Working Their Way Up? US Immigrants' Changing Labor Market Assimilation in the Age of Mass Migration

William J. Collins and Ariell Zimran
Vanderbilt University and NBER

Abstract

Whether immigrants advance in labor markets relative to natives is a fundamental question in immigration economics. It is difficult to answer this question for the Age of Mass Migration, when US immigration was at its peak. We document a substantial change in immigrant assimilation over this period: immigrants experienced substantial "catching up" relative to natives' occupational status from 1850 to 1880, but not from 1900 to 1930. We find that this change was not due to the large shift in immigrants' source countries. Instead, we show that it was rooted in a sizable change in natives' occupations that accompanied the structural transformation of the US economy.

1. Introduction

The 30 million European immigrants who entered the US in the century before 1921's closing of the "Golden Door" powerfully shaped American economic, demographic, and political development. Their experiences during this "Age of Mass Migration" form a cornerstone of broad interpretations of US history, especially regarding the opportunity for and realization of upward economic mobility. The idea that the US is (or was) the "land of opportunity" goes hand-in-hand with the idea that the US is "a nation of immigrants." These ideas, in turn, continue to inform contemporary views on immigration and social policy. A widely held and influential view is that European immigrants entered the US economy in low-paying occupations but experienced rapid economic and cultural assimilation despite facing discrimination and receiving little aid (Kennedy 1964; Smith et al. 2018).[1] Against this idealized historical benchmark, some question the capacity or willingness of many of today's immigrants to assimilate, and see this as a rationale for more restrictive and selective policies (e.g., Kelly 2018; Martin 2019).[2]

In this paper, we test the soundness of this view of historical immigrants' experiences. We show that assimilation patterns changed significantly over the Age of Mass Migration: European

---

[1] In the 1994 General Social Survey, 80 percent of respondents agreed or strongly agreed with the statement that, "The Irish, Italians, Jews, and many other minorities overcame prejudice and worked their way up. Today's immigrants should do the same without any special favors" (Smith et al. 2018).

[2] Ken Cuccinelli, at the time the acting head of US Citizenship and Immigration Services, stated that he would welcome immigrants "who can stand on their own two feet, be self-sufficient, pull themselves up by their bootstraps—again, as in the American tradition. My Italian-Irish heritage looks back at that. Most people in America look back at that. And that's what we expect going forward" (Martin 2019).

immigrants *did* narrow gaps in economic status relative to US natives in the mid-19th century, but later European immigrants did not. We also investigate *why* this pattern changed and find, surprisingly, that the dramatic shift of countries-of-origin within Europe (toward the South and East) had little to do with it. The answer instead lies in the deep structural transformation of the US economy between the middle and late 19th century.

Many scholars have studied immigrants' labor market outcomes and how they change with time spent in the US in both historical and recent settings. Indeed, whether and how quickly immigrants advance in labor markets relative to natives is one of the core questions in the economics of immigration (e.g., Borjas 2014, Abramitzky and Boustan 2017). The value of historical perspective on this question is widely recognized in this literature, but data limitations have obscured scholars' view of what actually happened between the 1840s and the 1920s. This period, on which we focus, predates the design and implementation of national longitudinal surveys or large-scale and representative administrative records. Instead, we conduct our analyses with individual-level census records that we link over time. To construct the datasets, we start with all white men ages 18 to 40 in the "complete-count" Census of Population records for 1850, the first year in which birthplace is recorded in the census; then, we use automated methods to search for the same men in the complete-count 1880 records by using information on their birthplace, year of birth, and name.[3] We repeat this approach with men in the complete-count records of the 1900 census, searching for their matches in 1930. The panel structure of the linked data avoids biases that arise in assimilation studies that use cross-sectional data, such as changing cohort quality and selective return migration (Chiswick 1978; Borjas 1985; Lubotsky 2007; Abramitzky et al. 2014). The datasets are sufficiently large that we can study outcomes for immigrants from specific source countries within the early and later immigrant cohorts. We show that our results are not sensitive to the linkage method we employ, which gives us confidence that we can reliably compare immigrants' and natives' lifecycle labor market outcomes over the full span of the Age of Mass Migration for the first time.

