous path dependence rather than fundamental productivity.

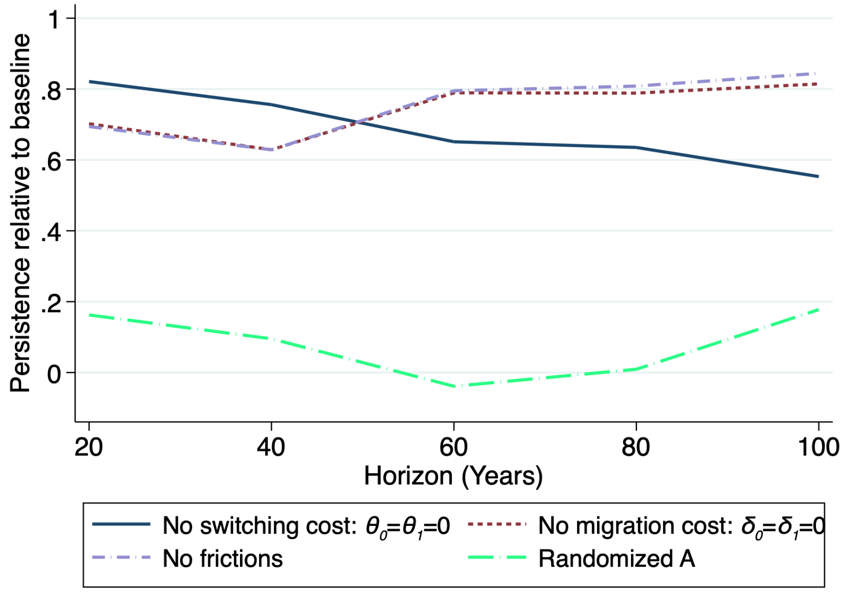
We use the model to decompose the contribution of mobility frictions (of various types) and path dependence to the persistence of specialization. We report results from four counterfactuals. Three of them modify mobility frictions: without migration costs, without switching costs, and without any costs. The fourth randomizes composite productivity in each year of the sample, holding investments fixed at their baseline values. That is, we randomize $A_{j,t}$ across states within each sector in every year from 1880 to 2020 as we simulate the economy forward from 1860.[32] For each exercise, we calculate the rank-rank elasticity of the location quotient for various time horizons. We take the ratio of these coefficients to their baseline values. The results are reported in Figure 3.

The three lines at the top of the figure represent the mobility cost exercises. Removing all mobility frictions (dashed lavender), persistence falls by 32% on impact, and attenuates somewhat at longer horizons. Removing switching frictions (dotted maroon) results in a sustained drop in persistence, equal to 19% on impact. This reflects the finding of Eckert and Peters (2018) that changes in sectoral composition in each state are orthogonal to changes in each state's share of national employment. Without switching costs, workers reallocate within each state according to its natural advantage. This is not the only component of the location quotient—total employment in each state enters as well—but local sectoral composition drives a steady share of persistence. Finally, although the cases of free migration and free switching have different effects on persistence, the frictionless economy coincides with the free-migration economy. This is because the cost of migration is much larger than the cost of switching sectors. Overall, these results indicate that mobility frictions play a substantial role in the persistence of regional specialization in the US.

The remainder of persistence comes from endogenous path dependence. The dashed mint line shows our fourth counterfactual with randomized $A_{j,t}$ and baseline mobility frictions. Persistence is almost completely eliminated.

[32] We do not perform a similar exercise with $\bar{A}_{j,t}$ because $\bar{A}_{j,t}$ exhibits inverse autocorrelation. Randomizing it actually increases its persistence.

Figure 3: Rank-rank elasticity of location quotients relative to baseline estimates



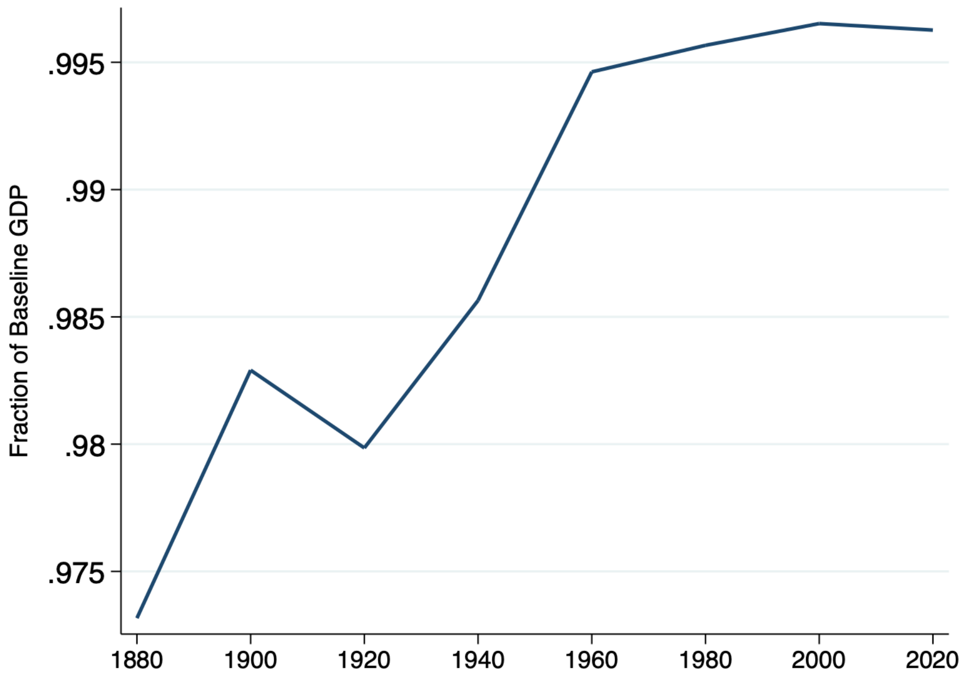
Notes: Each line represents the rank-rank elasticity of location quotient (LQ) within each sector, relative to the baseline estimate. Elasticities are calculated as follows. We compute LQ for each state-sector, rank these within each sector, and regress log rank on its lag at various horizons, controlling for state-year and sector-year dummies. The estimate for each horizon averages across the years in which that horizon is observed.

One might wonder if we overstate the role of mobility frictions in persistence by ascribing the mobility of young workers to the entire population. If young workers are the most mobile, then our estimates constitute an upper bound on. Appendix presents an alternate model where workers live for two periods, and are immobile in their old age. In this case, mobility frictions reduce persistence by 21%. In reality, older workers are quite mobile (Table A3), so this alternative estimate constitutes a lower bound for the role of mobility frictions in persistence.

5.3 Specialization and aggregate output

What is the value of specialization? Overall, we find that productivity differences across states are small. To illustrate this point, we flatten specialization across states. That is, we set $S_{j,t}$ such that all states share the national sectoral composition. $A_{j,t}$ is unchanged in this disequilibrium exercise. Figure 4 reports GDP in this counterfactual relative to the baseline. Specialization is not particularly valuable to the US economy. GDP falls by 3% in 1880; the loss shrinks to less than 1% after 1960. This is in line with declining specialization, reported in Crafts and Klein (2021). There is limited scope for aggregate misallocation. Keep this in mind as we proceed to simulate the effects of counterfactual mobility frictions.

Figure 4: GDP change without specialization



Notes: GDP relative to baseline from a shift share exercise in which efficiency units in each state are reallocated across sectors so that each state shares the national sectoral composition.

5.4 Mobility frictions and aggregate output

Does the persistence of state sectoral specialization matter for the aggregate economy? Persistence arises from two sources, agglomeration/path dependence and mobility frictions. Each has conceptually different implications for the aggregate economy. The former maintains the size of regional sectoral clusters, generating aggregate misallocation. The latter maintains size but also prevents idiosyncratic reallocation. We saw in the previous subsection that there is limited scope for aggregate misallocation. Here, we focus on the effect of mobility frictions on aggregate output.

