Human Capital, Ancestry, and Skill Composition

Murad Zeynalli*

October 15, 2025

[Click here for the latest version]

Abstract

The spatial correlation between worker skills and industry skill intensity is among the best documented features of US economic geography. Yet, the causal impact of human capital on the industrial skill composition of regions remains largely unknown, despite its fundamental role for local economies to adapt demographic changes. This paper studies how ancestry-induced shifts in historical human capital affect the contemporary industrial skill composition of US counties. Leveraging quasi-random origin-by-destination immigration patterns from 1860 to 2010, I isolate exogenous variation in skill-specific local working-age population at the county level for 1970-2010. I find that increases in medium- and high-skill worker shares raise employment and establishment shares in high-skill industries while reducing them in low-skill industries. The positive effects are concentrated in nontradable sectors, which expand in response to stronger local demand, whereas tradable sectors contract as the relative scarcity of low-skill labor reduces their local competitiveness. These findings align with a model in which firms producing differentiated goods employ labor of varying skill types with limited substitutability, implying that shifts in human capital composition induce partial but systematic reallocation of economic activity toward skill-intensive industries.

Keywords: Human Capital, Persistence, Ancestry, Workers, Establishments, Industries, Skills.

JEL Codes: D22, F22, J21, J24, N91, N92, R12.

---

*Department of Economics, Miami University (Email Address: zeynalm@miamioh.edu). I am extremely grateful to Daniel Keniston for his unwavering guidance and support in this project. I thank Areendam Chanda and Scott Abrahams for insightful feedback, and I appreciate valuable comments and suggestions from Federico Mandelman, Giovanni Peri, Gaurav Khanna, Santiago Pérez, Catalina Amuedo-Dorantes, and Jafar Jafarov. I am grateful to participants of the 2022–2024 LSU Department of Economics Brown Bag seminars, the 93rd Annual Meeting of the Southern Economic Association, the 39th Meeting of the European Economic Association and 76th European Meeting of the Econometric Society, the 3rd Mexican Summer School on the "Economics of Migration," and the 2025 Miami Department of Economics Brown Bag seminar for helpful discussions. All errors are my own.

1 Introduction

There exist significant differences in the skill composition of industries across the United States. This skill heterogeneity has been attributed to technological change and job polarization, as well as to local demand shocks and amenities.[1] However, an additional channel shaping the spatial sorting of skills and industries may lie in variation in human capital across local labor markets. Regions with more educated workforces can attract employers that rely more intensively on skilled labor, while areas with limited human capital may specialize in low-skill production. In this sense, shifts in local human capital supply can causally impact industrial specialization by altering the relative costs and availability of labor inputs. Yet, despite its theoretical relevance, the empirical evidence on how labor supply shocks shape the industrial skill composition of US regions remains limited.[2]

This paper provides empirical evidence on how exogenous shifts in local human capital, induced by historical immigration, affect the industrial skill composition of employment and establishments across US counties. Examining both employment and establishment outcomes captures how shifts in local human capital translate into adjustments in the allocation of labor and productive activity across industries. The importance of this analysis lies in its potential policy implications such that it highlights the mechanisms through which workforce composition drives regional industrial structure and inform ongoing debates about place-based and workforce development policies. For policymakers seeking to improve workers' earnings and promote firm creation, a central question is whether increasing local human capital can indeed generate "high-paying jobs."[3] For instance, policies that aim to attract high-tech or knowledge-intensive firms, often through education or training initiatives, may succeed only if firms endogenously respond to the augmented skill supply. Otherwise, the gains may dissipate if firms fail to enter, or if skilled workers relocate elsewhere. Understanding how the labor supply of different skill types affects the industrial skill distribution across US counties is thus essential for designing effective regional development strategies.

The primary empirical challenge lies in the endogeneity of local human capital with respect to unobserved labor demand shocks. For example, if a county is characterized by a high concentration of skill-intensive industries offering well-paid jobs and entrepreneurial opportunities, highly educated individuals may endogenously sort into that area. Such selective migration would gen-

[1] See Acemoglu (2011), Autor and Dorn (2013), Moretti (2010), and Diamond (2016).

[2] See Bartik (1991), Notowidigdo (2020), and Giannone (2022) for related evidence on local labor supply and industrial adjustment.

[3] For relevant discussions on skill upgrading and local wage premia, see Moretti (2004) and Greenstone et al. (2010).

erate spurious correlations between the local share of high-skill population and the employment or establishment shares of high-skill industries. To address this concern, I present three empirical patterns that mitigate the potential endogeneity of human capital.

First, much of the spatial variation in contemporary human capital reflects the path dependence and persistence (Allen and Donaldson, 2020; Voth, 2021) in historical immigration. Immigrants from different origin countries have arrived with systematically distinct skill levels and occupational specializations. This selection process (Borjas, 1987; Abramitzky et al., 2012; Clemens and Mendola, 2024) has generated cross-location differences in the composition of human capital by shifting the distribution of skills across industries (Goldin, 1994; Abramitzky et al., 2014; Boberg-Fazlić and Sharp, 2024). Second, historical settlement patterns of immigrants has produced heterogeneity in educational attainment across localities through the inter-generational transmission of human capital (Abramitzky and Braggion, 2006; Sequeira et al., 2020). Finally, immigrants have historically clustered in areas with pre-existing "ethnic enclaves," demonstrating the non-random nature of immigrant settlement (Altonji and Card, 1991; Card, 2001). "Social networks" have further amplified these patterns by attracting subsequent cohorts of immigrants with comparable skill endowments (Munshi, 2003).

Directed by the potential mechanisms above, I proceed with my identification strategy in the following steps. First, using detailed data on ancestry groups (Fulford et al., 2019), I predict ancestry stocks for each census year in 1970-2010 by leveraging historical origin-by-destination immigration patterns (*i.e.*, the "push" and "pull" factors) from 1860 to 2010 within an advanced version of the shift-share instrumental variable (SSIV) approach (Burchardi et al., 2019).[4] This construction addresses issues typical of the canonical shift-share design serial correlation in immigration patterns (Jaeger et al., 2018), correlations between levels or trends in economic outcomes and current immigration (Goldsmith-Pinkham et al., 2020; Borusyak et al., 2022), and the over-rejection problem (Adao et al., 2019). This procedure generates granular, quasi-random variation in ancestry, alleviating these concerns along with endogeneity arising from immigrant self-sorting into destinations. Second, utilizing this predicted ancestry, I isolate the exogenous population of each skill type, measured by the number of individuals with varying educational attainment at the county level for the decennial 1970-2010 period.

[4]Note that Burchardi et al. (2019) predict the number of individuals with reported ancestries at the county level, reflecting the changing nature of ethnic identity as a social construct. By contrast, I use a different measure of ancestry in the spirit of Fulford et al. (2019) that highlights genealogical construction by taking internal migration, generational transmission, and population growth of ancestries into account.

I make two major contributions to the relevant literature. My first contribution is to quantify the causal impact of human-capital-specific local population on the industrial skill composition of employment and establishment shares at the county level. The studies in the existing literature examine how exogenous changes in local labor supply, such as immigration inflows (Altonji and Card, 1991; Borjas, 2003; Ottaviano and Peri, 2012; Dustmann et al., 2017; Droller, 2018; Burchardi et al., 2019), demographic transitions (Maestas et al., 2023), or internal migration responses to local shocks (Monras, 2018), affect native labor market outcomes, local demand, and regional economic activity. In contrast, I focus on how the industrial skill distribution of workers and establishments adjusts to skill-specific exogenous shifts in the working-age population generated by predicted ancestry-based immigration patterns. Rather than studying aggregate effects on wages or employment levels, I analyze how local industries reallocate employment and establishments across skill cells that closely mirror the grouping of working-age individuals by educational attainment. To further capture heterogeneity in adjustment mechanisms, I distinguish between tradable and nontradable industries, whereby the former reflect exposure to global competition and the latter capture local demand multipliers.

My second contribution is to generate skill-specific exogenous variation in working-age population at the county level, thus enabling identification of the geographic distribution of workers and establishments. This approach addresses the challenge of disentangling supply-side responses from demand-driven changes in the local composition of skilled labor. Individuals with varying educational attainment tend to relocate to places where firms' demand for specific skills aligns with their human capital. For instance, concentrations of high-technology firms, such as those in Palo Alto, attract highly educated workers. Without accounting for this endogenous sorting, local shares of high-skill population and employment or establishment shares in high-skill industries would be spuriously correlated. To overcome this issue, I construct exogenous variation in the working-age population of each skill type at the county level by exploiting the predicted ancestry stocks explained above. This variation is driven by (*i*) heterogeneity in immigrant human capital as well as occupational and task specializations led by the immigrant selection process, (*ii*) the path dependence in the inter-generational transmission of human capital from earlier immigrant cohorts to their descendants, and (*iii*) the persistence of "ethnic enclaves" and "social networks" that shape subsequent immigrant inflows.

To guide the empirical analysis, I utilize a constant elasticity of substitution (CES) model with two distinct representative firms (white and blue collar), each specializing in producing differentiated output sold in a perfectly competitive market. The firms employ labor of different skill types (high or low skill) with varying intensities (Autor et al., 1998) and face imperfect substitutability across education groups (Katz and Murphy, 1992; Card and Lemieux, 2001). The model predicts that an increase in the share of high-skill labor supply raises the labor demand share of white-collar firms, whereas a rise in the share of low-skill labor supply reduces it. Conversely, the labor demand share of blue-collar firms increases with a larger share of low-skill labor supply and declines when the share of high-skill labor supply expands. The magnitude of these adjustments depends on the elasticity of substitution between skill types: higher substitutability dampens changes through smooth input reallocation within firms, while lower substitutability amplifies them through the reallocation of production across sectors.

I find that, relative to low-skill population share, increases in medium- and high-skill population shares reallocate employment and establishments toward high-skill industries. A one-percentage-point rise in these population shares raises high-skill industry employment by approximately 0.2–0.5 percentage points and reduces employment in low-skill industries by about 0.3–0.4 percentage points, with medium-skill industries showing smaller and mixed responses. Establishment shares exhibit comparable patterns: high-skill industries expand by roughly 0.3–1 percentage point, while low-skill industries contract modestly, and medium-skill industries experience slight declines following an increase in high-skill population share. The magnitudes of these effects suggest a moderate elasticity of substitution between high- and low-skill labor. Firms can substitute across skill types to some degree, but imperfect substitutability amplifies sectoral reallocation. Increases in high-skill labor supply thus expand high-skill industries and contract low-skill ones, while within-industry adjustments remain limited. These patterns overlap conceptually with Moretti (2004) and Shapiro (2006), who find that higher local human capital boosts productivity and shifts employment toward skill-intensive activities. However, unlike their focus on productivity and demand-side spillovers that raise aggregate employment and wages, my results highlight a reallocation mechanism, whereby exogenous increases in local human capital reshape the industrial composition itself, shifting establishments and employment toward more skill-intensive sectors.

To examine heterogeneity by industry tradability, I find that most positive effects in the baseline analysis originate from the nontradable sector, while negative effects primarily arise in the tradable sector. These results are consistent with nontradable industries expanding in response to stronger local demand (Moretti, 2010; Mian and Sufi, 2014) that amplifies the gains from a more skilled workforce. In contrast, tradable industries experience reduced local competitiveness (Autor and Dorn, 2013), reflecting the costs associated with the relative scarcity of low-skill labor. Additional heterogeneity by establishment size shows that the positive effects are concentrated among medium-sized establishments in the nontradable sector, whereas the negative effects are absorbed mainly by small establishments across both sectors. These patterns support a mechanism in which differences in labor productivity drive the direction of adjustment: tradability determines whether industries expand locally or adjust externally, and smaller and medium-sized establishments exhibit the greatest responsiveness.

Lastly, I carry out a set of robustness checks, which include the use of alternative instruments, various sample selections, and alternative tradability classification. Additionally, I implement placebo checks and randomization tests highlighted in Adao et al. (2019) to address the over-rejection problem typical of the SSIV designs. The results demonstrate that my main estimates are robust to all sensitivity checks.

Related Literature. This paper connects to several strands of research. First, it contributes to studies examining how industries and establishments respond to quasi-experimental labor market shocks. The literature exploits immigration policies as sources of exogenous local labor supply variation to assess their effects on workers and firms (Kerr et al., 2015; Bound et al., 2017; Khanna et al., 2018; Doran et al., 2022; Mahajan et al., 2024; Clemens and Lewis, 2022; Amuedo-Dorantes et al., 2023).[5] In this context, I evaluate how ancestry-induced exogenous shifts in local population composition affect employment and establishment shares across industries defined by skill intensity, showing that workers and establishments respond asymmetrically to changes in the local human capital mix.

Second, the paper adds to the literature emphasizing the persistence of historical events in shaping present-day economic outcomes (Nunn, 2014; Allen and Donaldson, 2020; Voth, 2021; Abramitzky and Braggion, 2006; Bleakley and Lin, 2012; Bazzi et al., 2020; Sequeira et al., 2020; Fulford et al., 2019; Rocha et al., 2017; Droller, 2018; Valencia Caicedo, 2019). Relative to these studies, I highlight how quasi-random variation in historical immigrant settlement patterns induced long-run differences in local human capital. Historical immigration at fine geographic scales remains a powerful predictor of educational attainment among the working-age population in US counties.

Third, I build on empirical work employing the shift–share instrumental variable (SSIV) approach to address endogeneity in immigration studies (Altonji and Card, 1991; Card, 2001; Bor-

[5] Ulltveit-Moe et al. (2019) quantifies industry and occupational adjustments to immigration-driven labor supply shocks.

jas, 2003, 2005; Dustmann et al., 2017; Ottaviano and Peri, 2012; Foged and Peri, 2016; Pandey and Chaudhuri, 2017). While these studies rely on the canonical Card-type design, I use an advanced SSIV framework following Burchardi et al. (2019).[6] This design accounts for origin- and destination-specific confounding shocks by leaving out relevant geographic units in both the push and pull components.

Finally, I complement research employing a CES production framework with imperfectly substitutable labor of different skill levels (Katz and Murphy, 1992; Card and Lemieux, 2001), in which representative firms specialize in using particular types of labor more intensively (Autor et al., 1998). These models emphasize the role of firm specialization in shaping labor demand across skill groups, providing theoretical grounding for the empirical work examined here.

The remainder of the paper is organized as follows. Section 2 presents the theoretical framework, introducing the model and deriving the comparative statics. Section 3 outlines the empirical strategy, addressing endogeneity concerns in human capital and describing the identification strategy. Section 4 details the data sources, explains the skill and tradability classifications, and reports summary statistics. Section 5 presents the main results and discussion. Section 6 examines the robustness of the findings. Section 7 concludes.

2 Theoretical Framework

In this section, I outline a theoretical framework that guides the subsequent empirical analysis. The model features two firm types producing differentiated output, each employing labor of a specific skill group more intensively, following Autor et al. (1998). Production occurs within a constant elasticity of substitution (CES) environment that allows for imperfect substitution across skills (Katz and Murphy, 1992; Card and Lemieux, 2001). Unlike the standard approach that focuses on wage adjustment, the framework emphasizes how variation in relative labor supplies shifts the equilibrium allocation of production across firm types, shaping the local industrial structure and patterns of regional development.

2.1 Setup

There exist two representative firms, blue and white collar, indexed by $i \in \{b, w\}$, each specializing in producing differentiated output, which they sell into a perfectly competitive market. Each type

[6] See also Terry et al. (2021) for a closely related identification strategy linking immigration to innovation and growth, as well as Burstein et al. (2020) and Caiumi and Peri (2024) for alternative extensions.

of firm produces output, Q_i, by combining labor of high-skill ($L_{i,h}$) and low-skill workers ($L_{i,l}$) according to a CES production function:

$$Q_i = \left(\beta_{i,h} L_{i,h}^{\frac{\sigma-1}{\sigma}} + \beta_{i,l} L_{i,l}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}. \tag{1}$$

Each firm utilizes labor of either high- or low-skill workers, $e \in \{h, l\}$ more intensively. The relative productivity parameters, $\beta_{i,h}$ and $\beta_{i,l}$, differ across firm types such that $\beta_{w,h} > \beta_{b,h}$ and $\beta_{b,l} > \beta_{w,l}$, with $\beta_{i,h} + \beta_{i,l} = 1$. Within each firm, white-collar firms employ high-skill labor more intensively ($\beta_{w,h} > \beta_{w,l}$), whereas blue-collar firms rely more heavily on low-skill labor ($\beta_{b,l} > \beta_{b,h}$). The parameter σ denotes the elasticity of substitution between labor types. These assumptions imply that low-skill labor is relatively more productive in blue-collar firms, whereas high-skill labor is relatively more productive in white-collar firms.

2.2 Comparative Statics

The differences in relative productivity parameters and the assumption of imperfect substitutability between high- and low-skill labor imply that $B_b > 1$, $B_w < 1$, $B_b > B_w$, and $Z > 0$.[7] Therefore, the response of labor demand share of each firm type to the relative labor supply of each skill type becomes

$$\frac{\partial s_w}{\partial n_h} = \left(\frac{B_b}{B_b - B_w} \right) \left(\frac{Z + B_w}{Z} \right) > 0, \tag{2}$$

$$\frac{\partial s_b}{\partial n_h} = -\left(\frac{B_w}{B_b - B_w} \right) \left(\frac{Z + B_b}{Z} \right) < 0, \tag{3}$$

$$\frac{\partial s_b}{\partial n_l} = \left(\frac{Z + B_b}{B_b - B_w} \right) > 0, \tag{4}$$

$$\frac{\partial s_w}{\partial n_l} = -\left(\frac{Z + B_w}{B_b - B_w} \right) < 0. \tag{5}$$

An increase in the relative supply of high-skill workers raises the labor-demand share of white-collar firms, while a rise in low-skill labor supply reduces it, as shown in equations (2) and (5). Conversely, the labor-demand share of blue-collar firms increases with greater low-skill labor supply and declines when high-skill labor becomes more abundant, as shown in equations (4) and (3). These results follow from differences in relative productivity across labor types. White-

[7]The full derivations of the model are presented in Appendix A.

collar firms have a comparative advantage in employing high-skill labor ($\beta_{w,h} > \beta_{b,h}$), whereas blue-collar firms are relatively more productive with low-skill labor ($\beta_{b,l} > \beta_{w,l}$). Thus, a higher supply share of high-skill labor reallocates production toward the high-skill–intensive sector, raising its employment share and lowering that of the low-skill–intensive sector. Likewise, a larger supply share of low-skill labor expands the relative size of blue-collar firms. These responses reflect firm specialization combined with imperfect substitutability between labor types.