A challenge to studying historical labor market outcomes is that individual wage information is not available in nationally representative data sources. The census did not collect wage data until 1940 and even then omitted earnings for farmers and other self-employed workers. Therefore, we focus on workers' occupational status. We start by characterizing occupational status according to

---

[3] As described below, subsequent analyses focus on non-southern whites. We focus on non-southern men because few immigrants resided in the South, because the Civil War's impact on the South would confound comparisons over this period (e.g., the application of occupational rankings that rely on post-Civil War data to the pre-Civil War South), and to maintain comparability to prior studies.

broad and simple categories—unskilled, operative, craft, farmer, and white collar. Then, we make full use of the censuses' detailed (three-digit) occupational information to rank workers using two different approaches. One approach relies on the complete-count records of the 1870 census to calculate the average level of wealth for men in each detailed occupational category; this "occupational score" is then assigned to each worker in our linked data according to his current occupation and transformed into a ranking. A second approach relies on average income circa 1900 for occupations reported in Preston and Haines (1991), which in turn relies heavily on Commissioner of Labor (1903), Douglas (1930), and Lebergott (1964). We supplement that information with an estimate of white farmers' average income using the Census of Agriculture in 1900.

As with other metrics of occupational status, which are commonly used in studies of historical economic mobility and immigrant assimilation, our baseline rankings will capture changes in status associated with changes in occupation but not changes in economic status *within* occupation. Later in the paper, however, we allow for differential changes in status within occupations over the lifecycle. This addresses the concern that farmers (or others) might have accumulated wealth and improved their relative status without changing occupations, which could confound results based on a fixed occupational score. Our results and interpretations are not sensitive to this alternative.

Our first main finding is that there was an important change in immigrants' occupational assimilation patterns over the Age of Mass Migration. The cohort of immigrants observed in 1850 and 1880 experienced substantial occupational upgrading relative to native-born men, reducing the gap in occupational status relative to natives from about 14 to 7 percentiles. When we examine a later cohort of immigrants observed in 1900 and 1930, we find little or no occupational upgrading relative to natives. This latter result is consistent with Abramitzky et al. (2014), who study 1900 to 1920, and it is reassuring since we employ different methods for linking the census records and assessing occupational status. Our finding for the mid-19th century is new to the literature, as is the comparison that it allows us to make with data from the 20th century. In sum, whereas the 19th-century patterns support the view that immigrants started behind natives in labor markets but gained ground with time, the early 20th-century experience was quite different. In this sense, the first part of the paper entails an illuminating expansion of scope relative to Abramitzky et al.'s (2014) work on the 20th century, one that juxtaposes and contrasts the early and later decades of the Age of Mass Migration, revealing and examining the changing patterns of immigrant assimilation between them.

The changing patterns of immigrants' outcomes relative to natives require investigation, and we consider several possibilities. The paper's second main finding is that the difference in immigrants' upgrading relative to natives in the 19th versus 20th century was *not* the product of

sharp changes in immigrants' source countries within Europe, shifting away from the North and West and toward the South and East. This finding is important in light of influential claims by social scientists, commentators, and policymakers circa 1900 that immigrants from the new sources were "inferior" to those who had previously come from the old sources.[4] Such claims helped to motivate the design and implementation of immigration restrictions in the 1920s that were severely biased against potential immigrants from eastern and southern Europe and remained in place for most of the 20th century (Walker 1896; Commissioner-General of Immigration 1903; Zolberg 2006; Benton-Cohen 2018). Instead, we show that the difference in immigrants' relative upgrading between the early and later parts of the Age of Mass Migration was rooted in sizable changes in the occupational distribution of young *natives* between 1850 and 1900. In 1850, natives in our sample were concentrated in farming whereas immigrants were largely engaged in unskilled labor. We show that men engaged in unskilled labor in 1850, whether immigrant or native, experienced high rates of occupational upgrading, whereas those engaged in farming did not. By 1900, the US economy's structural transformation away from agricultural employment had collapsed differences in occupations between immigrant and native workers under age 40. That is, there were far more young natives in "unskilled" jobs in 1900 than before. Natives in *both* periods upgraded at least as much as immigrants did *conditional* on the initial year's occupational category. Thus, the smaller immigrant-native differences in initial-year occupations (in 1900 compared to 1850) were associated with diminished average upgrading for immigrants relative to natives in the later period. In this sense, the "catching up" of immigrants in the 19th century is attributable to their arrival at a time when natives were primarily employed in farming, not to their skills or county-of-origin heritage.