Our model provides substantial flexibility for this analysis. We can separately manipulate migration costs, switching costs, and switching costs for migrants. In each case, we consider two counterfactuals, zero frictions and prohibitive frictions, relative to the baseline. For each counterfactual, we apply the modified mobility frictions and run the model forward from 1860. We then calculate the ratio of GDP to the baseline in each time period and take the average. Table 9 shows the results. Each row reports counterfactual migration costs; each column reports counterfactual switching costs.

Removing all frictions from the economy increases GDP by 31%. The huge increase in GDP comes from better matching of workers to labor markets. Applying the same exercise with composite productivity held fixed at baseline estimates (i.e., no agglomeration from concentration of workers), GDP still rises by 26%.

The bulk of the gain comes from removing migration costs. This is not surprising, as migration costs are much larger than switching costs, in utility terms. However, sector switching is still important. One way to see this is to consider prohibitive frictions instead of zero frictions. The US is already highly mobile by international standards, so by comparing an economy with prohibitive mobility costs to the baseline, we put a value on realized labor mobility. Intersectoral labor mobility is slightly more valuable than interregional mobility.[33] More to the point, the interaction between migration and sector-switching is important. Prohibiting switching for migrants ($\theta_1 = -\infty$) reduces GDP by 30% on average, compared to 45% from prohibiting migration. Our interpretation is that two-thirds of the benefit of realized migration comes from migrants' ability to switch sectors.

This is not to discount the potential benefit of reducing migration costs. Under $\theta_1 = -\infty$, removing migration costs increases GDP by 20%—more than two-thirds the gain achieved under baseline switching costs—though GDP remains below the baseline value.

Table 9: GDP in counterfactuals

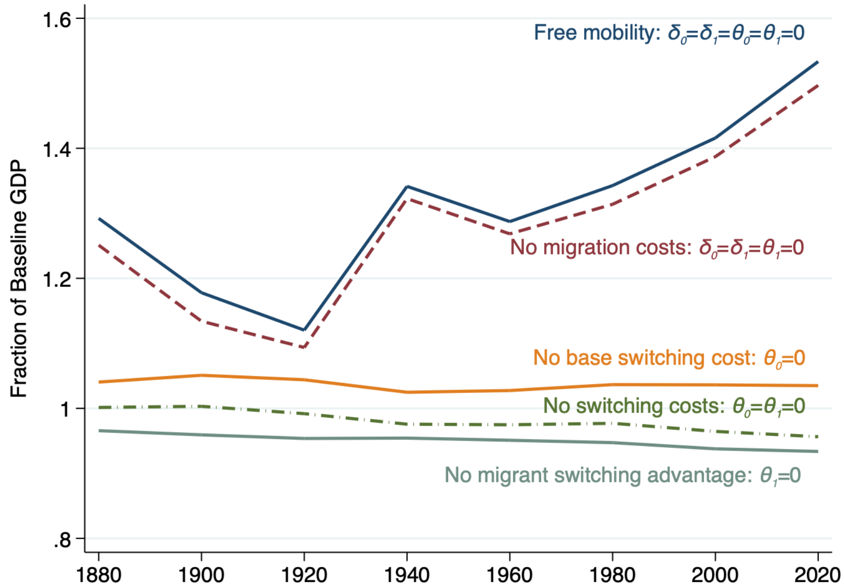
		μ^S			
		Baseline	None	Prohibitive	Prohibitive for migrants
	Baseline	100	98	46	70
μ^R	None	129	131	63	84
	Prohibitive	55	57	30	

Notes: Each cell corresponds to a counterfactual with mobility frictions set according to the row and column labels. In each case, we run the model forward from 1860; $\bar{A}_{j,t}$ and $B_{r,t}$ take on their estimated values. $A_{j,t}$ and $P_{s,t}$ will change in counterfactual equilibria. We take the ratio of GDP in the counterfactual to GDP in the baseline and multiply it by 100. We take the average of this ratio across all years of the simulation (1880-2020) and report this number in the table.

The value of removing mobility frictions changes over time. Figure 5 plots the ratio of counterfactual to the baseline for the scenarios described above, where mobility frictions are modified from 1860 onward. The gains from free migration are smaller in 1900 and 1920 compared to other years. The reasons for this are complex, reflecting the interaction of the preexisting employment allocation, shocks to fundamental productivity, the persistence of past productivity, and the divergent influence of amenities and productivity. The latter is easiest to quantify. We calculate employment-weighted average productivity for the baseline and frictionless economies and take the ratio of the latter to the former. We do the same for amenities. Figure 6 reports the results. Workers consistently enjoy better amenities and inferior productivity in the frictionless economy. In 1900 and 1920, the average value of amenities is 50% higher than the baseline.

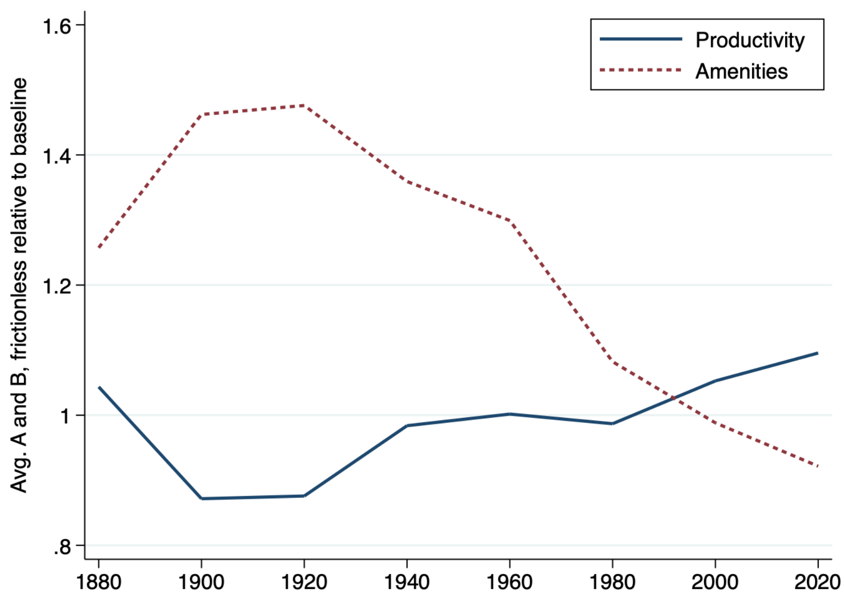
[33] Appendix Figure 8 shows trends for this comparison. The value of sectoral reallocation grows over time as the distance grows between the 1860 sectoral allocation and the allocation implied by exogenous productivity growth.

Figure 5: GDP without mobility frictions



Notes: Each line depicts GDP (relative to the baseline) in a counterfactual economy with modified mobility frictions simulated forward from 1860. $\bar{A}_{j,t}$ and $B_{r,t}$ take on their estimated values. $A_{j,t}$ and $P_{s,t}$ will change in counterfactual equilibria.

Figure 6: Productivity and amenities: Frictionless relative to baseline



Notes: Employment-weighted average productivity and amenities in the frictionless economy relative to the baseline. $\bar{A}_{j,t}$ and $B_{r,t}$ take on their estimated values. $A_{j,t}$ and $P_{s,t}$ will change in counterfactual equilibria.

5.5 Robustness

Much of the spatial economics literature relies on extreme value random utility models. This literature devotes much attention to the estimation of the shape parameter of this distribution. We chose to calibrate the shape parameter of our Fréchet distribution ν to match the variance of wages, in the style of Bryan and Morten (2019) and Hsieh et al. (2019). Here, we evaluate the robustness of our results to this choice.