The magnitude of these comparative statics depends on σ, which governs how easily firms can substitute between high- and low-skill labor within production. When labor inputs are more substitutable (higher σ), firms can adjust input mixes internally, leading to smaller changes in employment shares across firm types. In contrast, when labor inputs are less substitutable (lower σ), substitution within firms is limited, and the adjustment occurs mainly through reallocation of production across sectors, amplifying changes in employment and establishment shares.[8]

Empirically, this mechanism implies the following predictions. Regions or industries characterized by lower estimated substitutability between skill groups should exhibit stronger reallocation of employment and establishment shares in response to changes in labor supply composition. Conversely, when skills are more easily substituted, these compositional adjustments are expected to be attenuated.

3 Empirical Strategy

3.1 Estimating Equation and Potential Threats to Identification

Guided by the theoretical framework, the empirical analysis tests whether changes in the relative supply of different skill groups translate into shifts in the industrial composition of local labor markets. Specifically, I examine how the shares of medium- and high-skill population, measured relative to low-skill population, affect employment and establishment shares of industries of varying skill intensity across US counties over time. Considering both employment and establishment outcomes allows me to capture adjustment along two related margins of local industry structure, reflecting changes in the allocation of workers and establishments within local economies. The estimating equation takes the following form:

[8] In the limiting cases, as $\sigma \to 0$ (perfect complements) or $\sigma \to \infty$ (perfect substitutes), the model collapses to settings in which all adjustment occurs either entirely between or entirely within firms, respectively. Therefore, the analysis here assumes imperfect but finite substitution, $1 < \sigma < \infty$.

$$y_{e,d,t} = \alpha_{s(d)} + \alpha_t + \boldsymbol{n}_{e,d,t}\boldsymbol{\beta_e} + \varepsilon_{e,d,t}, \tag{6}$$

where $\boldsymbol{n}_{e,d,t}$ is a vector of medium- and high-skill population shares ($i.e.$, $e \in \{medium, high\}$) in county d at time t, measured by human capital levels. For each decennial census year, I define this variable as the number of working-age individuals with certain educational attainment by county relative to the entire working-age county population. I classify individuals with a high school degree and less as low skill, those with some college education as medium skill, and those with college degree and above as high skill. $y_{e,d,t}$ is the outcome of interest, which is employment and establishment shares of industry skill type e in county d at time t in separate specifications. The classification of industries into skill cells closely follows the human capital classification. In particular, I classify industries with average years of education corresponding to a high-skill degree and less as low skill, those with mean education equivalent to some college as medium skill, and those with average education comparable to college degree and above as high skill.

I define the employment share variable for each skill type as the number of workers in industries of each skill type by county divided by the entire county workforce in each census year. I construct the establishment share variable of each skill type as the number of establishments in industries of each skill type by county relative to the total number of establishments by county in each census year. $\alpha_{s(d)}$ and α_t represent state and time (decennial census year) fixed effects controlling for any time-invariant state-wide factors and aggregate shocks, as well as time-specific trends. The identifying variation, therefore, arises from changes across counties within the same state over time. The error term $\varepsilon_{e,d,t}$ captures all omitted factors. Standard errors are clustered at the county level to account for arbitrary serial correlation within counties, which is the level of treatment variation.[9]

The OLS estimates of the coefficient of interest, $\boldsymbol{\beta_e}$, however, are likely biased, since unobserved confounding factors may affect the industrial skill composition of both employment and establishments in counties despite the inclusion of state and time fixed effects. This problem may lead to reverse causality and omitted variable bias, which may emerge due to the local labor market and firm creation opportunities being likely correlated with local productivity shocks. Consequently, these opportunities may attract individuals with certain educational attainment to a specific locality. For instance, the major concentration of the oil and gas industry in Houston

[9] Clustering at the state level would be excessively conservative given the relatively small number of states and the sub-state nature of the analysis.

(Harris county) may draw in high-skill petroleum engineers. Furthermore, as this sector experiences a boom, it may result in an expansion of oil refineries (*i.e.*, establishments). Therefore, these local confounding shocks would generate spurious correlations between population human capital shares and the outcomes of interest. In the following subsection, I underline the mechanisms as motivating facts to account for this endogeneity concern.

3.2 What Explains Today's Varying Human Capital Levels across US Counties?

The first mechanism that contributes to the variation in the current economic conditions in US localities is the heterogeneity in skills and occupational specialization of immigrants in both historical and modern times. To the extent that retrospective immigrant selection process (Borjas, 1987; Abramitzky et al., 2012) across different origins is instrumental in affecting future economic outcomes in destinations, one can illustrate shifts in the skill composition of industries through variation in immigrant human capital.[10] For instance, some immigrants historically provided unskilled labor (Bergquist, 2007; Goldin, 1994) and crucial skills for industries (Malone, 1935), while others provided knowledge and innovation in various industries (Wittke, 1939; Boberg-Fazlić and Sharp, 2024).

Findings in Abramitzky et al. (2014) reveal substantial heterogeneity in the 1900-level occupational distribution of immigrants. The similar patterns on the task specialization, innovation, comparative advantage, and skills can still be observed amongst the relatively recent immigrants (Peri and Sparber, 2009; Hunt and Gauthier-Loiselle, 2010; Chiquiar and Hanson, 2005; Chiswick and Taengnoi, 2007; Hanson and Liu, 2023; Clemens and Mendola, 2024).

Building on the channel identified above, the second mechanism pertains to *persistence* and *path dependence*, the long-lasting impact of historical events (Nunn, 2014; Allen and Donaldson, 2020; Voth, 2021). From the human capital perspective, one likely channel through which the spatial differences in human capital levels exist at present can be attributed to historical immigration. In particular, settlements of earlier immigrants have induced varying degrees of educational attainment amongst local population via the transmission of human capital. For example, Sequeira et al. (2020) find that counties with a higher exposure to historical immigration have a larger level of educational attainment today.[11] Abramitzky and Braggion (2006) explain the comparative economic performance in the subsequent periods through the historical positive and negative selections

[10] The immigrant selection process is built upon the selection of workers into occupations pioneered in Roy (1951).
[11] For the contributions of immigrants to the US educational system, see Faust (1916).

amongst servants to mainland American colonies and those to West Indies, respectively.[12]

The final mechanism connects to historical immigration leading to the emergence of "ethnic enclaves" (Altonji and Card, 1991; Card, 2001). The location choice of prospective immigrants is heavily dictated by pre-existing immigrant settlements in a certain location. This can be corroborated by the diverse historical ancestry compositions across regions of the US such as the prevalence of Mexican ancestry in the Southwest, Italian ancestry in the Northeast, and German ancestry in the Midwest (Abramitzky and Boustan, 2017). These "social networks" are another contributing factor in attracting immigrants to specific areas. The existence of these networks is pivotal to immigrants arriving in the succeeding periods, since they might receive job referrals or acquire information on local economic conditions from their fellow countrymen (Munshi, 2003).

3.3 The Construction of Valid Instruments

Motivated by the mechanisms highlighted above, I generate the quasi-random variation in ancestry relying on historical settlement patterns at the granular level within an enhanced version of the shift-share instrumental variable (SSIV) approach. Using this predicted ancestry, I isolate plausibly exogenous working-age population with varying levels of educational attainment. These instruments correct for the previously identified endogeneity problem in human capital.

3.3.1 Predicting Ancestry Stocks

First, I predict the total number of individuals with ancestry o (*i.e.*, country), who resides in d (*i.e.*, county) at time t (*i.e.*, census year), $\mathcal{A}_{o,d,t}$, by following the procedure developed in Burchardi et al. (2019).[13] This method predicts ancestry stocks via the interactions of the historical "push" and "pull" factors along with a leave-out strategy. This helps identify variations in $\mathcal{A}_{o,d,t}$ that are plausibly exogenous to both county- and country-county-specific factors. Thus, I generate granular quasi-random variation in ancestry stocks at the county level from the simultaneous joint forces of the "push" and "pull" factors, by estimating the following equation:

[12] Valencia Caicedo (2019) documents the role of eighteenth century Jesuit presence in places across South America. Rocha et al. (2017) demonstrate the importance of historical immigrant settlement policies in Brazilian municipalities in underpinning current higher educational attainment. Droller (2018) finds that counties with higher shares of historical European immigrants have higher GDP per capita today due to better educational attainment amongst immigrants in Argentina.

[13] Unlike Burchardi et al. (2019), who predict county-level ancestries based on reported identities that evolve as social constructs, I employ an ancestry measure following Fulford et al. (2019) that reflects genealogical lineage by incorporating internal migration, intergenerational transmission, and ancestry-specific population growth.

$$\mathcal{A}_{o,d,t} = \alpha_{o,r(d)} + \alpha_{c(o),d} + X'_{o,d}\phi + \sum_{\tau=1860}^{t} \gamma_{r(d),\tau} \left(I_{o,-r(d),\tau} \frac{I_{-c(o),d,\tau}}{I_{-c(o),\tau}} \right) + u_{o,d,t}, \tag{7}$$

where $\mathcal{A}_{o,d,t}$ is the total number of individuals with ancestry o settled in county d at time t. $I_{o,-r(d),\tau}$ is the "push" factor that is the total number of immigrants from o at time τ settled in counties *outside* of the census division where d is located. $\frac{I_{-c(o),d,\tau}}{I_{-c(o),\tau}}$ is the "pull" factor: that is, the denominator shows the total number of immigrants in the United States at time τ coming from *outside* of the continent where o is located, whereas the numerator displays the total number of immigrants in location d at time τ, again from *outside* of the continent where o is located.[14] $\alpha_{o,r(d)}$ and $\alpha_{c(o),d}$ represent a series of country $\times$ destination census division and continent of origin $\times$ destination county fixed effects. $X'_{o,d}$ is a vector of the time-invariant distance and difference in latitude controls between o and d. My identification strategy in this step allows me to estimate equation (7) separately for each decennial census year between 1970 and 2010 utilizing all origin countries and all destination counties in my sample.[15] Finally, I obtain predicted ancestry stocks as follows:

$$\hat{\mathcal{A}}_{o,d,t} = \sum_{\tau=1850}^{t} \hat{\gamma}_{\tau} \left(I_{o,-r(d),\tau} \frac{I_{-c(o),d,\tau}}{I_{-c(o),\tau}} \right), \tag{8}$$

where $\hat{\gamma}_{\tau}$ are the coefficients that I estimate from equation (7).[16]

The following example provides some intuition. Using this approach, I predict a large stock of ancestry from Mexico relative to other ancestry groups from the Americas to Los Angeles county in the Pacific census division relative to other census divisions in the West in 1970 if the following occur. First, a considerably large number of Mexicans migrate to the United States in 1910, into census divisions excluding the Pacific census division in the West. Second, Los Angeles county is an appealing destination to foreign immigrants from *any* origins excluding the Americas in 1910. This method ensures that Mexican migrants with particular skill sets are left out to mitigate endogenous sorting into destinations.

[14] Appendix Table C.1 displays the allocation of states to census divisions.

[15] The sample comprises 146 origin countries and 3,141 destination counties harmonized using the 1990-level country and county definitions and borders.

[16] Note that I take the residuals of the interactions of push and pull factors with respect to all fixed effects and controls included in equation (7) similar to Terry et al. (2021). This procedure is necessary to conclude that $\hat{\mathcal{A}}_{o,d,t}$ hinges solely on local variations in relative predicted ancestry stocks by eliminating any remaining confounding factors between large and small counties as well as countries. This addresses the first threat, reverse causality, highlighted in the text.

On the other hand, had I used observed (realized) ancestry stocks in "push" and "pull" factors without the "leave-out" element (as in the canonical shift-share approach), this would induce a spurious correlation between the local ancestry composition and the outcomes owing to the persistence in local productivity or amenity shocks. Such shocks attract immigrants with different human-capital mixes, thereby inducing non-random destination choices.

3.3.2 Isolating Exogenous Population with Varying Levels of Educational Attainment

After predicting ancestry stocks, I proceed to the second step to isolate human-capital-specific working-age population by county over the 1970-2010 period, which is a key contribution in my identification strategy. To this end, I predict the total number of individuals with varying levels of educational attainment, corresponding to low-, medium-, and high-skill types, in US counties by utilizing the quasi-random variation in ancestry estimated in the previous step. Therefore, the estimating equation takes the following form:

$$N_{e,d,t} = \alpha_{s(d)} + \alpha_t + \sum_o \mu_{e,o} \cdot \hat{\mathcal{A}}_{o,d,t} + \nu_{e,d,t}, \tag{9}$$

where $N_{e,d,t}$ is the total number of individuals with educational attainment e residing in county d at time t. $\alpha_{s(d)}$ and α_t represent state and time fixed effects, respectively. $\hat{\mathcal{A}}_{o,d,t}$ is the predicted ancestry stock obtained in equation (8). I estimate equation 9 for individuals with each educational attainment separately. Finally, I obtain the main instruments for working-age population share with human capital level e in county d at time t, $n_{e,d,t}$, in equation (6) as follows:

$$\hat{N}_{e,d,t} = \sum_o \hat{\mu}_{e,o} \cdot \hat{\mathcal{A}}_{o,d,t}. \tag{10}$$

These instruments address endogeneity in the skill–specific local working-age population by exploiting exogenous variation from previously identified channels that shape local human capital. The channels are (*i*) the heterogeneity in immigrants' skills, comparative advantages, along with occupational specialization, (*ii*) the transmission of human capital from historical immigration to local population via persistence, and (*iii*) the formation of of "ethnic enclaves" and "social networks" induced by immigration.[17]

[17] Appendix Figures D.1, D.2, and D.3 illustrate the bivariate maps of endogenous human-capital-specific working-age population shares, and Appendix Figures D.4, D.5, and D.6 display the bivariate maps of exogenous skill-specific working-age population across US counties in 2010.

A potential concern is that the instruments might capture the persistence of ethnic or cultural traits, or simply reflect geographic patterns of historical immigrant settlement, rather than differences in human capital. In this context, however, the interpretation is distinct: immigrant groups arrived at different times and settled in different locations with markedly heterogeneous levels of education, literacy, and occupational skills. For instance, Borjas and Katz (2007) document that Mexican immigrants historically lagged in educational attainment relative to other groups, contributing to persistent wage gaps. Similarly, 45% of Southern European and 33% of Eastern European arrivals aged 14 and older were illiterate in 1899–1910 (Goldin, 1994). Linked census records further show systematic, country-specific differences in initial occupation-based earnings, with Southern and Eastern Europeans starting lower and Germans and Nordics closer to natives, consistent with differential skill endowments rather than cultural assimilation (Abramitzky et al., 2014). These historical patterns indicate that predicted ancestry captures the long-run transmission of human capital embedded in early settler populations, as opposed to cultural persistence or spatial clustering.[18]

Identifying Assumption. The advanced version of the SSIV design used in this study in the spirit of Burchardi et al. (2019) addresses one of the key critiques for the conventional Card-type instruments, which is serial correlation or autocorrelation among past and contemporaneous immigration stocks (Jaeger et al., 2018). As a remedy, this new method introduces the "leave-out" component for both "push" and "pull" factors as explained in detail above. Hence, a sufficient condition for the exogeneity of the constructed instruments rests on the assumption that the predicted immigration stock, $\hat{\mathcal{A}}_{o,d,t}$, is exogenous in equation (6). Formally,

$$\left(I_{o,-r(d),\tau} \frac{I_{-c(o),d,\tau}}{I_{-c(o),\tau}} \right) \perp \varepsilon_{e,d,t} \quad \forall o, \tau \leq t. \tag{11}$$

In other words, any confounding factors that drive changes in the industrial skill composition of a given county at time t ($\varepsilon_{e,d,t}$) do not systematically correlate with historical immigration from a certain origin to other census divisions at time τ ($I_{o,-r(d),\tau}$) interacted with the share of immigrants from other continents in that county in the same year $\left(\frac{I_{-c(o),d,\tau}}{I_{-c(o),\tau}}\right)$. The exogeneity condition for these instruments aligns with the "shock orthogonality" condition outlined in Goldsmith-Pinkham

[18]While cultural persistence may matter in other settings (Alesina and Tabellini, 2024; Gagliarducci and Tabellini, 2025), the channel here operates through inherited human capital differences. To guard against the possibility that results are driven by the dominance of any one ancestry, I exclude the five largest groups (English, German, Irish, Italian, and Mexican) in Appendix Table C.13. The coefficients remain stable in magnitude and significance, reinforcing the view that the mechanism reflects human capital transmission rather than cultural or locational persistence.

et al. (2020).[19]

The following example builds intuition for the plausibility of this assumption. Suppose that agriculture, a low-skill (LS) industry, experiences a productivity shock in Kern county at time t ($\varepsilon_{LS,Kern,t}$), which draws in Mexican immigrants with a comparative advantage in agriculture ($I_{Mexico,Kern,t}$). A sufficient restriction for exogeneity to hold is that this confounding shock together with any other country-county-specific factors that attract immigrants and shape changes in low-skill industries in that county does not affect the predicted ancestry from Mexico in Kern county at time t, $\hat{\mathcal{A}}_{o,d,t}$. The violation of this restriction would arise in the case, whereby the confounding shock: (i) systematically affected the settlement patterns of many immigrants excluding the Americas such as Filipino immigrants, and (ii) drew a significant number of Mexicans to counties outside of the Pacific region such as Lancaster county in Pennsylvania. I probe this possibility using a variant of the "leave-out push-pull" instruments in Section 6 addressing another problem emphasized in Adao et al. (2019).

3.4 Instrument Validation Exercises and Instrument Performance

To assess the validity of the identification strategy, I use the following reduced-form specification:

$$\Delta y^{std}_{e,d,t'} = \alpha_{s(d)} + \alpha_t + \hat{\boldsymbol{N}}^{\text{std}}_{e,d,t}\boldsymbol{\psi}_e + \epsilon_{e,d,t}, \tag{12}$$

where $\Delta y^{std}_{e,d,t'}$ is a standardized measure of the change in employment share of each industry-skill type at time t' (*i.e.*, between 1940 and 1960), and $\hat{\boldsymbol{N}}^{\text{std}}_{e,d,t}$ is a standardized measure of a vector of exogenous population of medium- and high-skill types over the 1970-2010 period. The regressions also include state ($\alpha_{s(d)}$) and time (α_t) fixed effects.

The rationale behind this specification is to examine whether the vector of human-capital-specific exogenous population, $\hat{\boldsymbol{N}}^{\text{std}}_{e,d,t}$, is correlated with pre-trends.[20] If the correlations between them exist, then they would point to the likely violations of the identifying assumption laid out above. In this light, $\boldsymbol{\psi}_e$ provide balance tests on pre-trends that probe whether these violations are present.