Our finding that immigrants in both periods (early and late in the Age of Mass Migration) failed to keep pace with natives who started in similar occupations is novel and contradicts the notion that European immigrants were at first under-placed in labor markets relative to natives with similar productive abilities and then gained ground rapidly with the accumulation of US-specific human capital (e.g., language skills). In fact, we show that even though immigrants in the early 20th century did improve their language skills, this did not translate into occupational gains relative to natives.

This paper advances the study of immigration in several ways. It builds new panel datasets that provide a clear and wide window on immigrants' lifecycle labor market outcomes, spanning the entire Age of Mass Migration. It uncovers new evidence of a *change* in immigrants' labor market assimilation patterns, revising our understanding of the facts of economic assimilation in US history.

---

[4] More specifically, the "old sources" were primarily Germany, Britain, and Ireland. The "new sources" were primarily Italy, the Russian Empire, and the Austro-Hungarian Empire.

And it provides an explanation for *why* these patterns prevailed and changed over time, emphasizing the economy's structural transformation and rejecting an alternative explanation that motivated ethnically biased immigration policies in the early 20th century. At a time when an oversimplified view of European immigrants' experience is held up as justification for US policy, these insights are timely and instructive. They challenge prevalent narratives about European immigrants' economic mobility in America—myths that continue to influence modern debates over immigration policy.

2. Immigrants in the US Labor Market during the Age of Mass Migration

Figure 2 (omitted for space) shows the share of white male natives and European immigrants, ages 18 to 65, working in two broad occupational categories—unskilled labor and farming—from 1850 to 1930.[5] The predominance of immigrants in unskilled work and natives in farming in 1850 is immediately apparent, as is the secular decline in farming as a share of the native labor force (and of the immigrant labor force after 1880). From 1850 to 1880, there is convergence in the shares of each group working in farming, and by 1930, the gaps between immigrants' and natives' shares in farming and unskilled jobs were smaller than in the mid-19th century. But these patterns of convergence are difficult, and possibly misleading, to interpret as evidence of labor market assimilation. Due to the continuous inflow and outflow of immigrants, cross-sectional data sources provide little useful insight into how immigrants who stayed in the US fared over their lifecycle. This has been a central challenge to studying immigrants' labor market outcomes during the Age of Mass Migration.

Data shortcomings, of course, did not forestall policy debates and decisions centered on immigrants' assimilation in the labor market and in other social dimensions. US natives' concerns over immigrants' poverty, lack of skills, insularity, and religion are at least as old as the onset of mass immigration. For instance, the influx of poor Irish immigrants in the 1840s and 1850s triggered an intensification of nativism and anti-Catholicism, epitomized by the rise of the "Know-Nothing" Party (Higham 1955; Anbinder 1992; Alsan et al. 2020). Later in the 19th century, the shift in source countries to southern and eastern Europe renewed vigorous debates over immigration policy (Goldin 1994; Williamson 1998; Benton-Cohen 2018). By this time, a common argument in favor of restricting European immigration involved unfavorable comparisons of the "new immigrants" to the "old immigrants," where "new" and "old" referred to place of origin. In 1896, Francis A. Walker, the first president of the American Economic Association, characterized the "new" immigrants as "beaten men from beaten races; representing the worst failures in the struggle for existence" (Walker

---

[5] The unskilled category includes service workers, farm laborers, and laborers.

1896). The Commissioner-General of Immigration (1903, p. 70), William Williams, bemoaned the "new" immigrants' urban settlement patterns, viewing them as evidence of the immigrants' low quality: "Notwithstanding the well-known demand for agricultural labor in the Western States, thousands of foreigners keep pouring into our cities, declining to go where they might be wanted because they are neither physically nor mentally fitted to go to these undeveloped parts of our country and do as did the early settlers from northern Europe."

In 1907, Congress authorized the era's most influential study of immigrants—the Dillingham Commission Reports. The Commission's investigation was unprecedented in scope and detail, including the collection of data on immigrant and native workers in a variety of industries. The Commission's policy recommendations included a literacy test, source-specific quotas, and continuing bans on Asian immigration (US Congress 1911). Hourwich (1912) and Douglas (1919) published critiques of the reports and of popular summaries of the Commission's work, most notably Jenks and Lauck's (1912) book, *The Immigration Problem*. They highlighted the limits of interpretation from cross-sectional data, especially in comparing "new" and "old" immigrants who had been in the US for different lengths of time. Nonetheless, several of the Commission's main recommendations eventually became law (Benton-Cohen 2018).