Table 10 shows key model outcomes over a range of values for ν. Results related to persistence are unaffected. This is because the Fréchet elasticity simply scales the product of labor market fundamentals and mobility frictions—which embed the sources of persistence—without changing their size relative to one other. The effect of mobility frictions on GDP is sensitive to the Fréchet parameter, mostly due to its role in selection. Increasing ν implies less dispersion in skill draws, hence less scope for comparative advantage, less misallocation, and smaller gains from removing mobility frictions.

Our results suggest that there is substantial misallocation of workers across states and sectors in terms of idiosyncratic skill. However, our model makes no distinction between state- and sector-specific skills. Therefore, we cannot say whether geographic or sectoral misallocation is more severe. The analysis of GDP should be viewed with caution pending a richer analysis of the composition of and correlation between state- and sector-specific skills.

Table 10: Robustness

	$\nu = 2$	3.5	6
Persistence without frictions	0.71	0.69	0.68
GDP without frictions	1.38	1.32	1.24

Notes: Counterfactual results for different values of the Fréchet elasticity ν. The first row reports twenty-year persistence in the frictionless counterfactual ($\mu_{i,j}^{\mathrm{Par}} = 1$ for all i, j) relative to the baseline the case. The second row reports GDP in the frictionless counterfactual relative to the baseline.

6 Conclusion

Regional specialization suggests differences in comparative advantage. But specialization also reflects the persistent influence of the past. In this paper, we distinguish two sources of endogenous persistence: labor mobility frictions that prevent workers from moving to more productive or higher amenity places; and endogenous investment arising modeled as an agglomeration externality. Using a dynamic spatial model, we show that mobility frictions explain one third of persistence and investment explains the rest. Further, we explore the implications of removing mobility frictions for aggregate output. Allowing workers to reallocate freely across space has a dramatic positive effect on GDP. However, this arises due to improved allocative efficiency in idiosyncratic skills; specialization itself has little value at the level of aggregation studied here. Our findings about the structure of workers' skills is a further contribution of this analysis.

Our work sheds light on the importance of migration to the aggregate economy. The costs of migration are much greater than the cost of switching sectors. At the same time, net migration flows are much smaller than gross flows. Migration is an important margin on which workers improve the match of their location- and sector-specific skills. This view is concordant with the finding of Eckert and Peters (2018) that net migration had little role to play in aggregate sectoral transformation. Nonetheless, migration is important to aggregate outcomes.

Our analysis of mobility across states and sectors together suggests avenues for future research related to misallocation. It appears that interstate migration could facilitate better matching of workers to sectors. To facilitate our analysis of persistence, we assume that workers' skills are drawn independently for each state-sector pair. Future work should delve into the structure of workers' skills: to what extent are skills state-specific or sector-specific? And to what extent are movements along each margin motivated by skill as opposed to preference? How are these shocks correlated across states and sectors? Panel data with wages are likely needed for this analysis. The answer has important implications for the aggregate implications of mobility frictions across states and sectors.

We provide novel estimates of mobility frictions across states and sectors. First, we observe that the cost of switching sectors is lower for interstate migrants. This could have implications for the incidence of place-based policies; but an answer to this question requires a more thorough understanding of the content of state- and sector-specific skills. Second, we observe that the structure of migration costs changed around 1940-1980. Future research should study the causes and consequences of this change.

The elasticity of labor supply in response to a local sectoral shock may change over time due to endogenous forces. The initial movement of workers from one location to another changes the destination (through agglomeration, in our model) and the composition of would-be migrants at the origin (Coate and Mangum, 2018, Zabek, 2019). A larger stock of migrants from a particular origin living in a destination may attract follow-on migration (Carrington et al., 1996). These features would make for interesting extensions to our analysis.

The interaction of geographic and sectoral switching costs may have implications for the analysis of local industrial policies and labor demand shocks. The incidence of place-based policies depends in part on geographic mobility (Busso et al., 2013, Notowidigdo, 2020). Likewise, the incidence of place-based policies targeting a particular industry depends on sectoral mobility as well. Our results suggest that migrants' sector-switching advantage may dilute the benefits of place-based industrial policies for local residents. The interaction of migration and sector-switching costs could also have implications for the welfare effects of trade shocks (Caliendo et al., 2019).

References

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2012). Europe's tired, poor, huddled masses: Self-selection and economic outcomes in the age of mass migration. *American Economic Review*, 102(5):1832–1856.

Adão, R., Beraja, M., and Pandalai-Nayar, N. (2020). Technological transitions with skill heterogeneity across generations.

Allen, T. and Donaldson, D. (2020). Persistence and path dependence in the spatial economy. *NBER Working Paper*, (28059).

Artu c, E., Chaudhuri, S., and McLaren, J. (2010). Trade shock and labor adjustment: A structural empirical approach. *American Economic Review*, 100(3):1008–1045.

Balassa, B. (1965). Trade liberalisation and "revealed" comparative advantage. *The Manchester School*, 33(2):99–123.

Bayer, P., McMillan, R., Murphy, A., and Timmins, C. (2016). A dynamic model of demand for houses and neighborhoods. *Econometrica*, 84(3):893–942.

Berkes, E., Gaetani, R., and Mestieri, M. (2020). Cities and technological waves. *Working paper*.

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *Quarterly Journal of Economics*, 127(2):587–644.

Borusyak, K., Hull, P., and Jaravel, X. (2020). Quasi-experimental shift-share research designs. *Review of Economic Studies*, Forthcoming.

Bryan, G. and Morten, M. (2019). The aggregate productivity effects of internal migration: Evidence from indonesia. *Journal of Political Economy*, 127(5):2229–2268.

Busso, M., Gregory, J., and Kline, P. (2013). Assessing the incidence and efficiency of a prominent place based policy. *American Economic Review*, 103(2):897–947.

Caliendo, L., Dvorkin, M., and Parro, F. (2019). Trade and labor market dynamics: general equilibrium analysis of the china trade shock. *Econometrica*, 87(3):741–835.

Caliendo, L., Parro, F., Rossi-Hansberg, E., and Sarte, P.-D. (2018). The impact of regional and sectoral productivity changes on the u.s. economy. *Review of Economic Studies*, 85:2042–2096.

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

Carrington, W. J., Detragiache, E., and Vishwanath, T. (1996). Migration with endogenous moving costs. *American Economic Review*, 86(4):909–930.

Coate, P. and Mangum, K. (2018). Fast locations and slowing labor mobility. *Working Paper*.

Combes, P.-P. and Gobillon, L. (2015). The empirics of agglomeration economies. In Henderson, J. V., Nijkamp, P., Mills, E. S., Cheshire, P. C., and Thisse, J. F., editors, *Handbook of Regional and Urban Economics*, chapter 5, pages 247–348. New York.

Crafts, N. and Klein, A. (2021). Spatial concentration of manufacturing industries in the united states: re-examination of long-run trends. *European Review of Economic History.*

Davis, D. R. and Weinstein, D. E. (2002). Bones, bombs, and break points: The geography of economic activity. *American Economic Review*, 92(5):1269–1289.

Desmet, K., Nagy, D. K., and Rossi-Hansberg, E. (2018). Spatial development. *Journal of Political Economy*, 126(3):903–983.

Desmet, K. and Rossi-Hansberg, E. (2014). Spatial development. *American Economic Review*, 104(4):1211–1243.

Dix-Carneiro, R. (2014). Trade liberalization and labor market dynamics. *Econometrica*, 82(3):825–885.

Duranton, G. and Puga, D. (2005). From sectoral to functional urban specialization. *Journal of Urban Economics*, 57(2):343–370.

Eckert, F. and Peters, M. (2018). Spatial structural change. *Working Paper.*

Ellison, G. and Glaeser, E. L. (1997). Geographic concentration in u.s. manufacturing industries: A dartboard approach. *Journal of Political Economy*, 105(5):889–927.