Figure 1 displays the results of these balance tests. All estimated coefficients are close to zero

[19] Note that this is a sufficient, but not a necessary, condition for instrument validity within the SSIV framework. For both conditions, see Borusyak et al. (2022).

[20] Since my design leverages persistence in local human capital, balance is best assessed using pre-trends rather than pre-period levels. In persistent settings, correlations between pre-study levels and the instrument are mechanical. The relevant orthogonality restriction is that the instrument does not predict pre-period changes conditional on state and time fixed effects.

and statistically insignificant, indicating the absence of systematic pre-trends in the relationship between the instruments and the outcome conditional on state and time fixed effects between 1940 and 1960. These findings provide reassuring evidence that the ancestry-driven shift in human capital of local population is uncorrelated with short- and long-run economic outcomes, an issue raised in Jaeger et al. (2018).

Figure 2 displays binned scatter plots, whereby exogenous working-age population of each human capital level is plotted against endogenous working-age population of each human capital level. Conditional on state and time fixed effects, binned scatter plots demonstrate that these variables are well-aligned.

4 Data

4.1 Sources

Ancestry. I use county-level data on ancestry stocks for 1970–2010 from Fulford et al. (2019). I briefly summarize how they construct expected ancestry for each person and aggregate it into population estimates to map the geographic distribution of ancestry. They build ancestry distributions at both national and county levels using census micro-samples for each decade from 1850 to 2010. The expected ancestry for first-generation immigrants is based on their countries of birth. For their US-born children (*i.e.*, second-generation immigrants), ancestry depends on the parents' birthplaces. Additionally, ancestry is traced through parents' birth states or respondents' own counties of residence, capturing internal migration and inter-generational transmission. For later generations, expected ancestry shares are constructed from past ancestry shares, reflecting how they evolve as internal migrants relocate and pass their ancestries to their children. In my sample, I harmonize all counties and county groups to 1990 boundaries for consistency across decennial census years.[21]

This approach emphasizes genealogical reconstruction by accounting for internal migration, generational transmission, and differential group growth. In contrast, studies using self-reported ancestry in the census since 1980 capture the changing nature of ethnic identity as a social construct.

Immigration. I utilize data on immigration obtained from the individual decadal census files of the Integrated Public Use of Microdata Series (IPUMS USA and NHGIS: Ruggles et al., 2022;

[21] The crosswalk files come from Burchardi et al. (2019) and Eckert et al. (2020b). See Appendix B for additional details.

Manson et al., 2022) samples from 1850-2000 as well as the 2006-2010 five-year sample of the American Community Survey (ACS). I weigh observations by the person weights in these sources. My dataset on immigration covers 3,141 US counties, 146 foreign countries, and 15 census waves. I provide additional details on the construction of immigration stock data in subsection B.1 of Data Appendix.

Working-Age Population, Employment, and Establishments. I acquire data on working-age population and employment at the county level using the individual census rounds from 1970-2000, and the 2006-2010 five-year waves. I define working-age population as individuals aged 16 to 64. As before, I weigh observations by the person weights. I obtain data on the number of establishments at the county level from the 1970-2010 extracts of the County Business Patterns (CBP) data published by the Census Bureau. I supplement these data with those compiled in Eckert et al. (2020a, 2022). The Census Bureau defines an establishment as a single physical location where services and production activities are undertaken.[22] I provide additional details on the constructions of working-age population, employment, and establishment data in subsections B.2, B.3, and B.4 of Data Appendix, respectively.

4.2 Skills, Human Capital, and Tradability Classification

A central component of my analysis is to allocate industries into skill cells. For this purpose, I construct the industry-based continuous skill measure in the following steps. First, I create an education variable and assign values to it based on the educational attainment (years of education) of each individual. Second, I aggregate the 1990-level census industry codes to a balanced panel of industries utilizing an industry concordance generated in Autor et al. (2019). Finally, I calculate mean education for workers employed by each industry in the balanced panel in each census year. I define industries with average years of education corresponding to a high school degree and less as low skill, those with average education corresponding to some college degree as medium skill, and those with mean education comparable to college degree and more as high skill. Appendix Table C.2 presents the industry-skill classification.

Classifying working-age individuals by human capital is identical to the procedure described above. Following the task-based framework (Autor et al., 2003), I classify labor into low-, medium-, and high-skill categories. High-skill workers perform abstract, capital-complementary tasks, while medium-skill workers are more concentrated in routine occupations subject to substitution (Goos

[22] An establishment is not necessarily identical to a firm or enterprise, which may own one or more establishments.

and Manning, 2007). Collapsing medium and high-skill into a single group would mask this heterogeneity.

Since tradability shapes how local shocks transmit, with nontradables absorbing local demand and tradables facing global competition, I also group industries into tradable and nontradable sectors. I classify agriculture, mining, and manufacturing industries as tradable, whereas I categorize all service industries as nontradable similar to the classification scheme in Burstein et al. (2020). Appendix Tables C.3, C.4, C.5, and C.6 display tradable and nontradable industries along with their skill types.

4.3 Summary Statistics

Table 1 reports summary statistics. Panel A shows substantial variation in the skill distribution of working-age population shares and in mean education levels across counties. Most working-age individuals hold a high school degree or less, a quarter have some college education, and the remainder possess a college degree or higher. Panel B presents employment and establishment shares, as well as the mean education of workers across industries, used in the baseline and tradability analyses.

The distribution of employment and establishment shares in medium-skill industries is similar; however, a larger share of workers are employed in low-skill industries, while establishments are more concentrated in high-skill industries. Employment and establishment shares in low-skill industries are comparable across tradable and nontradable sectors, but most workers and establishments are concentrated in medium- and high-skill nontradable industries. Although the mean education of medium-skill workers is similar across both sectors, workers in low- and high-skill nontradable industries tend to have slightly more schooling than those in corresponding tradable industries.

5 Results

In this section, I present my empirical findings by exploiting exogenous working-age population of each human capital level to explore the causal impacts of population shares on employment and establishment shares of industries across counties in the baseline and tradability analyses. Furthermore, I implement a heterogeneity analysis across major industries along with establishment size domains and probe the robustness of my baseline estimates on a series of checks. All regressions include state and time fixed effects to absorb time-invariant differences across states and common aggregate shocks, allowing identification to come from within-state, cross-county variation over time. Standard errors are clustered at the county level to control for serial correlation within counties across years in all specifications, which is the level at which the treatment varies. Clustering at a higher level (*e.g.*, by state) would be overly conservative given the limited number of states and the sub-state variation exploited in the analysis.

5.1 Baseline Analysis: Employment and Establishment Shares

Prior to the causal results, I report OLS estimates in Table 2 presenting the associations between population shares and employment shares in Panel A and establishment shares in Panel B based on estimating equation (6). The OLS estimates of medium- and high-skill population shares on employment and establishment shares of all industry-skill types are statistically significant. These results, however, reflect only correlations between the endogenous distribution of human capital and the industrial skill composition of counties rather than causal effects.

Table 3 reports IV estimates for the impacts of population shares on employment shares in Panel A and on establishment shares in Panel B. Endogenous medium- and high-skill population shares are instrumented with exogenous population of corresponding human capital levels as described in Section 3. All Anderson-Rubin Wald F-test p-values for the statistically significant point estimates are around 0%, suggesting that endogenous population shares are jointly significant, and that my instruments, exogenous population of given human capital levels, are valid. The table also presents Sanderson-Windmeijer first-stage F-statistics for each endogenous variable, which are sufficiently greater than the conventional threshold.[23]

In particular, I find that, relative to low-skill population share, a one-percentage-point increase in medium- and high-skill population shares leads to a decrease in employment share of low-skill industries by 0.39 and 0.32 percentage points, respectively. By contrast, this increase results in a rise in employment share of high-skill industries by corresponding 0.18 and 0.54 percentage points. Additionally, a one-percentage-point rise in medium-skill population share boosts employment share of medium-skill industries by 0.21 percentage points, while a rise in high-skill population by the same magnitude reduces employment share of those industries by 0.23 percentage points. One can observe similar patterns across the estimates for establishment shares. Specifically, establishment share of low-skill industries drops by 0.75 and 0.11 percentage points as

[23]The Sanderson–Windmeijer F-statistics provide a measure of instrument relevance in models with multiple endogenous regressors, testing the joint significance of the excluded instruments for each endogenous variable conditional on the others (Sanderson and Windmeijer, 2016).

medium- and high-skill population shares grow by a one percentage point. Conversely, establishment share of high-skill industries increases by 0.96 and 0.33 percentage points in response to a one-percentage-point increase in medium- and high-skill population shares. A one-percentage-point increase in high-skill population share leads to a 0.18 percentage points decline in establishment share of medium-skill industries.

Taken together, these empirical findings are in line with the model predictions introduced in Section 2. First, both employment and establishment shares of high-skill industries respond positively to a rise in medium- and high-skill population shares. This is equivalent to white-collar firms increasing their labor demand share for high-skill labor supply share due to their comparative advantage in employing high-skill labor more intensively.[24] Second, a rise in medium- and high-skill population shares reduces employment and establishment shares of low-skill industries. This corresponds to blue-collar firms having comparative advantage in using low-skill labor more intensively while employing high-skill labor less intensively. The relative magnitudes of these effects are consistent with a moderate elasticity of substitution between high- and low-skill labor, indicating that firms can partially substitute across skill types but that imperfect substitutability amplifies reallocation across industries.

To compare the establishment estimates, Sequeira et al. (2020) show that the aggregate establishment share rose in response to higher average immigrant shares during the Age of Mass Migration. Moretti (2004) finds that increases in the local share of college-educated workers lead to larger productivity gains for firms in cities. In a related study, Shapiro (2006) shows that most of the growth in urban employment of college graduates is explained by the improvement in productivity. Consistent with these findings, my estimates demonstrate that exogenous increases in county human capital reallocate establishments towards high-skill industries, raising the demand for high-skill labor. While these studies focus on productivity or aggregate employment effects, my analysis emphasizes compositional changes in industry structure, identifying how shifts in human capital supply reshape the distribution of establishments across industries differentiated by skill intensity.

[24] Although the theoretical framework introduces only two types of labor (low and high skills), the model predictions can be extended to a setting that includes a third type, medium skill, in addition to the existing two.

5.2 Tradability Analysis: Employment and Establishment Shares

The theory and empirical evidence in the existing literature suggest that local labor supply shocks generate heterogenous adjustments across firms, occupations, and industries (Acemoglu, 2011; Diamond, 2016; Burstein et al., 2020). Motivated by this, I distinguish between tradable and nontradable industries to capture heterogeneous adjustment margins. Nontradable industries are inherently local, with employment responding to shifts in local demand and supply, while tradable industries can reallocate across regions or absorb shocks through domestic and global markets. Failing to separate the two obscures these mechanisms.[25] In this subsection, I present my results pertinent to the effects of working-age population shares of all human capital levels, instrumented similarly as before, on the outcomes of interest in the tradable and nontradable sectors. I include state and time fixed effects and cluster standard errors by county in all specifications.

Table 4 reports the regression results of how population shares affect employment shares in the tradable and nontradable sectors in Panels A and B, respectively. The positive baseline effects originate in the nontradable sector, whereas the tradable sector absorbs the adverse impacts. Higher medium- and high-skill population shares lead to substantial declines in employment share of low-skill tradable industries, but sizable increases in employment share of nontradable industries of the corresponding skill types. In addition, a larger high-skill population share reduces employment share of medium-skill tradable industries, while modestly raising high-skill tradable employment share. Finally, an increase in high-skill population share has a small negative effect on low-skill nontradable employment share.

Table 5 presents the effects of population shares on establishment shares in the tradable and nontradable sectors in Panel A and Panel B, respectively. Relative to the baseline, the negative effects are concentrated in the tradable sector, while the positive effects arise in the nontradable sector. Higher medium- and high-skill population shares substantially reduce establishment share of low-skill tradable industries, and moderately lower share of high-skill tradable industries. By contrast, increases in medium- and high-skill population shares produce large gains in establishment share of high-skill nontradable industries. An increase in the high-skill population share slightly reduces the shares of high-skill tradable and low-skill nontradable industries. The positive effect of a larger medium-skill population share on low-skill nontradable establishments and the negative effect of a larger high-skill population share on medium-skill nontradable establishments are similar

[25] This distinction is standard in the literature on local labor markets and trade (Autor and Dorn, 2013; Mian and Sufi, 2014).

in magnitude, while the adverse effect on low-skill nontradable establishments is modest.

Overall, my findings show that nontradable industries expand when local medium- and high-skill labor rises, consistent with nontradables being locally consumed and adjusting to stronger local demand (Moretti, 2010; Mian and Sufi, 2014). In contrast, tradable industries contract, reflecting exposure to national and international competition and the fact that local labor-supply shocks are absorbed through cross-regional reallocation or reduced local competitiveness (Autor and Dorn, 2013). In short, positive adjustments occur in nontradables, where local demand absorbs and amplifies the gains from a more skilled workforce. However, negative adjustments concentrate in tradables, where competitive pressures transmit the costs of relative scarcity in low-skill labor. This pattern is consistent with the mechanism: industries differ in productivity across labor types, and whether those shifts translate into local expansion or external adjustment depends on sector tradability.

5.3 Employment Heterogeneity by Nativity

I isolate exogenous, human-capital-specific shifts in local population using quasi-random variation in the ancestral composition of counties. The identification relies on the inter-generational transmission of human capital and knowledge discussed in Section 3, and it captures the long-run variation in embedded human capital across ancestry groups. Accordingly, the estimates can be interpreted as reduced-form effects of ancestry-driven shifts in local human capital endowments.

While the design of my instruments does not directly permit separating individuals by nativity, I next assess whether these ancestry-induced variation in local human capital affect both native and immigrant workers. Appendix Table C.7 reports the results for native employment shares in Panel A and immigrant employment shares in Panel B. The coefficients indicate that both groups respond, but the effects are more pronounced for immigrants. This pattern suggests that local reallocation occurs more strongly along immigrant-intensive margins with higher adjustment elasticity, consistent with greater mobility, stronger industry networks, and faster occupational transitions among immigrant workers. Although the instrument derives from historical immigration patterns, the larger immigrant response reflects contemporaneous labor-market adjustment rather than mechanical relabeling of naturalized immigrants as "native."

5.4 Employment and Establishment Heterogeneity by Industry

The existing literature documents that historical events can generate both short-run and long-run changes in sectoral structure, industrialization, and local labor markets (Peters, 2022; Droller, 2018; Rocha et al., 2017). Guided by this evidence, I examine how exogenous shifts in local human capital reallocate employment and establishment shares across major industries. In this subsection, I report industry-level results without disaggregating by skill.[26] All specifications include state and year fixed effects, and standard errors are clustered by county.

Appendix Figure C.1 plots the coefficients for employment in Panel A and establishment shares in Panel B. The findings reveal that in medium- and high-skill population shares raise the services share on both margins and reduce the shares of mining and manufacturing (and agriculture for employment). These patterns mirror the tradability results: nontradables expand with local skill and demand, while tradables contract or reallocate. This sectoral reallocation aligns with US evidence that local demand and skill shifts tend to pull activity towards services and other nontradables and away from manufacturing (Moretti, 2010; Autor and Dorn, 2013; Notowidigdo, 2020; Mian and Sufi, 2014). In contrast to Peters (2022), who studies a German refugee shock that raises manufacturing and lowers agriculture, my estimates point to a services-intensive adjustment consistent with the nontradable expansion documented above.

5.5 Establishment Heterogeneity by Size

In this subsection, I examine heterogeneity by establishment size for all industries and separately for tradable and nontradable sectors. The goal is to identify which establishments are most affected by the ancestry-driven exogenous shift in local human capital, and whether this translates into additional industry–skill reallocation across size classes. Following the literature, I define small establishments as employing fewer than 20 workers, medium as 20–500, and large as more than 500 workers.[27] As before, I instrument endogenous skill-specific population shares with exogenous populations at corresponding human-capital levels. All specifications include state and year fixed effects, and standard errors are clustered by county.

Establishments in All Industries. Appendix Table C.8 reports the regression results showing how local population shares affect the distribution of small, medium, and large establishments

[26] Given limitations in the coverage of agricultural establishments in the CBP over the study period, I exclude agriculture from the analysis.

[27] For a similar classification, see Olney (2013).

across all industriy-skill types in Panels A, B, and C, respectively. The estimates demonstrate that the negative impacts of medium- and high-skill population shares are primarily absorbed by small establishments in low- and medium-skill industries, followed by medium-sized establishments in low-skill industries. However, medium-sized establishments in medium- and high-skill industries capture the positive effects of these population shifts, with stronger responses in medium-skill sectors. The estimates for large establishments are less precise, suggesting that they are less responsive to local human capital shocks. Taken together, the results show that establishment responses to human capital shocks are heterogeneous across both firm size and industry-skill types. Consistent with evidence from firm-dynamics models emphasizing differential adjustment costs and entry–exit responses across establishment sizes (Hopenhayn, 1992; Haltiwanger et al., 2013), the reallocation patterns observed here suggest that small and medium-sized establishments drive most of the compositional adjustment following local human capital shocks.

Establishments in the Tradable and Nontradable Sectors. Appendix Tables C.9 and C.10 present the responses of small, medium, and large establishment shares in the tradable and nontradable sectors to population shares in Panels A, B, and C, respectively. The estimates in Appendix Table C.9 for the tradable sector show that the negative effects of population shares on establishment share of low-skill industries reported in the baseline findings are almost entirely driven by small establishments, while the coefficients for medium and large establishments are imprecisely estimated. The estimated coefficients in Appendix Table C.10 for the nontradable sector reveal that the adverse impacts of high-skill population shares on low- and medium-skill industries are similarly concentrated among small establishments. In contrast, the positive effects of population shares on establishment shares in medium-skill industries are primarily accounted for by medium-sized establishments. These patterns align with the evidence mentioned above that smaller establishments are the most responsive margin of adjustment to local labor market shocks, while sectoral reallocation between tradable and nontradable activities serves as a key channel through which local economies absorb such shifts.

6 Robustness

I implement a battery of sensitivity checks and show that my results are robust to (*i*) an alternative construction of the shift-share instruments, (*ii*) various sample selections, (*iii*) an alternative tradability classification, (*iv*) placebo exercises, and (*v*) randomization tests.