Modern econometric analyses have produced mixed evidence as to whether 19th-century immigrants improved their status relative to natives as they gained experience in the US labor market. Blau (1980), for instance, drew on cross-sectional data tabulated and published by the Dillingham Commission and concluded that immigrants, especially from "new" sources, initially earned less than natives, but caught up over time. Later studies shifted attention to cross-sectional, micro-level datasets, which include wage information for workers in specific industries and states circa 1890. Hannon (1982), for instance, concluded that immigrants in Michigan's mining industry slowly caught up to those of natives, whereas Eichengreen and Gemery (1986) and Hanes (1996) concluded that immigrants did not gain in earnings relative to natives. In contrast, after allowing more flexible functional forms in age-earnings profiles for workers in Michigan and California, Hatton (1997) concluded that immigrants who arrived as adults started at a lower wage than similar natives, gained substantially in relative pay with age, but did not fully catch up to natives. Minns (2000) shifted attention to census microdata for 1900 and 1910, which provided a more geographically comprehensive view of immigrants and natives in the labor market, albeit without wage information. He found evidence consistent with occupational gains for immigrants relative to natives, at least outside the farm sector.

More recently and most closely related to our work, Abramitzky et al. (2014) extended

methods pioneered by Ferrie (1994, 1995, 1997, 1999) to construct a panel dataset of immigrants and natives between 1900 and 1920.[6] Focusing on average immigrant-native differences in occupational status and their evolution with immigrants' time in the US, they found almost no differential gains for immigrants, attributing prior studies' findings of immigrants' labor market assimilation in this period to data deficiencies that are addressed by the use of panel data.

Relative to the work of Abramitzky et al. (2014), we have added new analyses of mid-19th century immigrants and natives, which entails a significant expansion of coverage and reveals a fundamental *change* in overall assimilation patterns. The new datasets also enable us to make an important advance by explaining *why* patterns of occupational upgrading changed. Although our basic approach is similar in spirit to Abramitzky et al. (2014) and owes much to them, we linked the data using updated methods (but also linked it using their method to check robustness), worked from the complete-count census records in the base year (resulting in larger samples), and focused on occupational ranks derived from different approaches to measuring occupational status. Taken together, this enables us to provide consistent and comprehensive (all industries, all non-southern states) analyses of immigrants' and natives' labor market outcomes over both the early decades of the Age of Mass Migration, which were dominated by immigration from "old" source countries, and the later decades, which were dominated by immigration from "new" sources.

3. Building new datasets
Linking census records

We began by extracting all US- or European-born men from the complete-count records of the 1850 US census (Ruggles et al. 2015; Minnesota Population Center 2017). We then searched for the same men in the complete-count records of the 1880 census (Ruggles et al. 2015; Minnesota Population Center 2017), using linkage methods described below and in more detail in Appendix A. For men found at both points in time, we can observe and compare their occupations; this forms the basis for our analyses of the occupational status of immigrants relative to natives. The 1850 census was the first to record place of birth (thus identifying immigrants) and was implemented several years after the onset of mass migration from Europe. The 30-year span (to 1880) allows sufficient time to observe occupational upgrading over the lifecycle. We focus our subsequent analyses on non-southern white men because few immigrants located in the South in this period and because the Civil War's impact on the South may confound comparisons that span the war (e.g., applying an

---

[6] Ferrie focused on immigrants arriving in New York City in the mid-19th century and documented their occupational changes and wealth accumulation. He did not directly compare their upgrading to that of natives.

occupational ranking based on post-Civil War data to the South in 1850 would be tenuous). Our second dataset covers the early 20th century, linking men from the 1900 to the 1930 census records using the same method (Ruggles et al. 2019). The 1900 starting point allows us to include a large number of "new" immigrants, and the 1930 end point observes outcomes in the first census taken after the imposition of quotas but just before the full effects of the Great Depression.

The panel structure of the linked datasets is essential to study changes in immigrants' occupational status. Abramitzky et al. (2014) show that using either cross-section or repeated cross-section data can generate spurious evidence of labor market assimilation due to changing arrival-cohort quality and negatively selected return migration. As in their work, our panel datasets avoid potential confounds by limiting attention to a fixed sample of men who remained in the US. This design is appropriate for studies that are primarily interested in whether immigrants who stayed in the US experienced relative improvements in labor market status or other aspects of assimilation. Note, however, that the linked census data might not capture the lifetime experience of the "average immigrant" given that many returned to their country of origin or died before being observed for a second time (Bandiera et al. 2013); we return to this issue below.