Ellison, G., Glaeser, E. L., and Kerr, W. R. (2010). What causes industry agglomeration? evidence from coagglomeration patterns. *American Economic Review*, 100(3):1195–1213.

Fally, T. (2015). Structural gravity and fixed effects. *Journal of International Economics*, 97(1):76–85.

Gervais, A., Markusen, J. R., and Venables, A. J. (2020). Inter-city specialization and trade in functions versus sectors. *Working paper.*

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624.

Gollin, D., Lagakos, D., and Waugh, M. (2014). The agricultural productivity gap in developing countries. *Quarterly Journal of Economics*, 129(2):939–993.

Greenstone, M., Hornbeck, R., and Moretti, E. (2010). Identifying agglomeration spillovers: Evidence from winners and losers of large plant openings. *Journal of Political Economy*, 118(3):536–598.

Greenwood, M. J. (1997). Internal migration in developed countries. In Rosenzweig, M. R. and Stark, O., editors, *Handbook of Population and Family Economics*, chapter 12, pages 647–720. North Holland.

Guimaraes, P., Figueirdo, O., and Woodward, D. (2003). A tractable approach to the firm location decision problem. *Review of Economics and Statistics*, 85(1):201–204.

Hanlon, W. W. and Miscio, A. (2017). Agglomeration: A long-run panel data approach. *Journal of Urban Economics*, 99:1–14.

Helm, I. (2019). National industry trade shocks, local labour markets, and agglomeration spillovers. *Review of Economic Studies*, 87(3):13991431.

Herrendorf, B., Rogerson, R., and Valentinyi, A. (2013). Two perspectives on preferences and structural transformation. *American Economic Review*, 103(7):2752–2789.

Hsieh, C.-T., Hurst, E., Jones, C. I., and Klenow, P. J. (2019). The allocation of talent and us economic growth. *Econometrica*, 87(5):1439–1474.

Kaplan, G. and Schulhofer-Wohl, S. (2017). Understanding the long-run decline in interstate migration. *International Economic Review*, 58(1):57–94.

Kennan, J. and Walker, J. R. (2011). The effect of expected income on individual migration decisions. *Econometrica*, 79(1):211–251.

Kim, S. (1995). Expansion of markets and the geographic distribution of economic activities: The trends in u. s. regional manufacturing structure, 1860-1987. *Quarterly Journal of Economics*, 110(4):881–908.

Kline, P. and Moretti, E. (2014). Local economic development, agglomeration economies, and the big push: 100 years of evidence from the tennessee valley authority. *Quarterly Journal of Economics*, 129(1):275–331.

Krugman, P. (1991). Increasing returns and economic geography. *Journal of Political Economy*, 99(3):483–499.

Lin, J. and Rauch, F. (2020). What future for history dependence in spatial economics? *Regional Science and Urban Economics, forthcoming.*

Lind, N. and Ramondo, N. (2021). Trade with correlation. *Working Paper.*

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In Zarmbka, P., editor, *Frontiers in Econometrics*, chapter 4, pages 105–142. New York: Academic Press, New York.

Michaels, G., Rauch, F., and Redding, S. J. (2018). Task specialization in us cities from 1880 to 2000. *Journal of the European Economic Association*, 17(3):754–798.

Notowidigdo, M. J. (2020). The incidence of local labor demand shocks. *Journal of Labor Economics*, 38(3):687–725.

Oliveira, J. and Pereda, P. (2020). The impact of climate change on internal migration in brazil. *Journal of Environmental Economics and Management*, 103.

Ottinger, S. (2020). Immigrants, industries, and path dependence. *Working Paper.*

Pellegrina, H. S. and Sotelo, S. (2021). Migration, specialization, and trade: Evidence from the brazilian march to the west. *Working Paper*.

Redding, S. J. and Rossi-Hansberg, E. (2017). Quantitative spatial economics. *Annual Review of Economics*, 9:21–58.

Rossi-Hansberg, E., Sarte, P.-D., and Schwartzman, F. (2019). Cognitive hubs and spatial redistribution. *Working paper*.

Ruggles, S., Fitch, C. A., Goeken, R., Hacker, D. J., Nelson, M. A., Roberts, E., Schouweiler, M., and Sobek, M. (2021a). *IPUMS Ancestry Full Count Data: Version 3.0 [dataset]*. IPUMS, Minneapolis, MN.

Ruggles, S., Flood, S., Foster, S., Goeken, R., Pacas, J., Schouweiler, M., and Sobek, M. (2021b). *IPUMS USA: Version 11.0 [dataset]*. IPUMS, Minneapolis, MN.

Silva, J. S. and Tenreyro, S. (2006). The log of gravity. *Review of Economics and Statistics*, 88(4):641–658.

Simon, C. J. and Nardinelli, C. (2002). Human capital and the rise of american cities, 1900-1990. *Regional Science and Urban Economics*, 32:59–96.

Sjaastad, L. A. (1962). The costs and returns of human migration. *Journal of Political Economy*, 70(5):80–93.

Tombe, T. and Zhu, X. (2019). Trade, migration, and productivity: A quantitative analysis of china. *American Economic Review*, 109(5):1843–1872.

Young, A. (2013). Inequality, the urban-rural gap, and migration. *Quarterly Journal of Economics*, 128(4):1727–1785.

Zabek, M. (2019). Local ties in spatial equilibrium. *FEDS Working Paper*.

Appendix A Data Details

A.1 Restricted Complete Count Data

Our *Restricted data* come from the Confidential Complete Count Historical Census.[34] These data, available up to 1940 at time of writing, report full names and other characteristics for all individuals in the United States. We link individuals across census years by name and demographic traits following Abramitzky et al. (2012). Among the linked individuals, we restrict our attention to native-born white males observed in the labor force at some point

[34]Minnesota Population Center and Ancestry.com. IPUMS Restricted Complete Count Data: Version 1.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2013.

in the panel.[35] Childhood observations are used to construct intergenerational linkages. We link workers and children across each even cross-section from 1860 to 1940.[36]

The resulting panel allows us to track individuals across states and sectors and estimate labor mobility frictions. In our analysis, we study migration and sector-switching of young workers relative to their fathers; that is, workers who are between twenty and forty years old and were enumerated as a child of the household head two decades prior.

We focus on young workers in order to maintain a consistent definition of sector-switching. Young workers move relative to their place of birth and father's sector. For older workers, mobility might be more appropriately defined relative to their own prior location and sector. In the raw data, these groups demonstrate similar behavior in terms of migration and sector switching, as shown in Appendix Tables A4, A5, and A6. Focusing on young workers also facilitates comparison to the Public Use data, where we observe mobility at twenty-year frequency only relative to the state of birth. See Section A.2 for further discussion with regards to the Public Use data.

A.2 Public Use Data

Our *Public Use data* consists of Public Use Microdata Sample cross-sections for every even decade, 1860-2020. 1860-2000 data come from the decennial census; for the 2020 period, we use the 2014-18 ACS. We restrict our sample to workers with non-missing occupation and industry codes, per the IPUMS 1950 categorizations. We compute total employment by state of birth, state of residence, and sector.

Sectors are constructed as follows. First, we define *industries* based on the first digit of the IPUMS IND1950 variable, resulting in nine industries. We aggregate these into four *sectors*. The sectors and their component industries are listed in Table A2. In terms of geography, we focus on US states, the level at which workers' birthplace is reported throughout the Census data.