Alternative Instruments/Exogenous Population of Different Human Capital Levels. To assess the persistence and robustness of the main results, I examine whether ancestry-induced exogenous shifts in local human capital have had lasting effects on the industrial skill composition across US counties. This exercise sheds light on how historical shocks continue to shape local labor market structures. I predict ancestry stocks from 1970 to 2010 while restricting the push and pull factors to the 1850–1900 and 1850–1950 periods in separate specifications.[28] Using these predicted ancestry stocks, I isolate human-capital-specific exogenous local population shares following the baseline approach. Appendix Tables C.11 and C.12 present the corresponding regression results for employment and establishment shares in Panels A and B, respectively. The coefficient estimates across both periods closely resemble those in the baseline analysis, suggesting that the results are not sensitive to the historical window used to construct the instruments.

Overall, these findings point to two conclusions. First, historical immigrant settlement patterns have left lasting imprints on local labor markets through long-term human capital transmission. Second, the main results remain robust to alternative definitions of the immigration period. The comparability of the coefficients across these windows shows that the relationship between past immigrant settlement and today's industrial structure reflects persistent, deep-rooted channels of human capital transmission rather than artifacts of a particular historical period.

Various Sample Selections. Another concern is that the baseline estimates could be driven by the largest ancestry groups or counties that historically received the most immigrants. To ensure robustness to their exclusion, I conduct two exercises. First, I omit the five largest ancestry groups when isolating the human-capital-specific exogenous population and re-estimate the baseline regressions. Second, I exclude the five largest immigrant-receiving counties when constructing the instruments and repeat the analysis.[29] Appendix Tables C.13 and C.14 present the corresponding regression results. In both cases, the estimated coefficients remain highly consistent with the baseline results.

Alternative Tradability Classification. Next, I employ an alternative measure of tradability by re-classifying industries into tradable and nontradable sectors. The goal is to assess the sensitivity of the baseline estimates to the choice of classification. I adopt the industry tradability measure proposed in Mian and Sufi (2014), which is based on geographical Herfindahl–Hirschman Indices

[28] These periods roughly correspond to the Age of Mass Migration.

[29] The five largest ancestry groups are English, German, Irish, Italian, and Mexican, while the five largest destinations are Los Angeles, Cook, Harris, Miami-Dade, and Kings counties.

(HHIs); industries that are more geographically concentrated are considered more tradable.[30] Appendix Tables C.15 and C.16 report the effects on employment and establishment shares using this alternative measure. On the whole, the estimated coefficients remain broadly consistent with the baseline results, indicating that the findings are not sensitive to the tradability classification.

7 Conclusion

There are sizable differences in the skill distribution of industries across the United States. These differences may arise from technological change, job polarization, and local demand shocks. However, the role of human capital in shaping this distribution remains less understood. This paper examines how ancestry-induced exogenous shifts in the human capital composition of local populations affect the industrial skill structure of employment and establishment shares across US counties.

To address endogeneity in human capital, I employ a two-step empirical strategy. First, I predict ancestry stocks using a leave-out push–pull approach within a shift–share instrumental variable framework. Second, I exploit this quasi-random variation to isolate exogenous shifts in the skill-specific local working-age population. A simple CES model with two firm types, each specializing in either high- or low-skill labor, guides the empirical analysis and predicts differential labor demand responses to changes in the supply of skill groups.

The results show that increases in medium- and high-skill population shares raise employment and establishment shares of high-skill industries while reducing those of low-skill industries. These patterns are consistent with firms adjusting production towards sectors that use the increasingly abundant skill type more intensively. Moreover, the nontradable sector absorbs most of the positive effects, reflecting stronger local demand for skill-intensive services, while the tradable sector experiences relative contraction due to declining competitiveness in low-skill activities.

The findings carry important implications for workforce and regional development policy. Efforts to attract firms or foster local industrial growth should account for the composition of the local labor force rather than focusing solely on its magnitude. Policies that strengthen the alignment between local skill supply and industry demand (for instance, through targeted training, reskilling programs, or regional education initiatives) can enhance employment growth and encourage firm formation. By contrast, neglecting mismatches between workforce composition and firm needs may

[30] See Appendix Table I in the online appendix of their paper for the full classification scheme.

amplify spatial disparities in economic opportunity.

Future research could build on these results by examining how skill-driven reallocation across industries affects firm productivity, wage inequality, or local innovation capacity. Such extensions would help clarify whether the reorganization of industries in response to shifts in human capital merely redistributes activity or also raises aggregate efficiency and long-run growth potential.

References

Abramitzky, R. and Boustan, L. (2017). Immigration in american economic history. *Journal of economic literature*, 55(4):1311–1345. [Cited on page 11.]

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. [Cited on pages 2 and 10.]

Abramitzky, R., Boustan, L. P., and Eriksson, K. (2014). A nation of immigrants: Assimilation and economic outcomes in the age of mass migration. *Journal of Political Economy*, 122(3):467–506. [Cited on pages 2, 10, and 14.]

Abramitzky, R. and Braggion, F. (2006). Migration and human capital: self-selection of indentured servants to the americas. *The Journal of Economic History*, 66(4):882–905. [Cited on pages 2, 5, and 10.]

Acemoglu, D. (2011). Skills, tasks and technologies: Implications for employment and earnings. [Cited on pages 1 and 21.]

Adao, R., Kolesár, M., and Morales, E. (2019). Shift-share designs: Theory and inference. *The Quarterly Journal of Economics*, 134(4):1949–2010. [Cited on pages 2, 5, and 15.]

Alesina, A. and Tabellini, M. (2024). The political effects of immigration: Culture or economics? *Journal of Economic Literature*, 62(1):5–46. [Cited on page 14.]

Allen, T. and Donaldson, D. (2020). Persistence and path dependence in the spatial economy. Technical report, National Bureau of Economic Research. [Cited on pages 2, 5, and 10.]

Altonji, J. G. and Card, D. (1991). The effects of immigration on the labor market outcomes of less-skilled natives. In *Immigration, trade, and the labor market*, pages 201–234. University of Chicago Press. [Cited on pages 2, 3, 5, and 11.]

Amuedo-Dorantes, C., Arenas Arroyo, E., Mahajan, P., and Schmidpeter, B. (2023). Low-wage jobs, foreign-born workers, and firm performance. [Cited on page 5.]

Autor, D., Dorn, D., and Hanson, G. (2019). When work disappears: Manufacturing decline and the falling marriage market value of young men. *American Economic Review: Insights*, 1(2):161–178. [Cited on pages 17, 42, and 43.]

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the us labor market. *American economic review*, 103(5):1553–1597. [Cited on pages 1, 4, 21, 22, and 23.]

Autor, D. H., Katz, L. F., and Krueger, A. B. (1998). Computing inequality: have computers changed the labor market? *The Quarterly journal of economics*, 113(4):1169–1213. [Cited on pages 4 and 6.]

Autor, D. H., Levy, F., and Murnane, R. J. (2003). The skill content of recent technological change: An empirical exploration. *The Quarterly journal of economics*, 118(4):1279–1333. [Cited on page 17.]

Bartik, T. J. (1991). Who benefits from state and local economic development policies? [Cited on page 1.]

Bazzi, S., Fiszbein, M., and Gebresilasse, M. (2020). Frontier culture: The roots and persistence of "rugged individualism" in the united states. *Econometrica*, 88(6):2329–2368. [Cited on page 5.]

Bergquist, J. M. (2007). *Daily Life in Immigrant America, 1820–1870*. Greenwood Press, Westport, CT. [Cited on page 10.]

Bleakley, H. and Lin, J. (2012). Portage and path dependence. *The quarterly journal of economics*, 127(2):587–644. [Cited on page 5.]

Boberg-Fazlić, N. and Sharp, P. (2024). Immigrant communities and knowledge spillovers: Danish americans and the development of the dairy industry in the united states. *American Economic Journal: Macroeconomics*, 16(1):102–146. [Cited on pages 2 and 10.]

Borjas, G. J. (1987). Self-selection and the earnings of immigrants. *American Economic Review*, 77(4):531–553. [Cited on pages 2 and 10.]

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. *The quarterly journal of economics*, 118(4):1335–1374. [Cited on pages 3 and 5.]

Borjas, G. J. (2005). The labor-market impact of high-skill immigration. *American Economic Review*, 95(2):56–60. [Cited on page 6.]

Borjas, G. J. and Katz, L. F. (2007). The evolution of the mexican-born workforce in the united states. In *Mexican immigration to the United States*, pages 13–56. University of Chicago Press. [Cited on page 14.]

Borusyak, K., Hull, P., and Jaravel, X. (2022). Quasi-experimental shift-share research designs. *The Review of Economic Studies*, 89(1):181–213. [Cited on pages 2 and 15.]

Bound, J., Khanna, G., and Morales, N. (2017). Understanding the economic impact of the h-1b program on the united states. In *High-skilled migration to the United States and its economic consequences*, pages 109–175. University of Chicago Press. [Cited on page 5.]

Burchardi, K. B., Chaney, T., and Hassan, T. A. (2019). Migrants, ancestors, and foreign investments. *The Review of Economic Studies*, 86(4):1448–1486. [Cited on pages 2, 3, 6, 11, 14, 16, and 41.]

Burstein, A., Hanson, G., Tian, L., and Vogel, J. (2020). Tradability and the labor-market impact of immigration: Theory and evidence from the united states. *Econometrica*, 88(3):1071–1112. [Cited on pages 6, 18, 21, 33, and 46.]

Caiumi, A. and Peri, G. (2024). Immigration's effect on us wages and employment redux. Technical report, National Bureau of Economic Research. [Cited on page 6.]

Card, D. (2001). Immigrant inflows, native outflows, and the local labor market impacts of higher immigration. *Journal of Labor Economics*, 19(1):22–64. [Cited on pages 2, 5, and 11.]

Card, D. and Lemieux, T. (2001). Can falling supply explain the rising return to college for younger men? a cohort-based analysis. *The quarterly journal of economics*, 116(2):705–746. [Cited on pages 4 and 6.]

Cattaneo, M. D., Crump, R. K., Farrell, M. H., and Feng, Y. (2024). On binscatter. *American Economic Review*, 114(5):1488–1514. [Cited on page 32.]

Chiquiar, D. and Hanson, G. H. (2005). International migration, self-selection, and the distribution of wages: Evidence from mexico and the united states. *Journal of Political Economy*, 113(2):239–281. [Cited on page 10.]

Chiswick, B. R. and Taengnoi, S. (2007). Occupational choice of high-skilled immigrants in the united states. *International Migration*, 45(5):3–34. [Cited on page 10.]

Clemens, M. A. and Lewis, E. G. (2022). The effect of low-skill immigration restrictions on us firms and workers: Evidence from a randomized lottery. Technical report, National Bureau of Economic Research. [Cited on page 5.]

Clemens, M. A. and Mendola, M. (2024). Migration from developing countries: Selection, income elasticity, and simpson's paradox. *Journal of Development Economics*, 171:103359. [Cited on pages 2 and 10.]

Diamond, R. (2016). The determinants and welfare implications of us workers' diverging location choices by skill: 1980–2000. *American Economic Review*, 106(3):479–524. [Cited on pages 1 and 21.]

Doran, K., Gelber, A., and Isen, A. (2022). The effects of high-skilled immigration policy on firms: Evidence from visa lotteries. *Journal of Political Economy*, 130(10):2501–2533. [Cited on page 5.]

Droller, F. (2018). Migration, population composition and long run economic development: Evidence from settlements in the pampas. *The Economic Journal*, 128(614):2321–2352. [Cited on pages 3, 5, 11, and 23.]

Dustmann, C., Schönberg, U., and Stuhler, J. (2017). Labor supply shocks, native wages, and the adjustment of local employment. *The Quarterly Journal of Economics*, 132(1):435–483. [Cited on pages 3 and 6.]

Eckert, F., Fort, T. C., Schott, P. K., and Yang, N. J. (2020a). Imputing missing values in the us census bureau's county business patterns. Technical report, National Bureau of Economic Research. [Cited on pages 17, 42, and 43.]

Eckert, F., Gvirtz, A., Liang, J., and Peters, M. (2020b). A method to construct geographical crosswalks with an application to us counties since 1790. Technical report, National Bureau of Economic Research. [Cited on pages 16 and 41.]

Eckert, F., Lam, K.-l., Mian, A. R., Müller, K., Schwalb, R., and Sufi, A. (2022). The early county business pattern files: 1946-1974. Technical report, National Bureau of Economic Research. [Cited on pages 17, 42, and 43.]

Faust, A. B. (1916). *The Germans in the United States*. German University League, New York. [Cited on page 10.]

Foged, M. and Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, 8(2):1–34. [Cited on page 6.]

Fulford, S. L., Petkov, I., and Schiantarelli, F. (2019). Does it matter where you came from? ancestry composition and economic performance of us counties, 1850–2010. *Journal of Economic Growth*, 24(4):365–406. [Cited on pages 2, 5, 11, and 16.]

Gagliarducci, S. and Tabellini, M. (2025). Faith and assimilation: Italian immigrants in the us. *The Economic*

Journal, 135(670):1814–1851. [Cited on page 14.]

Giannone, E. (2022). Skill-biased technical change and regional convergence. *Unpublished manuscript*. Revise & resubmit at *Journal of Political Economy*. [Cited on page 1.]

Goldin, C. (1994). The political economy of immigration restriction in the united states, 1890 to 1921. In Goldin, C. and Libecap, G. D., editors, *The Regulated Economy: A Historical Approach to Political Economy*, pages 223–257. University of Chicago Press, Chicago. [Cited on pages 2, 10, and 14.]

Goldsmith-Pinkham, P., Sorkin, I., and Swift, H. (2020). Bartik instruments: What, when, why, and how. *American Economic Review*, 110(8):2586–2624. [Cited on pages 2 and 14.]

Goos, M. and Manning, A. (2007). Lousy and lovely jobs: The rising polarization of work in britain. *The review of economics and statistics*, 89(1):118–133. [Cited on page 17.]

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. [Cited on page 1.]

Haltiwanger, J., Jarmin, R. S., and Miranda, J. (2013). Who creates jobs? small versus large versus young. *Review of Economics and Statistics*, 95(2):347–361. [Cited on page 24.]

Hanson, G. H. and Liu, C. (2023). Immigration and occupational comparative advantage. *Journal of International Economics*, 145:103809. [Cited on page 10.]

Hopenhayn, H. A. (1992). Entry, exit, and firm dynamics in long run equilibrium. *Econometrica: Journal of the Econometric Society*, pages 1127–1150. [Cited on page 24.]

Hunt, J. and Gauthier-Loiselle, M. (2010). How much does immigration boost innovation? *American Economic Journal: Macroeconomics*, 2(2):31–56. [Cited on page 10.]

Jaeger, D. A., Ruist, J., and Stuhler, J. (2018). Shift-share instruments and dynamic adjustments: The case of immigration. *NBER Working Paper*, 24285. [Cited on pages 2, 14, and 16.]

Katz, L. F. and Murphy, K. M. (1992). Changes in relative wages, 1963–1987: supply and demand factors. *The quarterly journal of economics*, 107(1):35–78. [Cited on pages 4 and 6.]

Kerr, S. P., Kerr, W. R., and Lincoln, W. F. (2015). Skilled immigration and the employment structures of us firms. *Journal of Labor Economics*, 33(S1):S147–S186. [Cited on page 5.]

Khanna, G., Lee, M., et al. (2018). High-skill immigration, innovation, and creative destruction. Technical report, National Bureau of Economic Research. [Cited on page 5.]

Maestas, N., Mullen, K. J., and Powell, D. (2023). The effect of population aging on economic growth, the labor force, and productivity. *American Economic Journal: Macroeconomics*, 15(2):306–332. [Cited on page 3.]

Mahajan, P., Morales, N., Shih, K., Chen, M., and Brinatti, A. (2024). The impact of immigration on firms and workers: Insights from the h-1b lottery. [Cited on page 5.]

Malone, D. (1935). The intellectual melting-pot. *The American Scholar*, 4:444–459. [Cited on page 10.]

Manson, S., Schroeder, J., Van Riper, D., Kugler, T., and Ruggles, S. (2022). Ipums national historical geographic information system: Version 15.0. [Cited on page 17.]

Mian, A. and Sufi, A. (2014). What explains the 2007–2009 drop in employment? *Econometrica*, 82(6):2197–2223. [Cited on pages 4, 21, 22, 23, 25, 59, and 60.]

Monras, J. (2018). Economic shocks and internal migration. [Cited on page 3.]

Moretti, E. (2004). Workers' education, spillovers, and productivity: evidence from plant-level production functions. *American Economic Review*, 94(3):656–690. [Cited on pages 1, 4, and 20.]

Moretti, E. (2010). Local multipliers. *American Economic Review*, 100(2):373–377. [Cited on pages 1, 4, 22, and 23.]

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the us labor market. *The Quarterly Journal of Economics*, 118(2):549–599. [Cited on pages 2 and 11.]

Naqvi, A. (2024). Stata package: bimap. Version 2.0, Available at: `https://github.com/asjadnaqvi/stata-bimap`. [Cited on pages 61, 62, 63, 64, 65, and 66.]

Notowidigdo, M. J. (2020). The incidence of local labor demand shocks. *Journal of Labor Economics*, 38(3):687–725. [Cited on pages 1 and 23.]

Nunn, N. (2014). Historical development. *Handbook of economic growth*, 2:347–402. [Cited on pages 5 and 10.]

Olney, W. W. (2013). Immigration and firm expansion. *Journal of regional science*, 53(1):142–157. [Cited on page 23.]

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197. [Cited on pages 3 and 6.]

Pandey, M. and Chaudhuri, A. R. (2017). Immigration-induced effects of changes in size and skill distribution of the labor force on wages in the us. *Journal of Macroeconomics*, 52:118–134. [Cited on page 6.]

Peri, G. and Sparber, C. (2009). Task specialization, immigration, and wages. *American Economic Journal: Applied Economics*, 1(3):135–169. [Cited on page 10.]

Peters, M. (2022). Market size and spatial growth—evidence from germany's post-war population expulsions. *Econometrica*, 90(5):2357–2396. [Cited on page 23.]

Rocha, R., Ferraz, C., and Soares, R. R. (2017). Human capital persistence and development. *American Economic Journal: Applied Economics*, 9(4):105–136. [Cited on pages 5, 11, and 23.]

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers*, 3(2):135–146. [Cited on page 10.]

Ruggles, S., Flood, S., Sobek, M., Danika, B., Cooper, G., Richards, S., and Schouweiler, M. (2022). Ipums usa: Version 13.0 [dataset]. minneapolis, mn: Ipums, 2022. [Cited on page 16.]

Sanderson, E. and Windmeijer, F. (2016). A weak instrument f-test in linear iv models with multiple endogenous variables. *Journal of econometrics*, 190(2):212–221. [Cited on page 19.]