Our record linkage procedure follows that of Collins and Zimran (2019), which belongs to a class of methods derived from Ferrie's (1996) pioneering work. In brief, we began by eliminating from consideration men who are not unique in the base-year census (1850 or 1900) on their combination of name, birthplace, and birthyear (+/- 4 years). We then searched for each remaining man in the later-year census (1880 or 1930), allowing for inexact matches on the basis of name and birth year. We retained only unique matches for the analyses of occupations. This approach incorporates several features highlighted by Bailey et al. (2020) and Abramitzky et al. (2019) as useful in improving match quality, such as the use of string distance measures to compare names,[7] the elimination of non-unique individuals in the base year before linkage, and the restriction to cases in which a unique match is made between the initial and final censuses.

Table 1 (omitted for space) presents rates of successful linkage separately for immigrants and natives.[8] The linkage rates relative to the unique sub-sample are 25.7 percent for natives and 13.7 percent for immigrants. The former is approximately comparable to the rate of linkage relative to unique men achieved by Ferrie (1996). The latter is somewhat lower. The lower rate for immigrants likely reflects the large number of immigrant arrivals after 1850, which may prevent successful links

---

[7] Specifically, we use the SAS function SPEDIS, which is described in detail by Gershteyn (2000). This string distance measure is similar to the Jaro-Winkler distance but treats different types of edits differently.
[8] We attempt to link all males regardless of age and make the restriction to 18-40 year olds afterwards.

even for those who were unique in 1850. The linkage rates relative to the entire base population are 11.5 percent for natives and 5.7 percent for immigrants. These are lower than those achieved by other algorithms derived from Ferrie (1996) in other contexts (e.g., Beach et al. 2016). But as shown in Appendix F, the linkage rates are similar to those achieved using alternative methods described in detail by Abramitzky et al. (2019). Subsequent rows in Table 1 describe the linkage rates for 1900 to 1930, which are somewhat higher than in the 1850-80 data.

Several concerns arise in the use of linked census data. First, non-random selection into linkage is common in such studies. For instance, individuals with higher human capital may be more likely to survive and report consistent information to census enumerators, and name uniqueness may vary by country of origin or other characteristics.[9] Figure 3 (omitted for space) compares observable characteristics in the linked and full samples of census data. It reveals some evidence of positive selection into the linked sample, such as the over-representation of individuals with greater 1850 wealth holdings. To address this, we reweight the sample to make the distribution of 1850 observables in the linked sample match (as closely as possible) their distribution in the sample of men in 1850 who we attempted to link. Given the higher rates of return migration in this period (Bandiera et al. 2013), we weight the linked 1900-30 immigrant sample to look like the 1930 population of immigrants, aged 44-74, who reported arriving in the US before 1900 (rather than to resemble the 1900 sample that we attempted to link) and similarly reweight the 1900-30 native sample to match the 1930 native population aged 44-74.[10]

Second, some of the matches in the linked samples may be false positives, perhaps leading to spuriously high estimates of the probability of changing occupations over time. From the start, we adopted relatively strict match criteria to limit false positives with this concern in mind. Nonetheless, in Appendix E, we repeat our analysis using subsets of the linked samples that are limited to exact matches only; the results are qualitatively similar. Finally, Appendix F repeats our results using alternate linkage methods. Again, the results are qualitatively similar.

Studying occupations to learn about labor market assimilation

---

[9] Consistent with this concern, Appendix Tables A.1, A.2, and A.3 report rates of successful linkage for a variety of European source countries and reveal considerable variation in the rates of successful linkage.
[10] Weighting the linked 1900-30 sample to match the 1900 population of immigrants would make the (linked) sample of stayers match the average characteristics of the combined groups of stayers and return migrants. This is undesirable because our focus is on stayers, who are likely different from return migrants. Because 1880's census did not ascertain year of arrival, we cannot weight the 1850-80 sample in the same way as the 1900-30 sample; but the low rates of return migration in the early period mitigate our concern. We discuss the use of 1900-based weights for the 1900-30 sample later in the paper.