Throughout the paper, we analyze lifetime migration, indicated by a worker's residing outside their state of birth. We focus on workers age 20-39 and treat one period of the model as twenty years. One might be concerned that this is too short a time frame. Workers might return to their state of birth as they age. Table A3 mitigates this concern: if anything, lifetime migration rates increase with age. Older workers continue to move away from their state of birth as they age, albeit at a lower rate. In the model, we ascribe the mobility behavior of the young to all workers. In an extension, we consider a model where workers have two periods of working life and are immobile in the second period.

[35]We restrict our sample due to difficulties in matching members of the excluded groups. Names of foreign-born individuals are often misspelled or Americanized over time. Black Americans have fewer unique names as a consequence of slave naming conventions in the antebellum South. Women typically changed their name at marriage and comprised only a small fraction of the labor force during our sample period of 1860-1940.

[36]The bidecadal frequency accords with our model of overlapping generations and also skips the Census of 1890, which was lost in a fire.

A.3 Harmonization

The datasets described above provide two panels of flows. It is important to ensure consistency between the two datasets. In principle, both samples are derived from the same population. However, we note that the Restricted data understate the rate of gross migration relative to the Public Use data—s "0.20" in the former compared to s "0.29" in the latter.[37] We reweight the Restricted data flows so that $\sum_{i:r_i=r} L^{\text{micro}}_{i,j,t} = L^{\text{macro}}_{r,j,t}$.

Estimation of the model requires further balancing to ensure identification. First, we restrict the sample to state-sector destinations that receive inflows from at least two different states. Second, among the labor markets identified in the first step, we focus on a balanced panel of state-sectors, and restrict attention to states containing all four sectors throughout. Finally, we restrict the set of origins to fall within the set of destinations, so that the final sample of flows comprises a closed system. Tables A9 and A10 present these steps and their effect on sample size for the Public Use and Restricted datasets. See Section 4 for a formal discussion of these steps relative to model identification.

Appendix B A mechanical model

We deploy a simple mechanical model of labor mobility to gain additional insight into Fact 1. In particular, we will show that the fact implies that migrants must switch sectors at a higher rate than non-migrants. Consider an economy composed of two states, a and b, that each host two sectors, 1 and 2. Initially, the economy rests in a symmetric steady-state. Each state-sector employs a unit measure of infinitely-lived workers. Migration between states is exogenous: in each period, a fraction of the population μ migrates. Sector-switching behavior may differ for migrants and non-migrants. Let σ_k denote the probability that a worker initially in sector k leaves that sector given that they do not migrate; and let σ'_k denote the same for migrants. Each period, a worker can follow one of four paths: non-migrant stayer ($(1-\mu)(1-\sigma_k)$), non-migrant switcher ($(1-\mu)\sigma_k$), migrant stayer ($\mu(1-\sigma'_k)$), and migrant switcher ($\mu\sigma'_k$). In the initial steady-state, $\sigma_k = \sigma$ and $\sigma'_k = \sigma'$ for all k. The migrant fraction of each sector is $\mu(1-\sigma') + \mu\sigma' = \mu$.

Next, consider what happens when a shock hits the economy such that the employment share of sector 2 grows to $\alpha > .5$. μ is held fixed, but the switching rates must change. Consider the case in which $\sigma_2 < \sigma_1$ and $\sigma'_2 > \sigma'_1$. The migrant fraction in sector 2 in an arbitrary region is given by,

$$\frac{\mu(1-\sigma'_2+\sigma'_1)}{(1-\mu)(1-\sigma_2+\sigma_1)+\mu(1-\sigma'_2+\sigma'_1)}$$

If $\sigma_k = \sigma'_k$, then this expression reduces to μ. If this were true in reality, then we would find $\beta_1 = 0$ in equation (1). In order to obtain $\beta_1 > 0$, the previous expression must be larger than μ. This holds if,

$$\sigma'_1 - \sigma'_2 > \sigma_1 - \sigma_2$$

[37] Reasons for underestimation is a question I'm investigating in the literature and in Census documentation.

Appendix C Testing heterogeneity in ν

Our model assumes that ν is constant over time and across demographic groups. We test these assumptions using the following relationship, implied by the Fréchet distribution:

$$\frac{\mathrm{Var}(\mathrm{wage}_{i,j,t}(n))}{\mathbb{E}[\mathrm{wage}_{i,j,t}(n)]^2} := R_{i,j,t} = \frac{\Gamma\left(1-\frac{2}{\nu}\right)}{\Gamma\left(1-\frac{1}{\nu}\right)} - 1 \tag{32}$$

A larger value of ν implies less dispersion in wages within a given i, j cell. If we were to calibrate ν using equation (32), we would take the average of $R_{i,j,t}$ weighted by $L_{i,j,t}$ and choose the value of ν that solves that equation. To test for heterogeneity, we calculate $R_{i,j,t}$ and regress it on various ex ante worker characteristics that might influence the dispersion in skills and invalidate the assumption of homogeneity. All regressions control for origin-year and destination-year fixed effects. We evaluate heterogeneity in each characteristic based on that characteristic based on the F statistics from each regression. We use our two datasets as available. For analysis in the Public Use data, we focus on state-to-state flows. Table A16 reports F statistics for a range of variables. Significant variation appears across years, age groups, and sexes.

Next, we want to know whether the variation in ν is economically significant. To do that, we compute the employment-weighted average of $R_{i,j,t}$ for each group and calculate the corresponding value of ν. The results are shown in Tables A17 and A18. We calculate smaller values of ν, corresponding to greater wage dispersion, for older, female, and non-white workers. We also observe that the value of ν falls somewhat over time. However, the differences in ν are small.

Appendix D Multiple periods of working life with forward-looking workers

Suppose workers had two periods of working life, spanning ages 20-60. This is a necessary extension in order to incorporate capital in production (see below). The worker has preferences over consumption-utility—consumption multiplied by the amenity—in each of these periods and has access to a risk-free investment instrument I at price q_t, which pays off in the next period at price q_{t+1}. We assume that workers have perfect foresight and cannot move from j after their initial choice at the start of the first working period. Hence utility of moving from i to j is,

$$U_{i,j,t} = (B_{r,t}C_t)^{1-\beta}(B_{r,t+1}C_{t+1})^{\beta}\mu_{i,j,t} = \bar{B}_{jt}C_t^{1-\beta}C_{t+1}^{\beta}\mu_{i,j,t}$$

subject to the lifetime budget constraint,

$$C_t + \frac{1}{1+r_{t+1}}C_{t+1} = w_{j,t} + \frac{1}{1+r_{t+1}}w_{j,t+1} \equiv \bar{w}_{j,t}$$

where $1 + r_{t+1} = \frac{q_{t+1}}{q_t}$. Hence indirect utility is similar to what we had before:

$$U_{ijt} = \bar{B}_{r,t}\bar{w}_{j,t}\mu_{i,j,t}$$

This modification has two important implications. First, the supply of investment goods is equal to a fraction β of national labor income. Second, we will have to modify the labor market clearing condition to incorporate multiple generations. Including the old generation will simply involve adding the workers who moved to j in the previous period. However, model calibration now returns a more complicated object, $\bar{B}_{r,t}\bar{w}_{j,t}$. We can recover the actual model fundamentals from this object using their definitions, the calibrated value of β, and our chosen exogenous interest rate.

The foregoing model can be calibrated but is difficult to solve. We calibrate $V_{j,t} = \bar{B}_{r,t}\bar{w}_{j,t}$ and then decompose it by (1) assuming V is constant after 2020 and (2) the value of β. But it's difficult to solve. Workers have to decide where to live based on the entire future path of fundamentals. This might have some appeal for realism, but it is computationally costly and tangential to our research question. We highlight this extension as a target for future research.