Sequeira, S., Nunn, N., and Qian, N. (2020). Immigrants and the making of america. *The Review of Economic Studies*, 87(1):382–419. [Cited on pages 2, 5, 10, and 20.]

Shapiro, J. M. (2006). Smart cities: quality of life, productivity, and the growth effects of human capital. *The review of economics and statistics*, 88(2):324–335. [Cited on pages 4 and 20.]

Terry, S. J., Chaney, T., Burchardi, K. B., Tarquinio, L., and Hassan, T. A. (2021). Immigration, innovation, and growth. Technical report, National Bureau of Economic Research. [Cited on pages 6 and 12.]

Ulltveit-Moe, K. H., Moxnes, A., Bratsberg, B., Raaum, O., et al. (2019). Opening the floodgates: industry and occupation adjustments to labor immigration. Technical report, CEPR Discussion Papers. [Cited on page 5.]

Valencia Caicedo, F. (2019). The mission: Human capital transmission, economic persistence, and culture in south america. *The Quarterly Journal of Economics*, 134(1):507–556. [Cited on pages 5 and 11.]

Voth, H.-J. (2021). Persistence–myth and mystery. In *The handbook of historical economics*, pages 243–267. Elsevier. [Cited on pages 2, 5, and 10.]

Wittke, C. (1939). *We Who Built America*. The Press of Western Reserve University, Ann Arbor, MI. [Cited on page 10.]

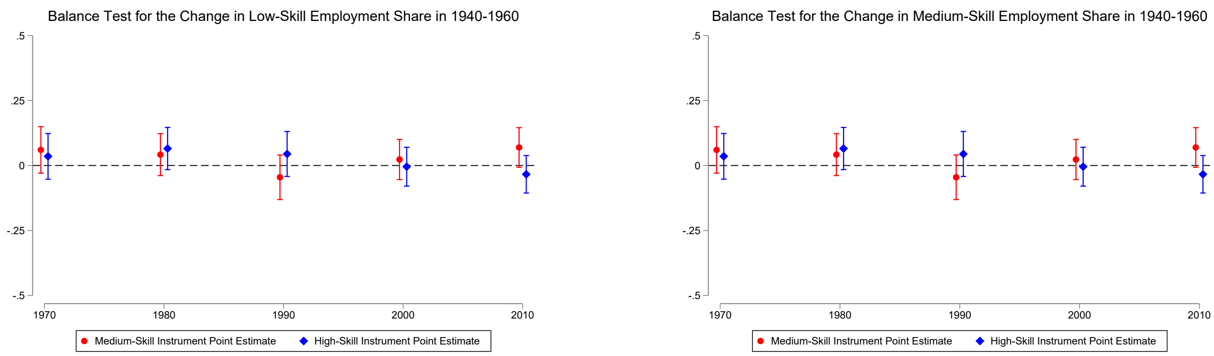

Figure 1: Balance Tests: The Changes in Employment Shares in 1940-1960

Notes: These figures display the results of the balance tests specified in equation (12). Coefficients report the effect of a one-standard-deviation increase in each instrument, human-capital-specific exogenous population, over the 1970-2010 period on a standardized measure of the changes in low- and medium-skill employment shares between 1940 and 1960. All regressions include state and time fixed effects, and standard errors are clustered by county for all specifications. Whiskers indicate 95% confidence intervals computed using standard errors clustered by county.

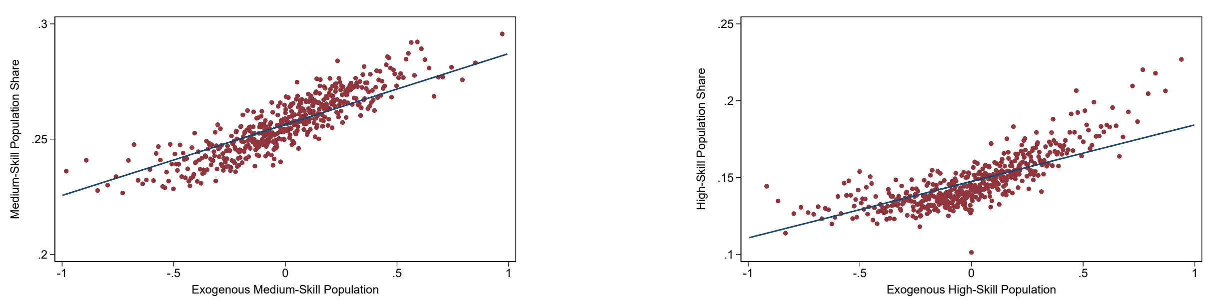

Figure 2: First-Stage Relationship

Notes: These figures display binned scatter plots of the first-stage estimates, whereby exogenous working-age population of each human capital level obtained in equation (10) is plotted against endogenous working-age population of each human capital level. The number of bins is fixed at 500. All regressions include state and time fixed effects, and standard errors are clustered by county for all specifications. The figures have been plotted using the "`binsreg`" package in Stata (Cattaneo et al., 2024).

Table 1: Summary Statistics

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Population			
$Working-Age\ Population\ Share_{e,d,t}$ (%)	59.57	25.61	14.73
	(13.29)	(8.00)	(6.68)
$Mean\ Education_{e,d,t}$	10.80	13.55	16.77
	(0.63)	(0.11)	(0.19)
Panel B: Industry			
Baseline			
$Employment\ Share_{e,d,t}$ (%)	31.79	45.99	22.22
	(25.53)	(17.13)	(11.75)
$Establishment\ Share_{e,d,t}$ (%)	23.98	44.09	31.82
	(20.16)	(15.15)	(16.18)
$Mean\ Education_{e,d,t}$	11.41	12.76	14.82
	(0.46)	(0.14)	(0.14)
Tradable Sector			
$Employment\ Share_{e,d,t}$ (%)	17.69	12.36	0.70
	(13.75)	(7.95)	(1.34)
$Establishment\ Share_{e,d,t}$ (%)	11.10	9.56	2.44
	(19.97)	(13.54)	(5.36)
$Mean\ Education_{e,d,t}$	11.27	12.62	14.44
	(0.54)	(0.22)	(0.21)
Nontradable Sector			
$Employment\ Share_{e,d,t}$ (%)	15.15	31.61	22.49
	(13.20)	(10.67)	(10.59)
$Establishment\ Share_{e,d,t}$ (%)	14.68	31.59	30.62
	(11.55)	(14.06)	(16.84)
$Mean\ Education_{e,d,t}$	11.56	12.82	14.86
	(0.39)	(0.12)	(0.16)
N	15,705	15,705	15,705

Notes: This table presents the mean, standard deviation (in parentheses), and the number of observations for both independent and dependent variables of all skill types (e) at the county (d) level for the decennial 1970-2010 period (t). Panel A displays the human-capital-specific variables. The working-age population share variable of each skill type at the county level in a certain decennial census year equals the number of working-age individuals with a given mean education level by county relative to total county population in that year. The employment share variable of each skill type at the county level in a given census year represents the number of workers of each skill type by county relative to the entire workforce in that county and census year. The establishment share variable of each skill type at the county level in a certain year is normalized by the total number of establishments by county in that year. Panel B reports the industry-skill-specific variables utilized in the baseline and tradability analyses. The determinant of the continuous skill measure corresponds to mean education of workers employed by each industry in each year. The measure of tradability is based on Burstein et al. (2020). The tradable sector includes the agriculture, mining, and manufacturing industries, whereas the nontradable sector covers all services industries. The employment share variable of each skill type in each sector and census year is normalized by the entire workforce in each sector and year. The establishment share of each skill type in each sector and year is normalized by the total number of establishments in each sector and year.

Table 2: OLS Estimates of Employment and Establishment Shares

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Employment Shares			
$Population\ Share_{MS,d,t}$	-0.610***	0.548***	0.062***
	(0.016)	(0.022)	(0.015)
$Population\ Share_{HS,d,t}$	-0.362***	-0.227***	0.588***
	(0.009)	(0.014)	(0.011)
R^2	0.973	0.913	0.943
Panel B: Establishment Shares			
$Population\ Share_{MS,d,t}$	-0.145***	0.086**	0.059**
	(0.039)	(0.036)	(0.030)
$Population\ Share_{HS,d,t}$	-0.270***	-0.300***	0.565***
	(0.019)	(0.019)	(0.021)
R^2	0.310	0.064	0.454
N	15,705	15,705	15,705
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table reports the OLS results of estimating equation (6). The dependent variables are employment and establishment shares of each industrial skill type at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are medium- and high-skill working-age population shares at the county level. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 3: IV ESTIMATES OF EMPLOYMENT AND ESTABLISHMENT SHARES

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Employment Shares			
$Population\,Share_{MS,d,t}$	-0.389***	0.212***	0.177***
	(0.047)	(0.059)	(0.036)
$Population\,Share_{HS,d,t}$	-0.315***	-0.229***	0.544***
	(0.022)	(0.027)	(0.020)
AR Wald F-Test P-value	0.000	0.000	0.000
Panel B: Establishment Shares			
$Population\,Share_{MS,d,t}$	-0.745***	-0.056	0.955***
	(0.135)	(0.133)	(0.114)
$Population\,Share_{HS,d,t}$	-0.110*	-0.175***	0.331***
	(0.065)	(0.061)	(0.055)
AR Wald F-Test P-value	0.000	0.006	0.000
SW First-Stage F-Stats	133; 364	133; 364	133; 364
N	15,705	15,705	15,705
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table reports the IV results of estimating equation (6). The dependent variables are employment and establishment shares of each industrial skill type at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with medium- and high-skill exogenous working-age population at the county level. Each specification presents the p-value for the Anderson-Rubin Wald F-test and the Sanderson-Windmeijer F-statistic for each endogenous variable. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 4: Employment Estimates in the Tradable and Nontradable Sectors

	Employment Shares			
	All	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)	(4)
Panel A: Tradable Sector				
Population Share$_{MS,d,t}$	-0.444***	-0.473***	0.020	0.011
	(0.076)	(0.062)	(0.035)	(0.012)
Population Share$_{HS,d,t}$	-0.601***	-0.290***	-0.343***	0.035***
	(0.035)	(0.029)	(0.020)	(0.006)
AR Wald F-Test P-value	0.000	0.000	0.000	0.000
N	15,705	15,705	15,705	12,564
Panel B: Nontradable Sector				
Population Share$_{MS,d,t}$	0.444***	0.015	0.246***	0.183***
	(0.076)	(0.030)	(0.051)	(0.038)
Population Share$_{HS,d,t}$	0.601***	-0.060***	0.100***	0.561***
	(0.035)	(0.014)	(0.025)	(0.021)
AR Wald F-Test P-value	0.000	0.000	0.000	0.000
N	15,705	15,705	15,705	15,705
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the IV results of estimating equation (6). The dependent variables are employment shares of each industrial skill type in the tradable and nontradable sectors at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with medium- and high-skill exogenous working-age population at the county level. Each specification presents the p-value for the Anderson-Rubin Wald F-test. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table 5: Establishment Estimates in the Tradable and Nontradable Sectors

	Establishment Shares			
	All	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)	(4)
Panel A: Tradable Sector				
Population Share$_{MS,d,t}$	-1.030***	-0.927***	-0.061	-0.042
	(0.172)	(0.140)	(0.094)	(0.045)
Population Share$_{HS,d,t}$	-0.181**	-0.109*	0.005	-0.077***
	(0.074)	(0.064)	(0.043)	(0.021)
AR Wald F-Test P-value	0.000	0.000	0.716	0.000
Panel B: Nontradable Sector				
Population Share$_{MS,d,t}$	1.030***	0.204***	-0.133	0.960***
	(0.172)	(0.066)	(0.103)	(0.115)
Population Share$_{HS,d,t}$	0.181**	-0.098***	-0.205***	0.484***
	(0.074)	(0.030)	(0.047)	(0.054)
AR Wald F-Test P-value	0.000	0.000	0.000	0.000
N	15,705	15,705	15,705	15,705
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Notes: This table reports the IV results of estimating equation (6). The dependent variables are employment shares of each industrial skill type in the tradable and nontradable sectors at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with medium- and high-skill exogenous working-age population at the county level. Each specification presents the p-value for the Anderson-Rubin Wald F-test. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Appendices

Appendix Contents

A Model Derivations

This appendix presents the full derivations of the equilibrium allocations described in Section 2. Let N_h and N_l represent the total labor supplies of high- and low-skill workers, respectively. Each representative firm $i \in \{b, w\}$ produces output Q_i according to a constant elasticity of substitution (CES) production function given by

$$Q_i = \left(\beta_{i,h} L_{i,h}^{\frac{\sigma-1}{\sigma}} + \beta_{i,l} L_{i,l}^{\frac{\sigma-1}{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}, \tag{A.1}$$

where $\sigma > 1$ denotes the elasticity of substitution between high- and low-skill labor inputs, and $\beta_{i,h}$ and $\beta_{i,l}$ represent relative productivity parameters for each labor type within firm i. The productivity parameters differ across firm types such that $\beta_{w,h} > \beta_{b,h}$ and $\beta_{b,l} > \beta_{w,l}$, with $\beta_{i,h} + \beta_{i,l} = 1$. Within each firm, white-collar firms employ high-skill labor more intensively ($\beta_{w,h} > \beta_{w,l}$), whereas blue-collar firms rely more heavily on low-skill labor ($\beta_{b,l} > \beta_{b,h}$).

Each representative firm maximizes profits subject to the aggregate labor-supply constraints:

$$\begin{aligned} \max_{L_{i,h}, L_{i,l}} \quad & \{Q_i - \omega_h L_{i,h} - \omega_l L_{i,l}\} \\ \text{s.t.} \quad & L_{b,h} + L_{w,h} = N_h, \\ & L_{b,l} + L_{w,l} = N_l, \end{aligned} \tag{A.2}$$

where ω_h and ω_l denote the common wages for high- and low-skill workers across firms.

First-order conditions for both types of labor demanded by each firm yield the following equilibrium allocations:

$$L_{b,l} = \frac{N_l B_w^{-1} - N_h Z^{-1}}{B_w^{-1} - B_b^{-1}}, \tag{A.3}$$

$$L_{w,l} = \frac{N_l B_b^{-1} - N_h Z^{-1}}{B_b^{-1} - B_w^{-1}}, \tag{A.4}$$

$$L_{b,h} = \frac{N_h B_w - N_l Z}{B_w - B_b}, \tag{A.5}$$

$$L_{w,h} = \frac{N_h B_b - N_l Z}{B_b - B_w}, \tag{A.6}$$

where

$$B_b = \left(\frac{\beta_{b,l}}{\beta_{b,h}} \right)^{\sigma}, \quad B_w = \left(\frac{\beta_{w,l}}{\beta_{w,h}} \right)^{\sigma}, \quad Z = \left(\frac{\beta_{b,l}^{\sigma} - \beta_{w,l}^{\sigma}}{\beta_{w,h}^{\sigma} - \beta_{b,h}^{\sigma}} \right)^{\frac{\sigma}{\sigma-1}}.$$

Let industry-level labor-demand shares, *i.e.*, the share of workers employed in each firm type, be denoted by

$$s_w = \frac{L_{w,h} + L_{w,l}}{\sum_i (L_{i,h} + L_{i,l})}, \qquad s_b = \frac{L_{b,h} + L_{b,l}}{\sum_i (L_{i,h} + L_{i,l})},$$

and labor-supply shares by

$$n_h = \frac{N_h}{N_h + N_l}, \qquad n_l = \frac{N_l}{N_h + N_l}.$$

The differences in the relative productivity parameters and the imperfect substitutability between labor types ensure that $B_b > 1$, $B_w < 1$, $B_b > B_w$, and $Z > 0$.

Substituting equilibrium allocations from equations (A.3)–(A.6) into these share expressions gives the comparative static responses of industry labor-demand shares with respect to labor-supply shares:

$$\frac{\partial s_w}{\partial n_h} = \left(\frac{B_b}{B_b - B_w}\right)\left(\frac{Z + B_w}{Z}\right) > 0, \tag{A.7}$$

$$\frac{\partial s_b}{\partial n_h} = -\left(\frac{B_w}{B_b - B_w}\right)\left(\frac{Z + B_b}{Z}\right) < 0, \tag{A.8}$$

$$\frac{\partial s_b}{\partial n_l} = \left(\frac{Z + B_b}{B_b - B_w}\right) > 0, \tag{A.9}$$

$$\frac{\partial s_w}{\partial n_l} = -\left(\frac{Z + B_w}{B_b - B_w}\right) < 0. \tag{A.10}$$

Equations (A.7)–(A.9) confirm that an increase in the relative supply of high-skill workers raises the labor-demand share of white-collar firms, while an increase in low-skill labor supply raises the labor-demand share of blue-collar firms. These results directly reflect firms' technological differences, whereby white-collar firms are more efficient in employing high-skill labor, and blue-collar firms are more efficient in employing low-skill labor. As shown in Section 2, the magnitude of these responses depends on the elasticity of substitution σ, which governs the degree to which production can substitute between skill groups.

B Data Appendix

B.1 Construction of Immigration Stock Data

I construct county-level data on immigration utilizing the individual census files of the IPUMS USA and NHGIS samples of the 1850 (full sample), 1860 (full sample), 1870 (10% sample), 1880 (10% sample), 1890 (10% sample), 1900 (5% sample), 1910 (1% sample), 1920 (10% sample), 1930 (5% sample), 1940 (1% sample), 1950 (1% sample), 1960 (5% sample), 1970 (1% Form 1 Metro sample), 1980 (5% State sample), 1990 (5% State sample), and 2000 (5% sample) waves, as well as the 2006-2010 five-year sample of the ACS. I define immigrants as "foreign-born" individuals born outside of the US and its territories in each decennial wave beginning from 1850 to 2000 in the census and the 2006-2010 five-year sample of the ACS. I define the measure of immigration stock, $I_{o,d,t}$, as the number of immigrants (foreign-born people) from origin (country) o, who resides in US destination (county) d in census year t. Countries of origin are based on the detailed birthplace variable in the census.

Both origins and destinations are either historical (prior to 1990) or modern (post-1990) geographical units. Therefore, to have consistent geographic units of countries and counties, I follow the method highlighted in Burchardi et al. (2019), in that I transform non-1990-level foreign countries of origin or detailed birthplaces to 1990-level foreign countries as well as non-1990-level counties, county groups, and PUMAs to 1990-level counties. Initially, I transition non-1990-level counties, county groups, and PUMAs to 1990-level counties in each census year utilizing the transition matrices created in Burchardi et al. (2019) and complementing it with county crosswalk files generated in Eckert et al. (2020b).[31] Afterwards, I generate a country crosswalk file and transform non-1990-level countries to 1990-level countries in each census wave. All transition matrices and crosswalk files rely on population-based weights.