We focus on the occupational status of immigrants relative to natives. This is primarily because the US census did not inquire about individuals' incomes until 1940, and even then it did not cover proprietors' or farmers' incomes. Occupational information, on the other hand, was collected in each census beginning in 1850, enabling our analysis of individual-level occupational changes in both the 19th and the 20th centuries in a fairly consistent manner. We rely on the detailed three-digit occupational coding provided by Ruggles et al. (2015, 2019) and the Minnesota Population Center (2017). We begin by using the codes to define five broad classes of occupations—farmer, white collar, craft, operative, and unskilled.[11] Our preliminary analysis entails simple comparisons (controlling only for age) of the shares of immigrants and natives in each occupational category and of changes in those shares over the men's lifecycle. This approach has the advantage of not requiring any assumptions about the relative standing of occupations.

We then introduce rank measures of occupational status. The literature has often relied on an occupational income score based on information from the 1950 census. We prefer not to use this measure because of its temporal distance from our study period, particularly in the early portions. Instead, we construct several alternative occupational scores. The first is based on the 1870 census questions about real and personal wealth (Collins and Zimran 2019; Ager et al. 2019; Craig et al. 2019). We use the complete-count sample of the 1870 census (Ruggles et al. 2020) to compute average total wealth (the sum of real and personal wealth) within each detailed (three-digit) occupational cell for men aged 30 to 65. This produces an occupational wealth score that is similar in spirit to the occupational income score commonly used in studies of historical labor markets, including immigrants' occupational assimilation. We use this score in two ways—directly (in logarithmic form) and to compute occupational ranks.[12] The ranks are expressed relative to all non-southern white men in the US labor force (ages 18-74 with reported occupations). Thus, for any given man, a change in rank over time depends on whether he changed occupations and on how the occupational distribution of all men in the labor force changed. We prefer the rank measure over the score because it is less sensitive to specific cardinal differences between scores.

As a robustness check, we also derive wealth scores based on age groups *within* detailed

---

[11] The broad occupational categories are defined using IPUMS occ1950 codes: white collar [1, 100) and [200, 500); farmer [100, 200); craft [500, 600); operative [600, 700); unskilled [700, 970]. The unskilled category includes service workers, farm laborers, and laborers. The 1900 and 1930 data include a significant fraction of individuals whose occupation codes are "Not Yet Classified" (occ1950 code 979). Our baseline analysis omits these individuals. In Appendix D, we classify these individuals by using the modal classification of the NYSIIS code of their occupational string (similar to Collins and Wanamaker 2020) and repeat the main analysis. Our results are qualitatively unaffected by the inclusion of these individuals.

[12] In the case of ties, we assign each individual in a category the average rank of the group.

occupational categories in 1870. This allows changes in scores over the lifecycle even for those who do not change occupations, and addresses (among other concerns) the possibility that individuals, especially farmers, might improve their economic status substantially without changing occupations. Specifically, we compute separate scores for individuals aged 18-40 and those aged 41-70 in the 1870 full count data. Since we observe individuals in the linked census data once when they are aged 18-40 and then again when they are aged 48-70, this approach to assigning scores allows changes in rank between the two census dates to reflect life-cycle changes in wealth holdings.[13]

We create another occupational income score using information that Preston and Haines (1991) compiled from Commissioner of Labor (1903), Douglas (1930), and Lebergott (1964). This provides estimates of annual income by detailed occupation circa 1900. Preston and Haines do not provide an estimate for farmers, so we used the 1900 Census of Agriculture to estimate average net income for farmers based on the reported value of farm production, cost of inputs, and estimates of in-kind income. Again, we use this measure directly (in logarithmic form) as well as to compute a ranking, as with the 1870-based score. Unlike the 1870 wealth-based score, data are not available to modify the wage-based score to capture changes over the lifecycle.

Finally, we also use the two ranking schemes to create an average occupational rank, which is the simple average of ranks based on the two occupational scores discussed above. This is our preferred measure for making consistent comparisons across the 1850-80 and 1900-30 periods. The main results are qualitatively insensitive to which ranking scheme is used.