Appendix E Multiple periods of working life with myopic workers

The baseline model employed in the body of the paper assigns common mobility to all workers. This might overstate labor mobility if older workers are less likely to migrate or switch sectors. Here, we present an extension in which workers have two periods of working life. Young workers make decisions based on $V_{j,t}$ as in the baseline model. The extension is that these workers continue to work in old age. This provides a more realistic notion of mobility in the economy, without overcomplicating the model. We highlight how this model changes the algebra of quantification.

Estimation is unchanged, since our baseline approach already focuses on young workers. Differences arise in calibration, where we target aggregate outcomes. Now, the labor market clearing condition targets employment outcomes for young workers, $L^y_{j,t}$. We calibrate $F_{j,t}$ to satisfy,

$$L^y_{j,t} = \sum_i \lambda_{i,j,t} L^y_{i,t-1} g_{t-1,t}$$

where $\lambda_{i,j,t}$ is a function of $V_{j,t}$ as in the baseline model. The remainder of the quantification process is unchanged.

The other difference arises in the equilibrium solution method. The wage loop is modified so that agglomeration responds to the supply of efficiency units from both generations, $S_{j,t} = S^y_{j,t} + S^y_{j,t-1}$. To solve the model in 1880, we set, $S^y_{j,1860} = L^y_{j,1860}$.

E.1 Implementation: a lower bound on the role of mobility frictions in persistence

This subsection expands on the counterfactual analysis of persistence in Section 5.2. In the main model, each generation of workers lives for one period. This model likely overstates the mobility of the workforce by assigning the behavior of young workers to everyone, when in fact older workers are less mobile (Table A3). Therefore, we interpret these results as an upper bound on the importance of mobility frictions for persistence. Appendix E presents an alternative model in which workers live for two periods and older workers are completely immobile. This will provide a lower bound.

Figure 7 repeats the persistence decomposition for the two-period model. Persistence falls by 21% at a twenty-year horizon. Over longer horizons, persistence falls even lower, in contrast to the one-period case where it converged somewhat toward the baseline. This reflects complicated dynamics that arise from older workers' contribution to agglomeration.

Appendix F Persistent productivity arising from investment

Here, we derive the endogenous evolution of local sectoral productivity as an investment process. Local sectoral productivity is augmented through investment. Consider the mass of firms operating in labor market $j = (r_j, s_j)$. Firms purchase investment goods G_{jt} to attain productivity equal to,

$$A_{jt} = (G_{jt})^{\alpha_1} A_{jt-1}^{\alpha_2} \bar{A}_{jt} \tag{33}$$

where $\bar{A}_{jt}$ is the fundamental component of productivity, a structural residual held fixed in counterfactual analysis.

Efficiency units, S, are hired in a competitive market at price w_{jt}. Investment goods are purchased at an exogenous price q_t. The firm's profit maximization problem is,

$$\max_{S,G} P_{st} A_{jt} S_{jt} - \omega_{jt} S_{jt} - q_t I_{jt}$$

where s is the sector associated with labor market j, and the maximization is subject to (33). The FOCs are

$$w_{jt} = P_{st} A_{jt} \tag{34}$$

$$q_t = \alpha_1 P_{st} \bar{A}_{jt} G_{jt}^{\alpha_1 - 1} A_{jt-1}^{\alpha_2} S_{jt} \tag{35}$$

We can solve for the firm's choice of G_{jkt} using the second FOC:

$$G_{jkt} = \left(\frac{\alpha_1 P_{st} A_{jt-1}^{\alpha_2} S_{jt}}{q_t} \right)^{\frac{1}{1-\alpha_1}} \tag{36}$$

Substituting (36) into (33), we obtain,

$$A_{jt} = \frac{\alpha_1 P_{st}}{q_t} \bar{A}_{jt}^{\frac{1}{1-\alpha_1}} S_{jt}^{\frac{\alpha_1}{1-\alpha_1}} A_{jt-1}^{\frac{\alpha_2}{1-\alpha_1}} \tag{37}$$

This equation clarifies the economic consequences of the firm's investment decision. Under our chosen production structure, investment scales with the size of the workforce, S. Hence, α_1 represents the strength of agglomeration forces in static equilibrium. Agglomeration has dynamic effects as well due to the presence of A_{jt-1} on the right-hand side of equation (37). The dynamic effect is governed by the elasticity α_2, which can be thought of as depreciation.

Appendix G Tables & Figures

Table A1: Persistence of Regional Specialization

	$h = 20$	$h = 40$	$h = 60$	$h = 80$	$h = 100$	$h = 120$	$h = 140$	$h = 160$
Persistence	0.869***	0.705***	0.550***	0.401***	0.261***	0.131***	0.0442	0.0384
	(0.0139)	(0.0185)	(0.0225)	(0.0250)	(0.0264)	(0.0278)	(0.0309)	(0.0537)
N	1182	1034	886	738	590	442	294	146

Notes: We regress log employment on its lag at horizon h with state-year and sector-year fixed effects, using data on state-by-sector employment from 1860-2020. With fixed effects, a regression in terms of employment is identical to a regression using the location quotient. The sample is pooled at each horizon, so each coefficient corresponds to persistence at the relevant horizon averaged across all years of the sample in which that horizon is observed. Further details on the data are provided in Section 2.1.

Table A2: Sectors and Industries

Sector	Industries
Agriculture	Agriculture
Manufacturing	Durable and Non-Durable Goods Manufacturing
Services	Finance, Insurance, & Real Estate
	Business, Personal, and Professional Services
	Public Administration
Other	Mining & Construction
	Transportation, Communication, & Utilities
	Wholesale & Retail Trade

Table A3: Gross migration by age group

Age group	1880	1900	1920	1940	1960	1980	2000	2020
20-39	0.45	0.40	0.41	0.31	0.37	0.40	0.45	0.43
40-59	0.62	0.53	0.51	0.46	0.41	0.44	0.47	0.53

Notes: Statistics from the Public Use data prior to balancing and harmonization. Balancing results in different samples for the two age groups, so the raw data provide a consistent comparison.

Table A4: Descriptive statistics by age group

	20-40	40-60	60-80
Share of obs.	58	30	12
Migration rate	24	22	18
Share of migrants	61	29	10
Switching rate	41	39	22
Share of switchers	62	31	7

Notes: Statistics from the Restricted data prior to balancing and harmonization. Balancing results in different samples for the two age groups, so the raw data provide a consistent comparison.

Table A5: Transition matrix for age 20-40

	Share in f	Agriculture	Manufacturing	Services	Other
Agriculture	0.45	0.62	0.11	0.07	0.20
Manufacturing	0.11	0.11	0.46	0.13	0.30
Services	0.24	0.10	0.26	0.28	0.36
Other	0.21	0.12	0.22	0.13	0.53
Share in k		0.34	0.20	0.14	0.32

Notes: Statistics from the Restricted data prior to balancing and harmonization. Balancing results in different samples for the two age groups, so the raw data provide a consistent comparison.

Table A6: Transition matrix for age 40-60

	Share in f	Agriculture	Manufacturing	Services	Other
Agriculture	0.50	0.76	0.05	0.05	0.14
Manufacturing	0.11	0.12	0.50	0.11	0.26
Services	0.19	0.18	0.19	0.32	0.31
Other	0.19	0.14	0.14	0.11	0.61
Share in k		0.46	0.15	0.12	0.28

Notes: Statistics from the Restricted data prior to balancing and harmonization. Balancing results in different samples for the two age groups, so the raw data provide a consistent comparison.

Table A7: Transition matrix for non-migrants

	Share in f	Agriculture	Manufacturing	Services	Other
Agriculture	0.48	0.65	0.09	0.07	0.19
Manufacturing	0.10	0.12	0.44	0.12	0.31
Services	0.23	0.12	0.25	0.27	0.37
Other	0.19	0.13	0.20	0.13	0.53
Share in k		0.38	0.18	0.13	0.31

Notes: Statistics from the Restricted data prior to balancing and harmonization.