B.2 Construction of Working-Age Population Data

I construct variables for human-capital-specific working-age population shares using each decennial census wave in 1970, 1980, 1990, 2000, and the 2006-2010 five-year sample of the ACS in the following steps. First, I restrict the sample to individuals aged 16-64. Second, I allocate them into skill cells based on their educational attainment (*i.e.*, years of schooling). Finally, I take the number of individuals with a given human capital level by county relative to the entire working-

[31] The additional details on the construction of county crosswalk files and the weighting scheme can be found in those papers alongside the following websites: (1) and (2).

age population by county. To have consistent geographic units of counties, I follow the procedure described above.

B.3 Construction of Employment Data

I construct variables for all, native, and immigrant employment shares of each industry-skill type utilizing each decennial census wave in 1970, 1980, 1990, 2000, and the 2006-2010 five-year sample of the ACS. The measure of employment is the number of individuals employed, and the continuous skill measure corresponds to average years of education of workers employed by each industry. I further distinguish employment shares by industry tradability and discuss how I generate these shares below.

Initially, I restrict the sample to only households and drop individuals working in the military occupations along with those classified as "unemployed" and "not in the labor force." Afterwards, I aggregate the 1990-level census industry codes to a balanced panel of industries utilizing an industry concordance outlined in Autor et al. (2019).[32] Then I allocate workers to three industry-skill cells based on mean education of workers in industries. Later, I take the number of workers by county and industry-skill types relative to the entire workforce by county. I proceed with the construction of immigrant and native employment shares similarly, in that I take the number of workers by county, industry-skill types, and nativity relative to the entire workforce by county. Lastly, I generate the employment shares for workers of all industry-skill types in the tradable and nontradable sectors through distinguishing them by industry tradability.

B.4 Construction of Establishments Data

I construct establishment share of each industry skill type by county along with industry tradability using the 1970, 1980, 1990, 2000, and 2010 extracts of the County Business Patterns (CBP) data published by the Census Bureau and supplement them with those assembled in Eckert et al. (2020a, 2022).[33] The data provide further important details on establishment size, defined as the number of workers employed in a given establishment, and industry, in which an establishment operates. Apart from establishment and industry characteristics, they also provide employment figures during the pay period including March 12, first-quarter payroll, and annual payroll. Consistent with the literature, I define small establishments as those employing fewer than 20 workers, medium

[32] For additional details, see their paper and David Dorn's website. The crosswalk file I utilize is "[C4] Census *ind*1990 to *ind*1990*dd*."

[33] For further details on digitization, cleaning, imputation on employment counts, and consolidation of these panels under consistent industry codes, see their papers, and Fabian Eckert's website.

establishments as those employing between 20 and 500 workers, and large establishments as those employing more than 500 workers. I carry out the following steps to create establishment shares.

First, using the industry concordance files generated in Eckert et al. (2020a, 2022), I crosswalk all industry codes to a unique industry classification, $SIC87$. Afterwards, I map the $SIC87$ industry codes to a balanced panel of industries compiled in Autor et al. (2019).[34] This step is necessary for not only making them compatible with the census industry codes but also identifying the industry-skill rankings discussed in the text. Next I take the number of establishments by county and industry-skill types relative to the total number of establishments by county. I further differentiate establishment share of each industry-skill type by the industry tradability.

[34]See David Dorn's website for additional details. The crosswalk file I utilize is "[C8] $SIC87$ to $ind1990ddx$."

C Appendix Tables

Table C.1: Allocation of States to Census Divisions

Census Division	**State**
New England	Connecticut, Maine, Massachusetts, New Hampshire, Rhode Island, Vermont
Middle Atlantic	New Jersey, New York, Pennsylvania
East North Central	Illinois, Indiana, Michigan, Ohio, Wisconsin
West North Central	Iowa, Kansas, Minnesota, Missouri, Nebraska, North Dakota, South Dakota
South Atlantic	Delaware, District of Columbia, Florida, Georgia, Maryland, North Carolina, South Carolina, Virginia, West Virginia
East South Central	Alabama, Kentucky, Mississippi, Tennessee
West South Central	Arkansas, Louisiana, Oklahoma, Texas
Mountain	Arizona, Colorado, Idaho, Montana, Nevada, New Mexico, Utah, Wyoming
Pacific	Alaska, California, Hawaii, Oregon, Washington

Table C.2: Industry Skill Classification

Low Skill	Medium Skill		High Skill	
Agricultural production, crops	Agricultural production, livestock	Apparel and accessory stores, except shoe	Agricultural chemicals	Accounting, auditing, and bookkeeping services
Apparel and accessories, except knit	Bakery products	Auto and home supply stores	Aircraft and parts	Advertising
Carpets and rugs	Blast furnaces, steelworks, rolling, and finishing mills	Automobile parking and carwashes	Beverage industries	Air transportation
Logging	Canned, frozen, and preserved fruits, and vegetables	Automotive rental and leasing, without drivers	Computers and related equipment	Alcoholic beverages
Meat products	Cement, concrete, gypsum, and plaster products	Automotive repair and related services	Construction and material handling machines	Apparel, fabrics, and notions
Private households	Coal mining	Barber shops	Drugs	Banking
Scrap and waste materials	Cutlery, handtools, and general hardware	Beauty shops	Electrical machinery, equipment, and supplies, n.e.c	Book and stationery stores
Services to dwellings and other buildings	Cycles and miscellaneous transportation equipment	Bowling centers	Engines and turbines	Business services, n.e.c
	Dairy products	Bus service and urban transit	Forestry	Catalog and mail order houses
	Dyeing and finishing textiles, except wool and knit goods	Department stores	Guided missiles, space vehicles, and parts	Child day care services
	Fabricated structural metal products	Detective and protective services	Industrial and misc. chemicals	Colleges and universities
	Farm machinery equipment	Eating and drinking places	Machinery, n.s	Computer and data processing services
	Fishing, hunting, and trapping	Farm-product raw materials	Medical, dental, and optical instruments and supplies	Credit agencies, n.e.c
	Food industries, n.s	Food stores, n.e.c	Metal mining	Direct selling establishments
	Footwear, except rubber and plastic	Fuel dealers	Miscellaneous petroleum and coal products	Drug stores
	Furniture and fixtures	Gasoline service stations	Newspaper publishing and printing	Drugs, chemicals, and allied products
	Glass and glass products	Groceries and related products	Oil and gas extraction	Educational services, n.e.c
	Grain mill products	Grocery stores	Ordnance	Electric and gas, and other combinations
	Household appliances	Hardware stores	Paints, varnishes, and related products	Electric light and power
	Iron and steel foundries	Hotels and motels	Petroleum refining	Electrical goods
	Knitting mills	Laundry, cleaning, and garment services	Pottery and related products	Electrical repair shops
	Landscape and horticultural services	Lodging places, except hotels and motels	Printing, publishing, and allied industries, except newspapers	Elementary and secondary schools
	Leather products, except footwear	Lumber and building material retailing	Radio, TV, and communication equipment	Engineering, architectural, and surveying services
	Machinery, except electrical, n.e.c	Lumber and construction materials	Scientific and controlling instruments	Farm supplies
	Manufacturing industries, n.s	Misc. entertainment and recreation services	Soaps and cosmetics	Funeral service and crematories
	Metalworking machinery	Misc. general merchandise stores	Tobacco manufactures	Furniture and home furnishings
	Misc. food preperations and kindred products	Misc. repair services	Furniture and home furnishings stores	Gas and steam supply systems
	Misc. nonmetallic mineral and stone products	Misc. vehicle dealers	Gift, novelty, and souvenir shops	Hardware, plumbing and heating supplies
	Misc. fabricated metal products	Motor vehicles and equipment	Health services, n.e.c	Hospitals
	Misc. fabricated textile products	Nursing and personal care facilities	Household appliance stores	Insurance
	Misc. manufacturing industries	Railroads	Labor unions	Legal services
	Misc. paper and pulp products	Retail bakeries	Liquor stores	Libraries
	Misc. plastics products	Retail florists	Machinery, equipment, and supplies	Management and public relations services
	Motor vehicles and motor vehicle equipment	Retail nurseries and garden stores	Membership organizations, n.e.c	Metals and minerals, except petroleum
	Nonmetallic mining and quarrying, except fuels	Sanitary services	Misc. personal services	Misc. professional and related services
	Other primary metal industries	Shoe stores	Misc. retail stores	Misc. wholesale, durable goods
	Other rubber products, and plastics footwear and belting	Taxicab service	Misc. wholesale, nondurable goods	Museums, art galleries, and zoos
	Paperboard containers and boxes	Trucking service	Jewelry stores	Job training and vocational rehabilitation services
	Plastics, synthetics, and resins	Vending machine operators	Music stores	Offices and clinics of chiropractors
	Primary aluminum industries	Warehousing and storage	Offices and clinics of dentists	Offices and clinics of optometrists
	Pulp, paper, and paperboard mills	Railroad locomotives and equipment	Offices and clinics of physicians	Paper and paper products
	Sawmills, planing mills, and millwork	Ship and boat building and repairing	Personnel supply services	Professional and commercial equipment and supplies
	Structural clay products	Sugar and confectionery products	Radio and television broadcasting and cable	Radio, tv, and computer stores
	Tires and inner tubes	Toys, amusement, and sporting goods	Real estate, including real estate-insurance offices	Religious organizations
	Veterinary services	Wood buildings and mobile homes	Research, development, and testing services	Residential care facilities, without nursing
	Yarn, thread, and fabric mills	All construction	Retail trade, n.s	Savings institutions, including credit unions
	Metal forgings and stampings		Security, commodity brokerage, and investment companies	Services incidental to transportation
			Sewing, needlework, and piece goods stores	Social services, n.e.c
			Sporting goods, bicycles, and hobby stores	Telephone communications
			Theaters and motion pictures	U.S. postal service
			Utilities, n.s	Video tape rental
			Water supply and irrigation	Water transportation
			Wholesale trade, n.s	Administration of economic programs
			Administration of human resources programs	Administration of environmental quality and housing programs
			General government, n.e.c	Justice, public order, and safety
			National security and international affairs	Public finance, taxation, and monetary policy

Notes: This detailed industry-skill classification is based on mean education of workers employed by a given industry as described in the text. Since this classification is a continuous measure of industry-skill types varying over time, I present the skill types for only the 2010-level data. Workers in low-skill, medium-skill, and high-skill industries have average years of education corresponding to a high-school degree and less, some college education, and college education and above, respectively.

Table C.3: Tradable and Nontradable Industries

Tradable Industries	**Nontradable Industries**
Agriculture, forestry and fisheries	Retail trade
Mining	Personal services
Transportation equipment	Professional and related services
Professional and photographic equipment and watches	Transportation
Petroleum and coal products	Wholesale trade, durables
Toys, amusement, and sporting goods	Wholesale trade, nondurables
Printing, publishing and allied industries	Communications
Apparel and other finished textile products	Business and repair services
Manufacturing industries, others	Finance, insurance, and real estate
Machinery and computing equipment	Entertainment and recreation services
Rubber and miscellaneous plastics products	Utilities and sanitary services
Textile mill products	
Chemicals and allied products	
Leather and leather products	
Electrical machinery, equipment, and supplies	
Furniture and fixtures	
Tobacco manufactures	
Food and kindred products	
Lumber, woods products (except furniture)	
Paper and allied products	
Stone, clay, glass and concrete products	

Notes: This broad industry tradability classification is based on Burstein et al. (2020). I classify all goods industries (agriculture, mining, and manufacturing) as tradable and all services industries as nontradable. I also drop the construction industry from the analysis.

Table C.4: Low-Skill Tradable and Nontradable Industries

Tradable Industries	**Nontradable Industries**
Agricultural production, crops	Private households
Apparel and accessories, except knit	Scrap and waste materials
Carpets and rugs	Services to dwellings and other buildings
Logging	
Meat products	

Notes: This detailed industry-skill classification across tradability is based on mean education of workers employed by a given industry as described in the text. Since this classification is a continuous measure of industry-skill types varying over time, I present the skill types for only the 2010-level data. In this case, workers in low-skill industries have average years of education corresponding to a high-school degree and less.

Table C.5: Medium-Skill Tradable and Nontradable Industries

Tradable Industries	**Nontradable Industries**
Agricultural production, livestock	Apparel and accessory stores, except shoe
Bakery products	Auto and home supply stores
Blast furnaces, steelworks, rolling, and finishing mills	Automobile parking and carwashes
Canned, frozen, and preserved fruits, and vegetables	Automotive rental and leasing, without drivers
Cement, concrete, gypsum, and plaster products	Automotive repair and related services
Coal mining	Barber shops
Cutlery, handtools, and general hardware	Beauty shops
Cycles and miscellaneous transportation equipment	Bowling centers
Dairy products	Bus service and urban transit
Dyeing and finishing textiles, except wool and knit goods	Department stores
Fabricated structural metal products	Detective and protective services
Farm machinery equipment	Eating and drinking places
Fishing, hunting, and trapping	Farm-product raw materials
Food industries, n.s	Food stores, n.e.c
Footwear, except rubber and plastic	Fuel dealers
Furniture and fixtures	Gasoline service stations
Glass and glass products	Groceries and related products
Grain mill products	Grocery stores
Household appliances	Hardware stores
Iron and steel foundries	Hotels and motels
Knitting mills	Laundry, cleaning, and garment services
Landscape and horticultural services	Lodging places, except hotels and motels
Leather products, except footwear	Lumber and building material retailing
Machinery, except electrical, n.e.c	Lumber and construction materials
Manufacturing industries, n.s	Misc. entertainment and recreation services
Metalworking machinery	Misc. general merchandise stores
Misc. food preperations and kindred products	Misc. repair services
Misc. nonmetallic mineral and stone products	Misc. vehicle dealers
Misc. fabricated metal products	Motor vehicles and equipment
Misc. fabricated textile products	Nursing and personal care facilities
Misc. manufacturing industries	Railroads
Misc. paper and pulp products	Retail bakeries
Misc. plastics products	Retail florists
Motor vehicles and motor vehicle equipment	Retail nurseries and garden stores
Nonmetallic mining and quarrying, except fuels	Sanitary services
Other primary metal industries	Shoe stores
Other rubber products, and plastics footwear and belting	Taxicab service
Paperboard containers and boxes	Trucking service
Plastics, synthetics, and resins	Vending machine operators
Primary aluminum industries	Warehousing and storage
Pulp, paper, and paperboard mills	
Railroad locomotives and equipment	
Sawmills, planing mills, and millwork	
Ship and boat building and repairing	
Structural clay products	
Sugar and confectionery products	
Tires and inner tubes	
Toys, amusement, and sporting goods	
Veterinary services	
Wood buildings and mobile homes	
Yarn, thread, and fabric mills	

Notes: This detailed industry-skill classification across tradability is based on mean education of workers employed by a given industry as described in the text. Since this classification is a continuous measure of industry-skill types varying over time, I present the skill types for the 2010-level data. In this case, workers in medium-skill industries have average years of education corresponding to some college education.

Table C.6: High-Skill Tradable and Nontradable Industries

Tradable Industries	**Nontradable Industries**	
Agricultural chemicals	Accounting, auditing, and bookkeeping services	Furniture and home furnishings stores
Aircraft and parts	Advertising	Gas and steam supply systems
Beverage industries	Air transportation	Gift, novelty, and souvenir shops
Computers and related equipment	Alcoholic beverages	Music stores
Construction and material handling machines	Apparel, fabrics, and notions	Personnel supply services
Drugs	Banking	Professional and commercial equipment and supplies
Electrical machinery, equipment, and supplies, n.e.c	Book and stationery stores	Hardware, plumbing and heating supplies
Engines and turbines	Business services, n.e.c	Health services, n.e.c
Forestry	Catalog and mail order houses	Household appliance stores
Guided missiles, space vehicles, and parts	Child day care services	Hospitals
Industrial and misc. chemicals	Colleges and universities	Insurance
Machinery, n.s	Computer and data processing services	Jewelry stores
Medical, dental, and optical instruments and supplies	Credit agencies, n.e.c	Job training and vocational rehabilitation services
Metal mining	Direct selling establishments	Labor unions
Miscellaneous petroleum and coal products	Drug stores	Liquor stores
Newspaper publishing and printing	Drugs, chemicals, and allied products	Libraries
Oil and gas extraction	Educational services, n.e.c	Machinery, equipment, and supplies
Ordnance	Electric and gas, and other combinations	Management and public relations services
Paints, varnishes, and related products	Electric light and power	Membership organizations, n.e.c
Petroleum refining	Electrical goods	Metals and minerals, except petroleum
Pottery and related products	Electrical repair shops	Misc. personal services
Printing, publishing, and allied industries, except newspapers	Elementary and secondary schools	Misc. professional and related services
Radio, TV, and communication equipment	Engineering, architectural, and surveying services	Misc. retail stores
Scientific and controlling instruments	Farm supplies	Misc. wholesale, durable goods
Soaps and cosmetics	Funeral service and crematories	Misc. wholesale, nondurable goods
Tobacco manufactures	Furniture and home furnishings	Museums, art galleries, and zoos
	Offices and clinics of chiropractors	Offices and clinics of dentists
	Offices and clinics of optometrists	Offices and clinics of physicians
	Paper and paper products	Radio and television broadcasting and cable
	Radio, tv, and computer stores	Real estate, including real estate-insurance offices
	Religious organizations	Research, development, and testing services
	Residential care facilities, without nursing	Retail trade, n.s
	Savings institutions, including credit unions	Security, commodity brokerage, and investment companies
	Services incidental to transportation	Sewing, needlework, and piece goods stores
	Social services, n.e.c	Sporting goods, bicycles, and hobby stores
	Telephone communications	Theaters and motion pictures
	U.S. postal service	Utilities, n.s
	Video tape rental	Water supply and irrigation
	Water transportation	Wholesale trade, n.s

Notes: This detailed industry-skill classification across tradability is based on mean education of workers employed by a given industry as described in the text. Since this classification is a continuous measure of industry-skill types varying over time, I present the skill types for only the 2010-level data. In this case, workers in high-skill industries have average years of education corresponding to college education and above.