4. Empirical strategy and summary statistics

Empirical strategy

In our analysis, our initial focus is on immigrants' "unconditional assimilation"—that is, we ask whether the average occupational status of immigrants improved relative to that of natives over the 30 years in which we observe each cohort, adjusting only for differences in age. This approach, which also underlies Abramitzky et al.'s (2014) analysis, answers a simple but fundamental question about the Age of Mass Migration, and it shows how immigrants' and natives' relative occupational status changed with time in the US. By comparing these patterns across the two time periods that we study, this analysis shows how these assimilation patterns changed as the US economy transformed

---

[13] Dividing into these "young" and "old" age cells is a simple but effective way of addressing concerns over different upgrading over time within occupations. In principle, we could create finer cells, such as birthplace-specific cells. But doing so would prevent us from distinguishing between immigrants who had been in the US for some time from recent arrivals (the 1870 census did not record time of arrival), thus eliminating one of the gains from the use of panel data.

into an urban and industrial powerhouse.

The 19th-century US censuses never inquired about immigrants' year of arrival. Thus, it is not possible to use that information in comparisons of the early and late phases of the Age of Migration. Later in the paper, we discuss whether differences in the distribution of time-in-the-US might confound comparisons of the early and later periods, and we conclude that this is unlikely.

We estimate a series of regressions of the form

$$y_{it} = \beta_t F_i + g(a_{it}) + \varepsilon_{it} \tag{1}$$

where y_{it} is an outcome for individual i in year t, F_i is an indicator for foreign birth, and $g(a_{it})$ is a quartic function in individual i's age in year t. We first describe differences in occupational categories, then in occupational scores and ranks, and finally differences in some rough measures of human capital. Although human capital is an input into labor market outcomes, we provide the separate analysis for a direct view on differences in basic skills that workers bring to the market, including the ability to speak English, which has long been central to discussions of immigrants' assimilation (e.g., US Congress 1911, Borjas 2015).

We estimate equation (1) separately for each outcome for each census year t. The coefficient β_t represents the difference in occupational status between natives and immigrants in that census year, conditional on age. To measure changes between census years for each cohort, we estimate equation (1) twice by seemingly unrelated regression, once using data from the early census year and once using data from the later census year. A measure of assimilation is then given by $\beta_t - \beta_{t-30}$. The difference will be positive in cases where immigrants increased average levels of y relative to natives over time, indicating assimilation in cases where immigrants were initially behind natives.[14]

We also use estimates of equation (1) for the five occupational category measures to compute dissimilarity indices between the two groups according to

$$D_t = \sum_k \frac{1}{2} |\hat{\beta}_{tk}|, \tag{2}$$

where $\hat{\beta}_{tk}$ is the estimated coefficient from estimating equation (1) with occupational category indicator k as the dependent variable. This is analogous to the standard dissimilarity index (Duncan and Duncan 1955), but the difference in the fraction of each nativity group in each occupational category is estimated while controlling for the different age structures of the two groups.

After characterizing assimilation patterns using these measures, we compare immigrants'

---

[14] For instance, if immigrants start off with a lower occupational rank than natives, then a positive value of $\beta_t - \beta_{t-30}$ indicates that this gap closed over thirty years, implying that immigrants were more like natives (so long as it is not large enough for immigrants to overtake natives).

occupational upgrading to that of natives who start in occupations with similar economic status by estimating equations of the form

$$R_{it} - R_{it-30} = \delta F_i + \gamma R_{it-30} + g(a_{it-30}) + \eta_i \qquad (3)$$

where $R_{it} - R_{it-30}$ is the change in worker's average rank measure between the two periods, R_{it-30} is the worker's rank in the initial year based on the detailed occupational coding and scoring described above, and all other notation is the same as in equation (1). The coefficient δ measures the average upgrading of immigrants relative to natives, conditional on the worker's initial occupational status and a quartic in age. These "conditional assimilation" results enable us to determine whether changes in assimilation (or lack thereof) documented by estimating equation (1) reflect changes in lifetime mobility or in initial occupational distributions.

Summary statistics

Figure 4 depicts the broad occupational distributions for men in the linked samples for easier comparison. In 1850, immigrants and natives had quite different occupational distributions; the modal immigrant held an unskilled job and the modal native was a farmer. Immigrants were more likely than natives to hold craft or operative occupations, whereas natives were more likely to hold white-collar occupations. In 1880, the occupational distributions retained these properties, but both immigrants and natives experienced considerable movement *into* the "farmer" category over their lifecycles (often upgrading from farm laborer).