Table A8: Transition matrix for migrants

	Share in f	Agriculture	Manufacturing	Services	Other
Agriculture	0.45	0.47	0.14	0.12	0.28
Manufacturing	0.10	0.14	0.38	0.16	0.32
Services	0.25	0.11	0.23	0.29	0.37
Other	0.21	0.14	0.21	0.17	0.48
		0.28	0.20	0.18	0.35

Notes: Statistics from the Restricted data prior to balancing and harmonization.

Table A9: Harmonization in the Public Use Data

Step	1860	1880	1900	1920	1940	1960	1980	2000	2020
0. Raw population (millions).	4.08	7.69	12.71	20.78	26.06	28.27	53.85	62.99	70.89
1. Keep states w/all secs. & years.	0.98	0.97	0.95	0.93	0.94	0.94	0.93	0.92	0.91
2. Keep origins s.t. Step 1.	0.72	0.78	0.77	0.75	0.85	0.85	0.83	0.75	0.73

Notes: Each row after Row 0 reports one step in the balancing procedure and the share of employment remaining after implementing restrictions up to that point.

Table A10: Harmonization in the Public Use Data

Step	1880	1900	1920	1940
0. Raw population (millions).	0.27	0.40	0.60	1.46
1. Merge Public Use data.	1.00	1.00	1.00	1.00
2. Restrict to positive flows in both.	0.99	0.99	1.00	1.00

Notes: Each row after Row 0 reports one step in the balancing procedure and the share of employment remaining after implementing restrictions up to that point.

Table A11: Summary of Public Use Data cross sections

	1880	1900	1920	1940	1960	1980	2000	2020
Sectoral employment shares								
Agriculture	0.52	0.39	0.25	0.15	0.05	0.02	0.02	0.02
Manufacturing	0.12	0.15	0.25	0.27	0.31	0.23	0.15	0.09
Services	0.16	0.20	0.22	0.26	0.33	0.42	0.49	0.53
Other	0.20	0.26	0.29	0.33	0.32	0.33	0.35	0.35
Regional employment shares								
Northeast	0.28	0.27	0.27	0.30	0.26	0.22	0.20	0.18
Midwest	0.34	0.35	0.34	0.31	0.30	0.28	0.27	0.25
South	0.37	0.35	0.33	0.32	0.32	0.35	0.38	0.40
West	0.02	0.03	0.06	0.07	0.12	0.15	0.14	0.17
Gross migration	0.30	0.24	0.24	0.23	0.30	0.31	0.31	0.28
Net migration	0.13	0.06	0.06	0.06	0.09	0.07	0.04	0.06
Net sectoral reallocation	0.02	0.14	0.14	0.10	0.11	0.10	0.09	0.05

Notes: Historical facts from the harmonized data.

Table A12: Summary of Restricted Data cross sections

	1880	1900	1920	1940
Sectoral employment shares				
Agriculture	0.52	0.39	0.25	0.15
Manufacturing	0.12	0.15	0.25	0.27
Services	0.16	0.20	0.22	0.26
Other	0.20	0.26	0.29	0.33
Regional employment shares				
Northeast	0.28	0.27	0.27	0.30
Midwest	0.34	0.35	0.34	0.31
South	0.37	0.35	0.33	0.32
West	0.02	0.03	0.06	0.07
Gross migration	0.31	0.24	0.24	0.23
Net migration	0.14	0.06	0.06	0.06
Net sectoral reallocation	0.10	0.14	0.14	0.10
Gross sectoral reallocation	0.41	0.47	0.55	0.63

Notes: Historical facts from the harmonized data.

Table A13: Migrant fraction regression for alternate samples

	(1)	(2)	(3)	(4)	(5)	(6)
	Labor markets		States		Sectors	
	Growing	Shrinking	Growing	Shrinking	Growing	Shrinking
$\Delta \frac{L_{j,t}}{L_{r,t}}$	0.47^{***}	0.37^{***}	0.24^{***}	0.14^{**}	0.15^{**}	0.14^{*}
	(0.12)	(0.11)	(0.06)	(0.06)	(0.06)	(0.08)
N	635.00	384.00	516.00	668.00	740.00	296.00
R2	0.86	0.89	0.88	0.73	0.87	0.89
F	16.38	11.59	17.73	6.35	6.00	3.15

Notes: Each regression restricts the sample based on whether employment is growing or shrinking. Labor markets: $\Delta \frac{L_{j,t}}{L_{r_j,t}} \lesseqgtr 0$. States: $\Delta L_{r,t} \lesseqgtr 0$. Sectors: $\Delta L_{s,t} \lesseqgtr 0$. Total workforce normalized to 1 so that $L_{j,t}$, $L_{r,t}$, and $L_{s,t}$ represent shares. Standard errors in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A14: Employment RCA of workers by place of origin

	Agriculture	Manufacturing	Services	Other
US & Canada	0.51	1.22	1.19	1.03
Central & South America	0.60	0.96	1.34	1.07
Europe	0.47	1.24	1.14	1.06
Asia & Middle East	1.10	0.63	1.67	0.96
Africa	0.17	1.17	1.85	0.93
Oceania	0.82	0.88	1.95	0.72

Notes: The table reports revealed comparative advantage for the sector indicated in each columns, in terms of 1940 employment, for different immigrant groups (S_s^m in the text), averaged by continent.

Table A15: Regression estimates of mobility frictions

(a) Young Workers

	(1)	(2)
Fixed cost, 1880	4.175***	4.137***
	(0.211)	(0.213)
Fixed cost, 1900	4.268***	4.243***
	(0.156)	(0.156)
Fixed cost, 1920	3.638***	3.623***
	(0.142)	(0.141)
Fixed cost, 1940	3.424***	3.409***
	(0.123)	(0.121)
Fixed cost, 1960		2.210***
		(0.109)
Fixed cost, 1980		1.232***
		(0.106)
Fixed cost, 2000		1.503***
		(0.105)
Fixed cost, 2020		1.606***
		(0.131)
Log dist., 1880	-1.379***	-1.374***
	(0.0349)	(0.0350)
Log dist., 1900	-1.415***	-1.413***
	(0.0242)	(0.0244)
Log dist., 1920	-1.301***	-1.301***
	(0.0217)	(0.0217)
Log dist., 1940	-1.297***	-1.297***
	(0.0184)	(0.0182)
Log dist., 1960		-1.054***
		(0.0161)
Log dist., 1980		-0.874***
		(0.0153)
Log dist., 2000		-0.904***
		(0.0153)
Log dist., 2020		-0.941***
		(0.0193)
Switching cost	-0.925***	-0.818***
	(0.0155)	(0.0120)
Migrate × switch	0.345***	0.361***
	(0.0253)	(0.0163)
N	87320	174936

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

(b) Old Workers

	(1)	(2)
Fixed cost, 1880	4.706***	4.652***
	(0.503)	(0.503)
Fixed cost, 1900	3.897***	3.865***
	(0.223)	(0.226)
Fixed cost, 1920	4.694***	4.672***
	(0.139)	(0.140)
Fixed cost, 1940	4.367***	4.359***
	(0.121)	(0.121)
Fixed cost, 1960		4.366***
		(0.128)
Fixed cost, 1980		2.894***
		(0.127)
Fixed cost, 2000		2.446***
		(0.128)
Fixed cost, 2020		2.327***
		(0.141)
Log dist., 1880	-1.401***	-1.391***
	(0.0858)	(0.0860)
Log dist., 1900	-1.276***	-1.272***
	(0.0366)	(0.0371)
Log dist., 1920	-1.445***	-1.443***
	(0.0217)	(0.0219)
Log dist., 1940	-1.411***	-1.411***
	(0.0187)	(0.0187)
Log dist., 1960		-1.423***
		(0.0198)
Log dist., 1980		-1.147***
		(0.0192)
Log dist., 2000		-1.044***
		(0.0193)
Log dist., 2020		-1.038***
		(0.0215)
Switching cost	-1.319***	-1.163***
	(0.0180)	(0.0138)
Migrate × switch	0.543***	0.537***
	(0.0281)	(0.0190)
N	57600	115200

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Notes: PPML estimates of mobility frictions. Panel A15a shows the results for workers age 20-39. Panel A15b shows the results for workers age 40-59. In each panel, column (1) shows the results from the Restricted data and column (2) shows the results from the disaggregated Public Use data. The sample of origins and destinations is smaller for older workers due to balancing. Estimated mobility frictions for the two age groups are qualitatively similar: Switching is less costly for migrants than for non-migrants. Migration has a positive fixed "cost" and a negative distance elasticity. The distance elasticity declines in the latter half of the sample, but the change takes place twenty years later for older workers, as one would expect if these workers were more mobile in their younger years. In addition, the correlation of estimated $W_{j,t}$ across the two samples is 0.95.