Table C.7: Employment Estimates by Nativity

	All	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)	(4)
Panel A: Native Employment Shares				
Population Share$_{MS,d,t}$	0.045	-0.332***	0.091	0.044
	(0.076)	(0.046)	(0.070)	(0.067)
Population Share$_{HS,d,t}$	-0.088**	-0.192***	-0.384***	0.496***
	(0.037)	(0.021)	(0.032)	(0.030)
AR Wald F-Test P-value	0.006	0.000	0.000	0.000
Panel B: Immigrant Employment Shares				
Population Share$_{MS,d,t}$	-0.045	-0.709***	-0.091	0.800***
	(0.076)	(0.088)	(0.075)	(0.088)
Population Share$_{HS,d,t}$	0.088**	-1.012***	0.167***	0.845***
	(0.037)	(0.051)	(0.035)	(0.045)
AR Wald F-Test P-value	0.006	0.000	0.000	0.000
N	15,705	15,705	15,705	15,705
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are native and immigrant employment shares of each industrial skill type at the county level in separate regressions as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population at the county level. Each specification presents the p-value for the Anderson-Rubin Wald F-test. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

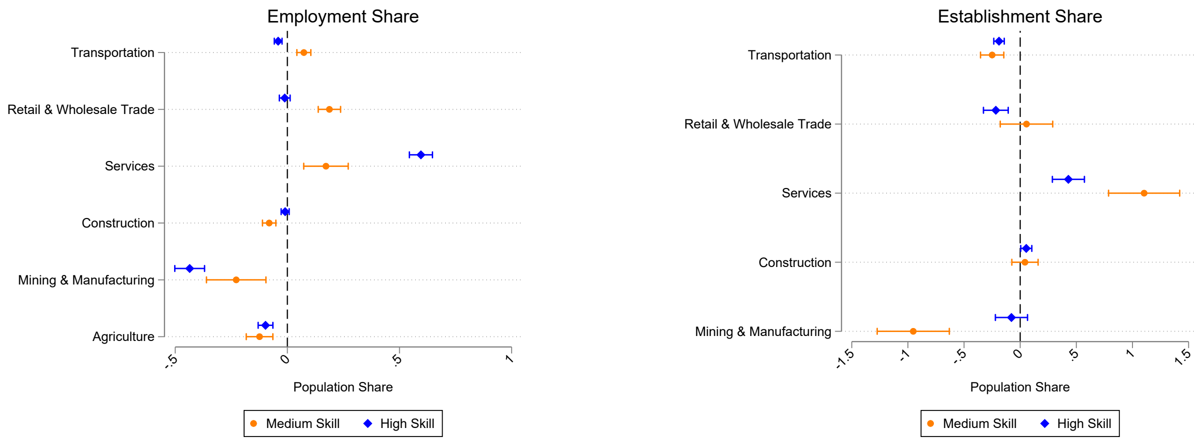

Figure C.1: Employment and Establishment Estimates by Industry

Notes: These figures display the coefficient plots of estimating equation (6), but the outcome variables are not disaggregated by industrial skill types. The dependent variables are employment and establishment shares of major industries at the county level in separate specifications. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population at the county level. The "Services" sector also includes "Finance, Insurance, and Real Estate" along with "Public Administration," and the "Transportation" sector includes "Communications" and "Other Public Utilities." Since the CBP data do not provide the full coverage of establishments operating in agriculture during the study period, I exclude them from the analysis. All regressions include state and time fixed effects, and standard errors are clustered by county for all specifications. Whiskers indicate 95% confidence intervals computed using standard errors clustered by county.

Table C.8: Establishment Estimates by Size

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Small Establishments			
$Population\ Share_{MS,d,t}$	-0.820***	-0.380	1.607***
	(0.179)	(0.230)	(0.244)
$Population\ Share_{HS,d,t}$	-0.228***	-0.217***	0.335***
	(0.036)	(0.047)	(0.043)
Panel B: Medium Establishments			
$Population\ Share_{MS,d,t}$	-0.285***	0.272***	0.187***
	(0.109)	(0.084)	(0.039)
$Population\ Share_{HS,d,t}$	-0.015	0.061*	0.048**
	(0.050)	(0.037)	(0.019)
Panel C: Large Establishments			
$Population\ Share_{MS,d,t}$	-0.055*	-0.016	-0.010
	(0.033)	(0.032)	(0.015)
$Population\ Share_{HS,d,t}$	-0.002	0.004	0.014*
	(0.013)	(0.014)	(0.008)
N	14,896	14,896	14,896
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are small, medium, and large establishment shares of each industrial skill type at the county level in separate regressions as displayed in Panels A, B, and C, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population at the county level. Consistent with the literature, small establishments are defined as those employing fewer than 20 workers, medium establishments are defined as those employing between 20 and 500 workers, and large establishments are defined as those employing more than 500 workers. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.9: Establishment Estimates by Size in the Tradable Sector

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Small Establishments			
$Population\ Share_{MS,d,t}$	-0.416***	-0.067**	-0.024**
	(0.065)	(0.034)	(0.010)
$Population\ Share_{HS,d,t}$	-0.120***	-0.030*	-0.009*
	(0.030)	(0.017)	(0.005)
Panel B: Medium Establishments			
$Population\ Share_{MS,d,t}$	-0.544***	0.077	0.020
	(0.112)	(0.079)	(0.025)
$Population\ Share_{HS,d,t}$	-0.055	0.027	-0.002
	(0.050)	(0.033)	(0.012)
Panel C: Large Establishments			
$Population\ Share_{MS,d,t}$	-0.046	-0.009	-0.011
	(0.033)	(0.032)	(0.014)
$Population\ Share_{HS,d,t}$	-0.001	0.003	0.010
	(0.013)	(0.015)	(0.008)
N	14,896	14,896	14,896
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are small, medium, and large establishment shares of each industrial skill type in the tradable sector in separate regressions as displayed in Panels A, B, and C, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population at the county level. Consistent with the literature, small establishments are defined as those employing fewer than 20 workers, medium establishments are defined as those employing between 20 and 500 workers, and large establishments are defined as those employing more than 500 workers. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.10: ESTABLISHMENT ESTIMATES BY SIZE IN THE NONTRADABLE SECTOR

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Small Establishments			
$Population\ Share_{MS,d,t}$	-0.098*	-0.258***	0.833***
	(0.060)	(0.094)	(0.098)
$Population\ Share_{HS,d,t}$	-0.099***	-0.292***	0.417***
	(0.026)	(0.042)	(0.044)
Panel B: Medium Establishments			
$Population\ Share_{MS,d,t}$	0.248***	0.162***	0.136***
	(0.024)	(0.040)	(0.046)
$Population\ Share_{HS,d,t}$	0.060***	0.035*	0.048***
	(0.012)	(0.020)	(0.018)
Panel C: Large Establishments			
$Population\ Share_{MS,d,t}$	0.000	-0.004	0.002
	(0.000)	(0.003)	(0.006)
$Population\ Share_{HS,d,t}$	-0.000	0.002	0.005**
	(0.000)	(0.001)	(0.002)
N	14,896	14,896	14,896
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are small, medium, and large establishment shares of each industrial skill type in the nontradable sector in separate regressions as displayed in Panels A, B, and C, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population at the county level. Consistent with the literature, small establishments are defined as those employing fewer than 20 workers, medium establishments are defined as those employing between 20 and 500 workers, and large establishments are defined as those employing more than 500 workers. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.11: Employment and Establishment Estimates Using Instruments for the 1850-1900 Period

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Employment Shares			
$Population\ Share_{MS,d,t}$	-0.381***	0.187***	0.194***
	(0.048)	(0.061)	(0.037)
$Population\ Share_{HS,d,t}$	-0.312***	-0.238***	0.550***
	(0.022)	(0.027)	(0.020)
AR Wald F-Test P-Value	0.000	0.000	0.000
Panel B: Establishment Shares			
$Population\ Share_{MS,d,t}$	-0.731***	-0.080	0.958***
	(0.138)	(0.137)	(0.119)
$Population\ Share_{HS,d,t}$	-0.086	-0.201***	0.333***
	(0.065)	(0.062)	(0.055)
AR Wald F-Test P-Value	0.00	0.000	0.000
SW First-Stage F-Stats	147; 374	147; 374	147; 374
N	15,705	15,705	15,705
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are employment and establishment shares of each industrial skill type at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population. Exogenous working-age population of each skill type has been generated using the predicted ancestry stocks based on the immigrant leave-out push-pull variables for only the 1850-1900 period. Each specification reports the p-value for the Anderson-Rubin Wald F-Test and the Sanderson-Windmeijer F-statistic for each endogenous variable. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.12: Employment and Establishment Estimates Using Instruments for the 1850-1950 Period

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Employment Shares			
$Population\ Share_{MS,d,t}$	-0.391***	0.206***	0.185***
	(0.048)	(0.060)	(0.037)
$Population\ Share_{HS,d,t}$	-0.317***	-0.229***	0.546***
	(0.022)	(0.027)	(0.020)
AR Wald F-Test P-Value	0.000	0.000	0.000
Panel B: Establishment Shares			
$Population\ Share_{MS,d,t}$	-0.744***	-0.070	0.958***
	(0.136)	(0.135)	(0.117)
$Population\ Share_{HS,d,t}$	-0.097	-0.193***	0.334***
	(0.064)	(0.061)	(0.055)
AR Wald F-Test P-Value	0.000	0.000	0.000
SW First-Stage F-Stats	138; 365	138; 365	138; 365
N	15,705	15,705	15,705
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are employment and establishment shares of each industrial skill type at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population. Exogenous working-age population of each skill type has been generated using the predicted immigration stocks based on the immigrant leave-out push-pull variables for only the 1850-1950 period. Each specification reports the p-value for the Anderson-Rubin Wald F-Test and the Sanderson-Windmeijer F-statistic for each endogenous variable. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.13: Robustness - Excluding the Five Largest Ancestry Groups

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Employment Shares			
$Population\ Share_{MS,d,t}$	-0.350***	0.152**	0.198***
	(0.049)	(0.061)	(0.037)
$Population\ Share_{HS,d,t}$	-0.330***	-0.211***	0.540***
	(0.022)	(0.027)	(0.020)
AR Wald F-Test P-Value	0.000	0.000	0.000
Panel B: Establishment Shares			
$Population\ Share_{MS,d,t}$	-0.723***	-0.109	0.988***
	(0.138)	(0.136)	(0.117)
$Population\ Share_{HS,d,t}$	-0.111*	-0.162***	0.318***
	(0.063)	(0.060)	(0.054)
AR Wald F-Test P-Value	0.000	0.019	0.000
SW First-Stage F-Stats	149; 355	149; 355	149; 355
N	15,705	15,705	15,705
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are employment and establishment shares of each industrial skill type in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population. While isolating exogenous working-age population of each skill type, I exclude the five largest ancestry groups from the list of predicted ancestry stocks. These groups are English, German, Irish, Italian, and Mexican. Each specification reports the p-value for the Anderson-Rubin Wald F-Test and the Sanderson-Windmeijer F-statistic for each endogenous variable. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.14: ROBUSTNESS - EXCLUDING THE FIVE LARGEST COUNTIES

	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)
Panel A: Employment Shares			
$Population\ Share_{MS,d,t}$	-0.395***	0.219***	0.176***
	(0.047)	(0.058)	(0.036)
$Population\ Share_{HS,d,t}$	-0.318***	-0.226***	0.544***
	(0.022)	(0.026)	(0.020)
AR Wald F-Test P-Value	0.000	0.000	0.000
Panel B: Establishment Shares			
$Population\ Share_{MS,d,t}$	-0.762***	-0.036	0.989***
	(0.134)	(0.131)	(0.111)
$Population\ Share_{HS,d,t}$	-0.116*	-0.164***	0.347***
	(0.064)	(0.060)	(0.053)
AR Wald F-Test P-Value	0.000	0.008	0.000
SW First-Stage F-Stats	132; 372	132; 372	132; 372
N	15,680	15,680	15,680
State FE	Yes	Yes	Yes
Time FE	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variables are employment and establishment shares of each industrial skill type at the county level excluding the five most "immigrant-receiving" counties in separate regressions as reported in Panels A and B, respectively. The excluded counties are Los Angeles, Cook, Harris, Miami-Dade, and Kings counties. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population. Each specification reports the p-value for the Anderson-Rubin Wald F-Test and the Sanderson-Windmeijer F-statistic for each endogenous variable. All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.15: Robustness - Alternative Tradability Classification

	Employment Shares			
	All	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)	(4)
Panel A: Tradable Sector				
$Population\ Share_{MS,d,t}$	-0.491***	-0.559***	0.171***	-0.102***
	(0.098)	(0.080)	(0.049)	(0.032)
$Population\ Share_{HS,d,t}$	-0.483***	-0.277***	-0.428***	0.222***
	(0.049)	(0.037)	(0.028)	(0.020)
Panel B: Nontradable Sector				
$Population\ Share_{MS,d,t}$	0.491***	0.354***	0.012	0.125***
	(0.098)	(0.045)	(0.043)	(0.040)
$Population\ Share_{HS,d,t}$	0.483***	-0.020	0.030	0.474***
	(0.049)	(0.022)	(0.022)	(0.022)
N	15,705	15,705	15,705	15,705
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variable is employment share of each industrial skill type in the tradable and nontradable sectors at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population. In classifying workers by the tradability domain, I use an alternative tradability measure highlighted in Mian and Sufi (2014) based on geographical Herfindahl-Hirschman Indices (HHIs). All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

Table C.16: Robustness - Alternative Tradability Classification

	Establishment Shares			
	All	Low Skill	Medium Skill	High Skill
	(1)	(2)	(3)	(4)
Panel A: Tradable Sector				
$Population\ Share_{MS,d,t}$	-0.971***	-1.052***	0.076	0.006
	(0.148)	(0.155)	(0.126)	(0.062)
$Population\ Share_{HS,d,t}$	-0.099	-0.168**	0.076	-0.006
	(0.068)	(0.070)	(0.055)	(0.026)
Panel B: Nontradable Sector				
$Population\ Share_{MS,d,t}$	0.971***	-0.099	-0.104	1.174***
	(0.148)	(0.081)	(0.098)	(0.089)
$Population\ Share_{HS,d,t}$	0.099	-0.155***	-0.213***	0.467***
	(0.068)	(0.037)	(0.046)	(0.047)
N	15,705	15,705	15,705	15,705
State FE	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes

Notes: This table presents the IV results of estimating equation (6). The dependent variable is establishment share of each industrial skill type in the tradable and nontradable sectors at the county level in separate specifications as displayed in Panels A and B, respectively. The independent variables are endogenous medium- and high-skill working-age population shares instrumented with exogenous medium- and high-skill working-age population. In classifying establishments by the tradability domain, I use an alternative tradability measure highlighted in Mian and Sufi (2014) based on geographical Herfindahl-Hirschman Indices (HHIs). All regressions include state and time fixed effects. Standard errors are clustered by county for all specifications, and *, **, and *** denote statistical significance at the 10%, 5%, and 1% levels, respectively.

D Appendix Figures

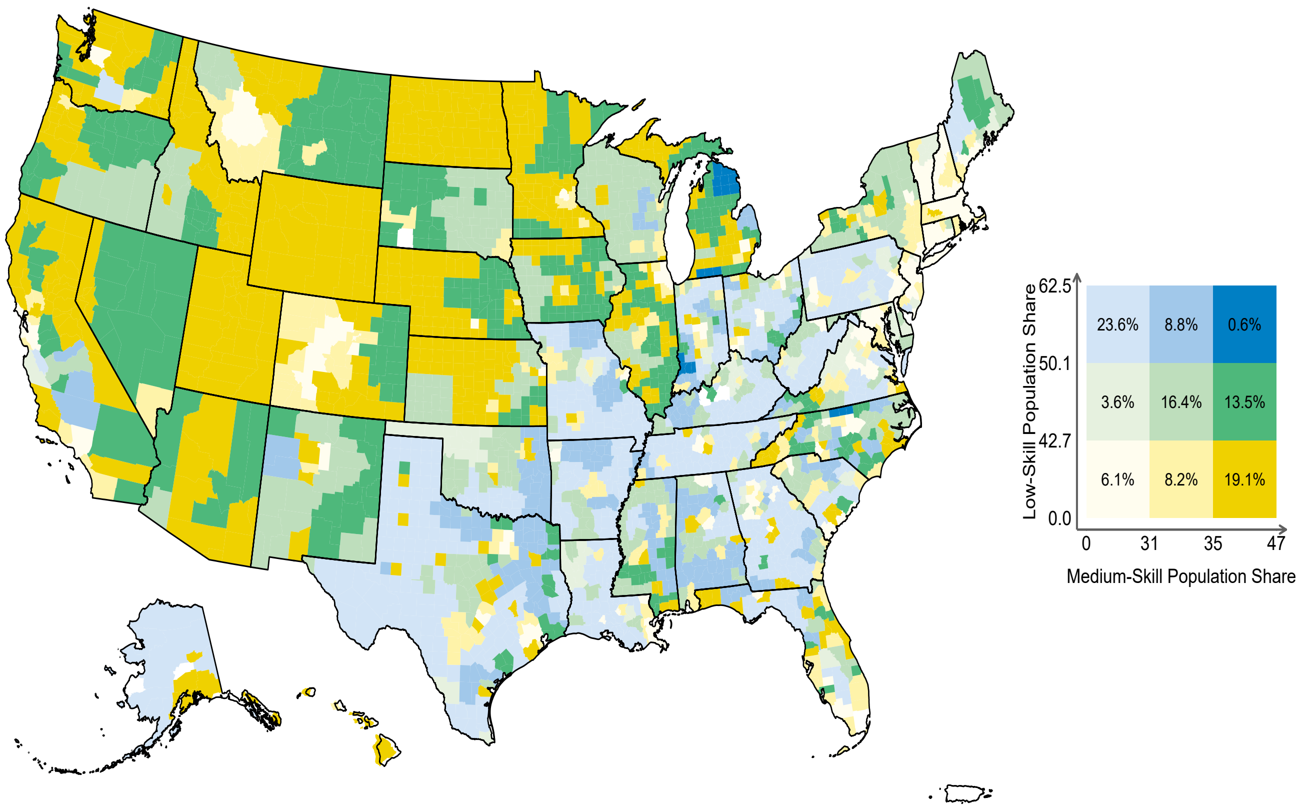

Figure D.1: The Bivariate Map of Low- and Medium-Skill Working-Age Population Shares in 2010

Notes: This bivariate map illustrates the distribution of low- and medium-skill working-age population shares across US counties in 2010. Lighter colors indicate lower concentrations, while darker colors represent higher concentrations for each respective skill type. The map has been plotted using the "`bimap`" package in Stata (Naqvi, 2024).

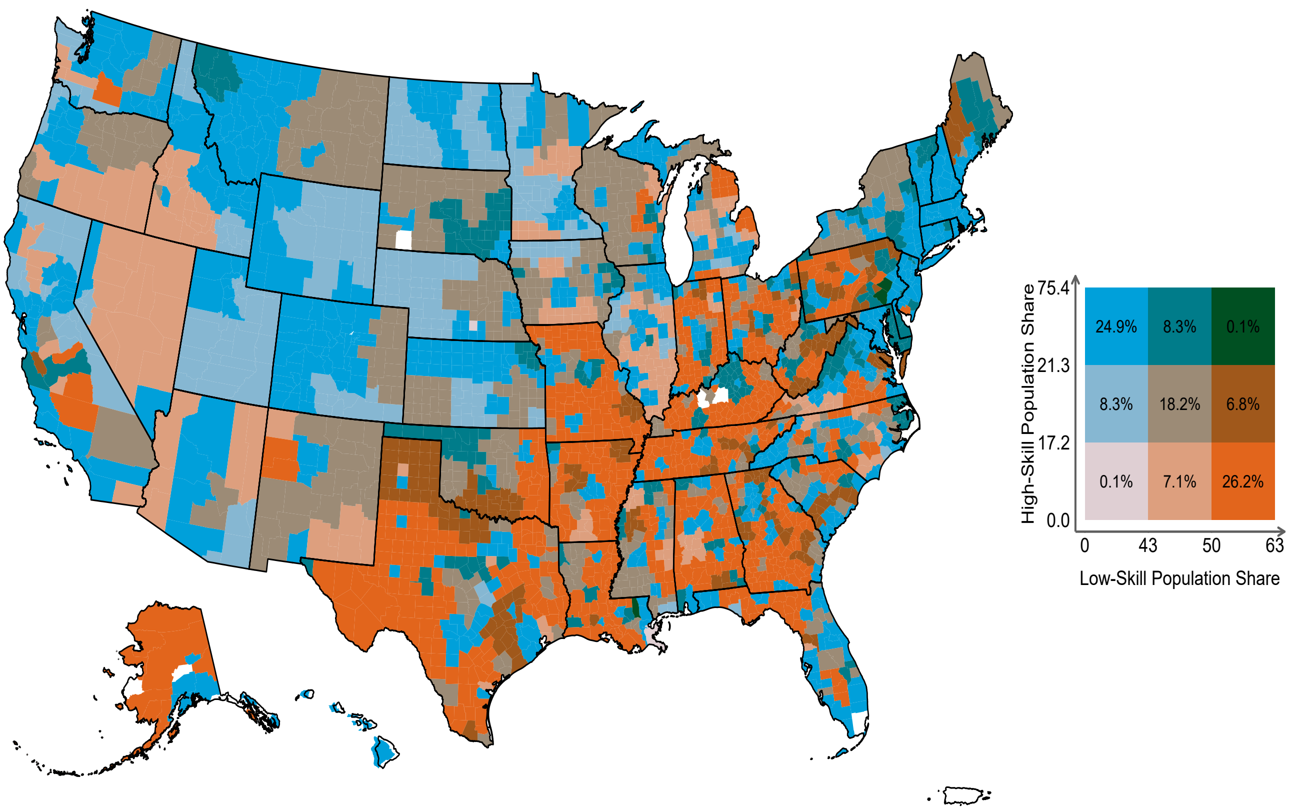

Figure D.2: The Bivariate Map of Low- and High-Skill Working-Age Population Shares in 2010

Notes: This bivariate map illustrates the distribution of low- and high-skill working-age population shares across US counties in 2010. Lighter colors indicate lower concentrations, while darker colors represent higher concentrations for each respective skill type. The map has been plotted using the "`bimap`" package in Stata (Naqvi, 2024).

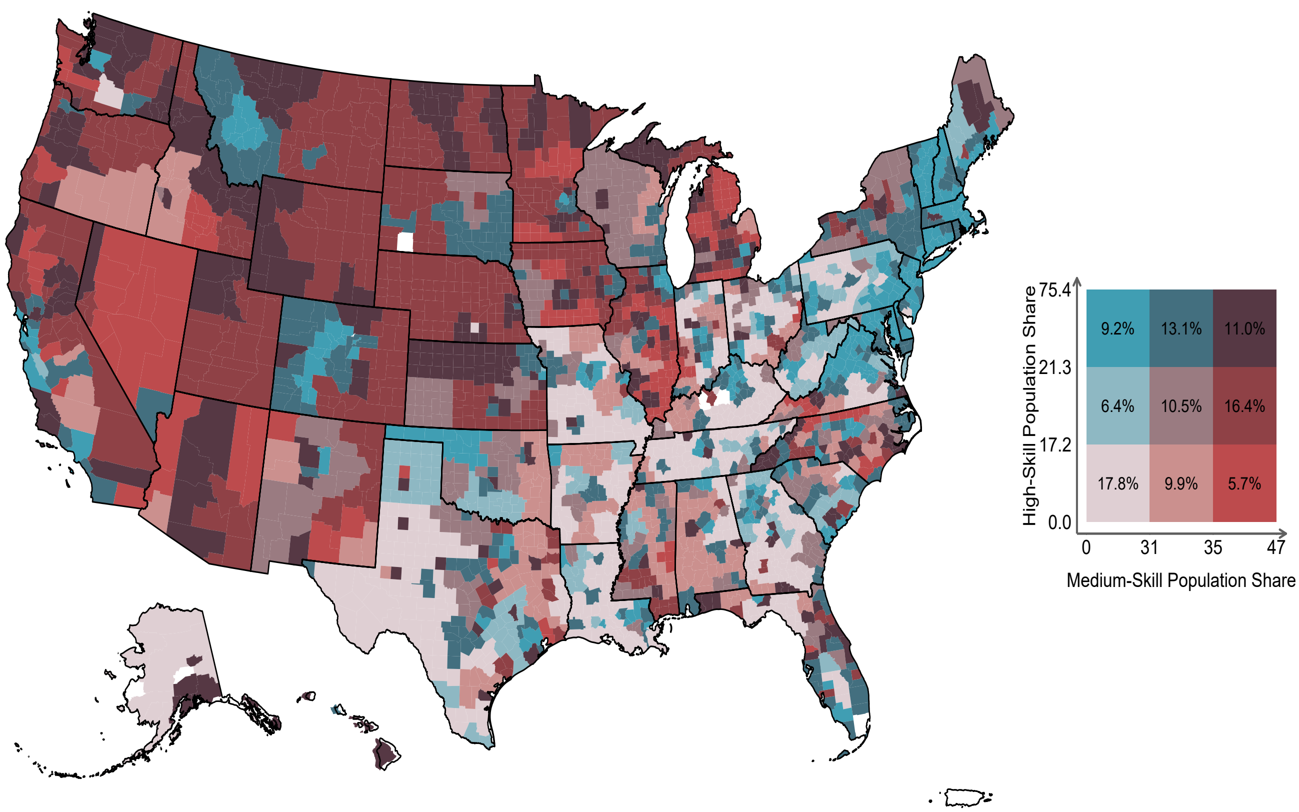

Figure D.3: The Bivariate Map of Low- and High-Skill Working-Age Population Shares in 2010

Notes: This bivariate map illustrates the distribution of medium- and high-skill working-age population shares across US counties in 2010. Lighter colors indicate lower concentrations, while darker colors represent higher concentrations for each respective skill type. The map has been plotted using the "`bimap`" package in Stata (Naqvi, 2024).

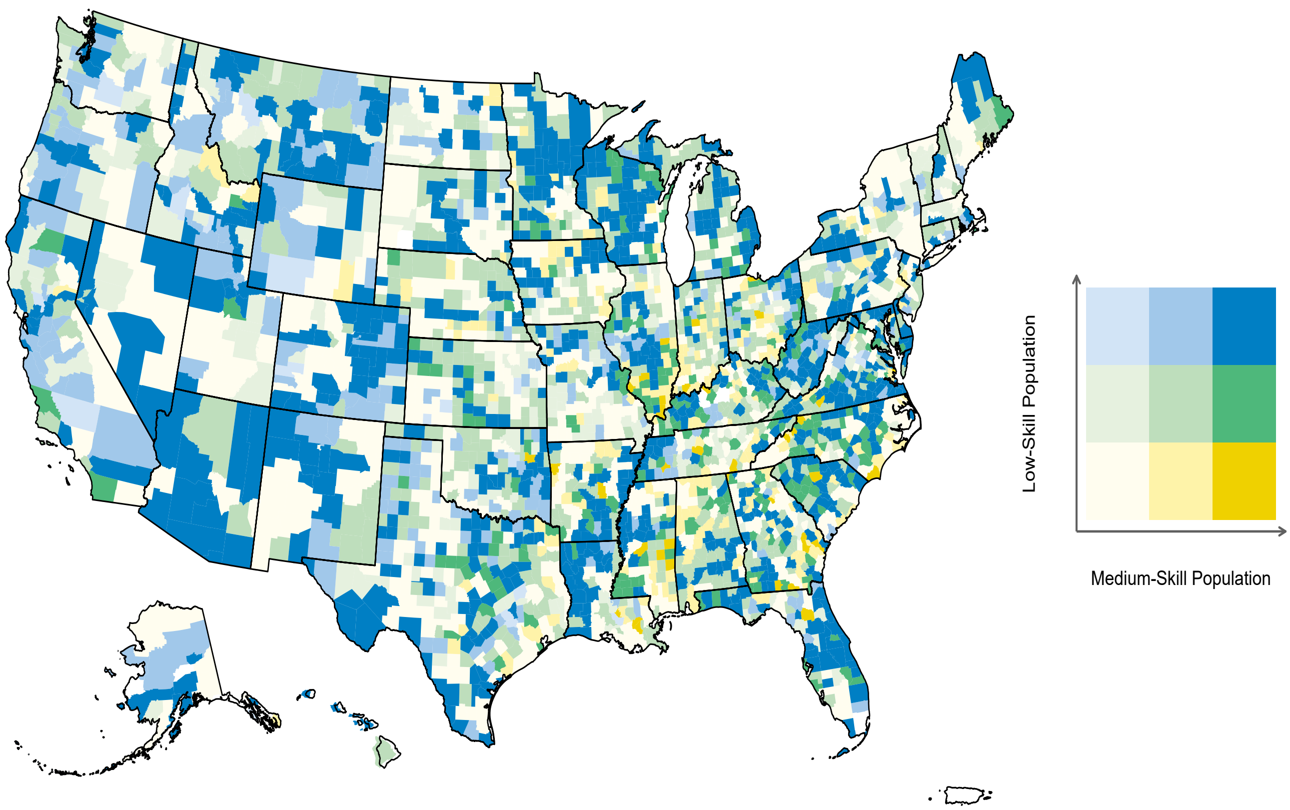

Figure D.4: The Bivariate Map of Low- and Medium-Skill Exogenous Population in 2010

Notes: This bivariate map illustrates the distribution of low- and medium-skill exogenous population across US counties in 2010. I regress each instrument on county and time fixed effects and obtain the residuals. The color coding visualizes the 200 quantiles of the residuals across counties. Lighter colors indicate lower quantiles, while darker colors represent higher quantiles for each respective skill type. The map has been plotted using the "`bimap`" package in Stata (Naqvi, 2024).

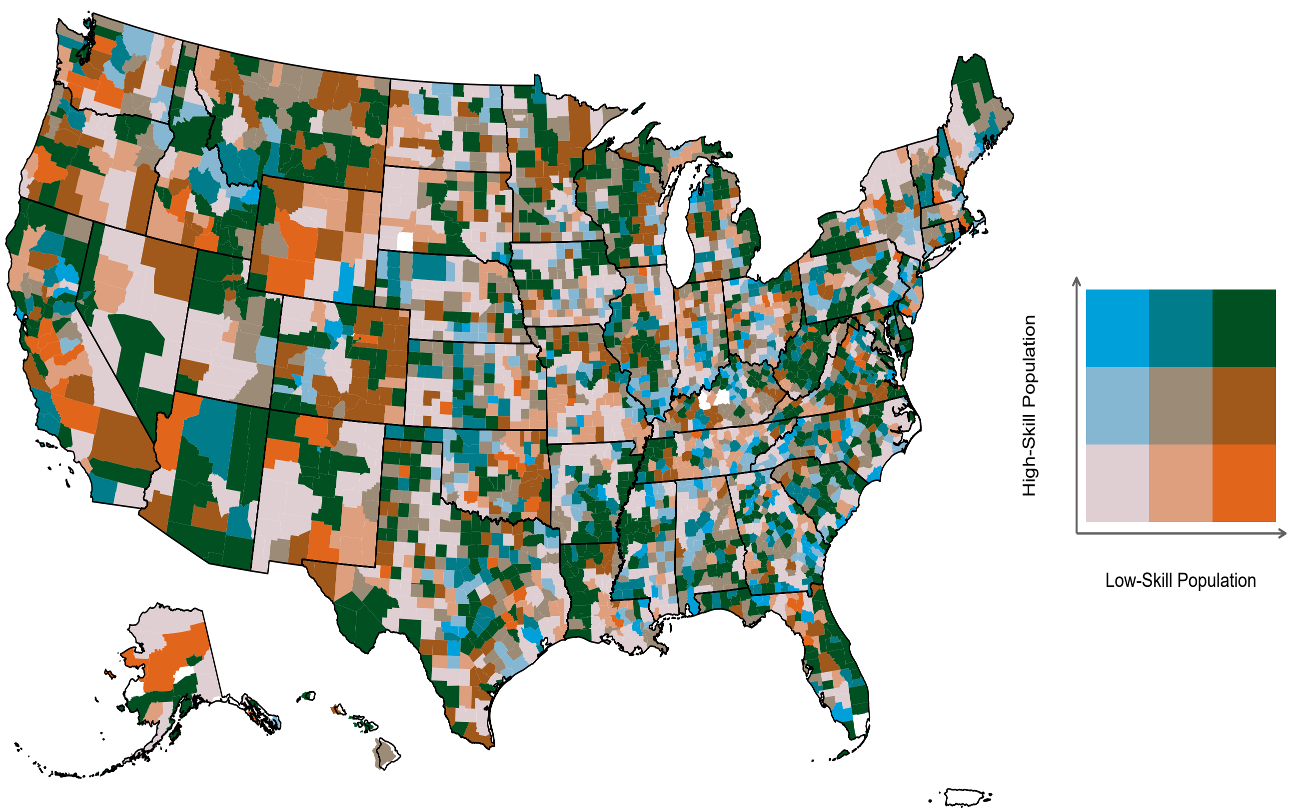

Figure D.5: The Bivariate Map of Low- and High-Skill Exogenous Population in 2010

Notes: This bivariate map illustrates the distribution of low- and high-skill exogenous population across US counties in 2010. I regress each instrument on county and time fixed effects and obtain the residuals. The color coding visualizes the 200 quantiles of the residuals across counties. Lighter colors indicate lower quantiles, while darker colors represent higher quantiles for each respective skill type. The map has been plotted using the "`bimap`" package in Stata (Naqvi, 2024).

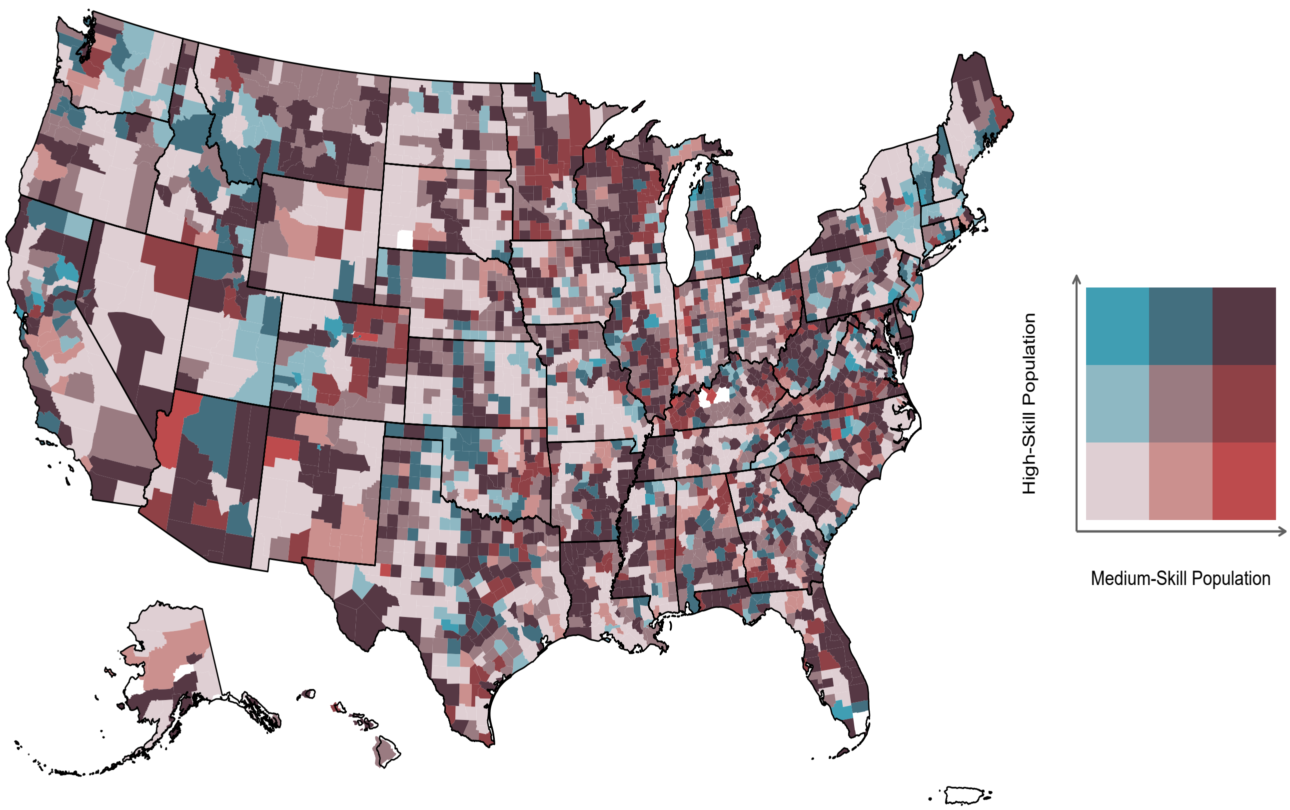

Figure D.6: The Bivariate Map of Medium- and High-Skill Exogenous Population in 2010

Notes: This bivariate map illustrates the distribution of medium- and high-skill exogenous population across US counties in 2010. I regress each instrument on state and time fixed effects and obtain the residuals. The color coding visualizes the 200 quantiles of the residuals across counties. Lighter colors indicate lower quantiles, while darker colors represent higher quantiles for each respective skill type. The map has been plotted using the "`bimap`" package in Stata (Naqvi, 2024).