Table A16: Testing for heterogeneity in wage dispersion by ex ante worker characteristics

	F statistic	
	Restricted	Public Use
State of origin	0.22	2.29
Sector of origin	0.01	
Father's ownership	0.01	
Father's urbanicity	0.01	
Father's occupation score	0.01	
Year		817.93
Age		714.15
Sex		78.67
Race		8.82

Table A17: Implied values of ν for different demographic groups

Group	ν
Age 20-39	2.67
Age 40-59	2.62
Age 60+	2.52
Male	2.73
Female	2.71
White	2.75
Black	2.70
Other	2.59

Table A18: Implied values of ν over time

Year	ν
1940	2.92
1960	3.05
1980	2.97
2000	2.64
2020	2.58

Table A19: Persistence: First Stage

	(1)
1900	0.389***
	(0.0422)
1920	0.323***
	(0.0488)
1940	0.351***
	(0.0526)
1960	0.290***
	(0.0386)
1980	0.225***
	(0.0356)
2000	0.209***
	(0.0171)
2020	0.218***
	(0.0149)
N	1184
R2	0.527
F	136.4

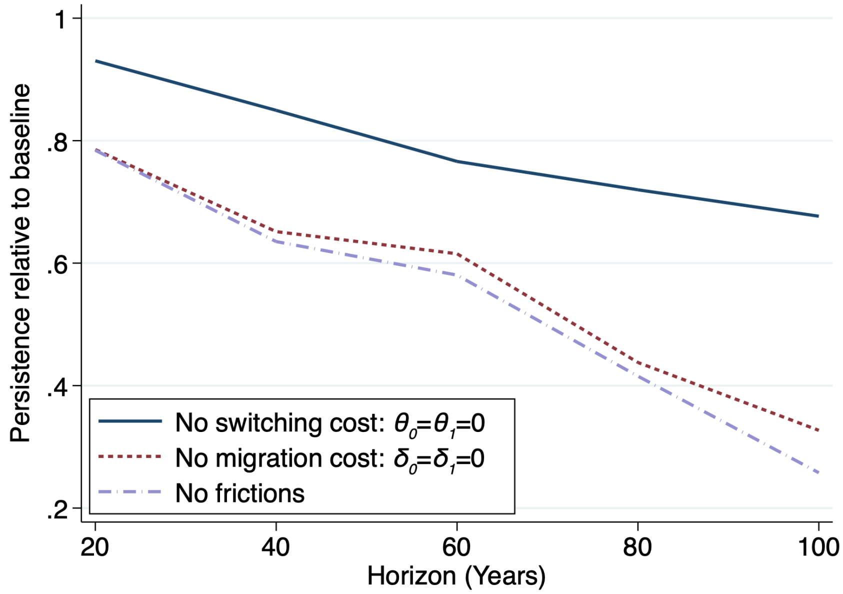
Notes: First stage regression of log efficiency units, $S_{j,t}$, on embodied skill (equation (30)) interacted with bidecadal dummies. Controls include $\log A_{j,t-1}$ and sector-year fixed effects. 1880 is excluded because productivity in 1860 is not calibrated from the data. Standard errors are clustered by sector-year and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Table A20: 2SLS Estimates of α_1 and α_2 for S_s^m defined in different years

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	1900	1920	1940	1960	1980	2000	2020
logS	0.103***	0.0994***	0.0980***	0.0961***	0.0964***	0.0965***	0.0971***
	(0.0262)	(0.0261)	(0.0261)	(0.0261)	(0.0265)	(0.0266)	(0.0267)
lagloga	0.668***	0.671***	0.672***	0.673***	0.673***	0.673***	0.672***
	(0.0613)	(0.0614)	(0.0614)	(0.0615)	(0.0618)	(0.0618)	(0.0619)
N	1036	1036	1036	1036	1036	1036	1036
CDW F	31.76	31.40	31.21	31.26	30.58	30.44	30.13

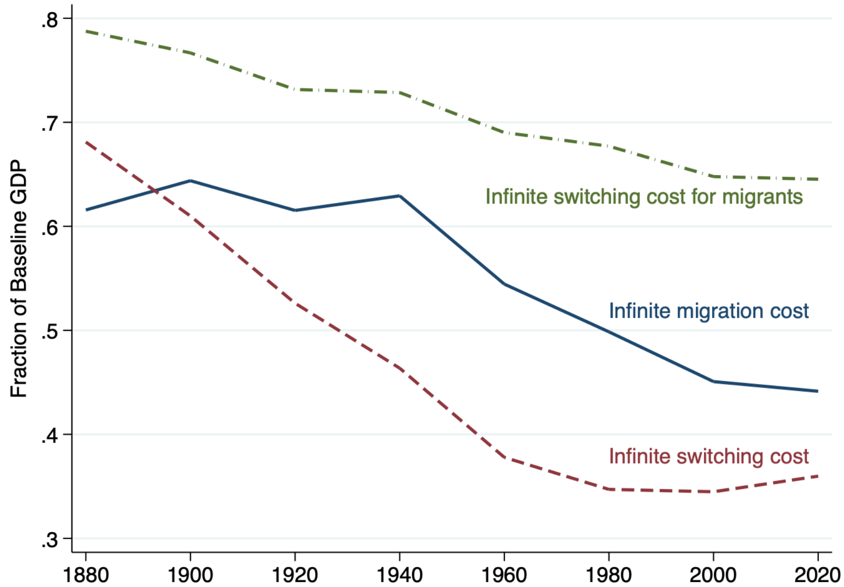
Notes: 2SLS estimates of composite productivity on efficiency units and lag productivity (equation (29)). Log efficiency units is instruments with log embodied skill (equation (30)). Controls include $\log A_{j,t-1}$ and sector-year fixed effects. In each column, immigrants' revealed comparative advantage, used to construct the instrument, is computed in the year indicated in the column title. The main text presents results for 1940. Standard errors are clustered by sector-year and shown in parentheses. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$.

Figure 7: Rank-rank elasticity of location quotients relative to baseline estimates



Notes: Each line represents the rank-rank elasticity of location quotient (LQ) within each sector, relative to the baseline estimate. Elasticities are calculated as follows. We compute LQ for each state-sector, rank these within each sector, and regress log rank on its lag at various horizons, controlling for state-year and sector-year dummies. The estimate for each horizon averages across the years in which that horizon is observed.

Figure 8: GDP without mobility frictions



Notes: Each line depicts GDP (relative to the baseline) in a counterfactual economy with modified mobility frictions simulated forward from 1860. $\bar{A}_{j,t}$ and $B_{r,t}$ take on their estimated values. $A_{j,t}$ and $P_{s,t}$ will change in counterfactual equilibria.