In contrast, moving to the later cohorts, in 1900 the modal immigrant and native both held unskilled occupations. By 1930, both groups had reduced their presence in unskilled work. Immigrants were over-represented in craft and operative labor, while natives remained over-represented in white collar and farming occupations. Comparing the 1850 and 1900 occupational distributions highlights changes in the American economy that will have important implications later in our analysis, especially the decline in the share of natives who were farmers and the rise in the share holding unskilled occupations.

5. Unconditional gains in occupational status in the 19th and 20th centuries

Table 3 presents estimates of equation (1) for broad occupational categories. Columns (1) and (2) pertain to 1850 and 1880, respectively, and column (3) presents the measure of assimilation ($\beta_t - \beta_{t-30}$). Columns (4) to (6) present analogous figures for the 1900-30 cohort. The starkest evidence of convergence in occupations between immigrants and natives comes from the first two rows for the

13

1850-80 cohort. In 1850, the foreign born were 23.8 percentage points more likely to hold an unskilled occupation than were natives; by 1880, the gap had shrunk to only 12.8 percentage points. Meanwhile, natives were 27.8 percentage points more likely than immigrants to be farmers in 1850, but by 1880 the gap had declined to 18.7 percentage points. Gaps in white collar and craft occupations declined only slightly, and the gap slightly widened in operative jobs.

In 1900-30, natives and immigrants began with a statistically indistinguishable probability of holding an unskilled occupation (conditional on age);[15] by 1930, however, the foreign-born were 4.6 percentage points more likely to hold an unskilled occupation. Immigrants narrowed the gap in farming, but only by a small amount (from 10.7 to 8.4 percentage points), and they did not narrow the gap in white-collar work. Overall, Table 3 reinforces the visual impression from Figure 4: the 1850-80 period was characterized by wide differences in initial broad occupations and by strong convergence, whereas the 1900-30 period was characterized by comparatively smaller initial differences and limited convergence.

In Table 4, we summarize the occupational category differences in the form of the dissimilarity index computed as in equation (2). Specifically, the table reports dissimilarity indices for each year based on the five major categories defined above and on the estimates of Table 3. The statistic 0.31 for 1850 in column (1) indicates that 31 percent of immigrants would have to change their occupational category to match the occupational distribution of natives, or vice versa. The dissimilarity index declined from 0.31 to 0.21 between 1850 and 1880, reflecting the considerable occupational assimilation visible in Figure 4 and Table 3. On the other hand, although there is a slight decline in the index between 1900 and 1930, from 0.15 to 0.13, it is much smaller than in the earlier period, supporting the conclusion from Table 3 of limited assimilation in the 20th century. It is notable that the index in 1900 is already at a much lower level than for the earlier cohort in 1850 or even in 1880—the differences between natives and immigrants had greatly compressed. Indeed, it is plausible that the rising tide of calls for immigration restriction reflected the more direct contact between immigrants and natives in labor markets (Goldin 1994; Hatton and Williamson 1998).

It is possible that the broad occupational categories examined to this point obscure important variation within each category; moreover, such broad categories provide a limited sense of whether men move up the occupational ladder with time. Therefore, we incorporate more detailed occupational definitions into the analysis and rank occupations based on what is known about the average economic status of men holding them. Table 5 repeats the estimation of equation (1) for the

---

[15] Note that the adjustment for age eliminates the 1900 difference in unskilled work that appears in Figure 4.

occupational income scoring and ranking schemes described above. Column (3) again reports *changes* in immigrants' status relative to natives over time, with positive numbers indicating that immigrants advanced faster than natives. Substantial assimilation for the 1850-80 cohort is again evident, with the immigrant-native gap in average occupational rank closing by nearly 7 percentiles; using the 1870 wealth-based scoring, the gap narrows by 8 percentiles. All methods of scoring and ranking indicate substantial labor market assimilation in the 1850-80 dataset. In contrast, for the 1900-30 cohort, there is evidence of slightly *widening* gaps in status using every metric.

In Appendix B, we rule out several concerns regarding the robustness of the results described above. In particular, we show that our results are not meaningfully affected by excluding second-generation immigrants in the sample of natives, immigrants in the 1900-30 sample who had been in the US for more than 10 years, or men aged 31-40 in the initial year. In Appendix B, we also use occupational scores that are based on the wealth holdings of younger (18-40) and older (41-70) men within each detailed occupational category in the 1870. For example, if the age-wealth profile were steeper among farmers than other occupations, this approach would allow that fact to register in the assignment of scores and ranks. The results are similar to those from our baseline analysis.[16]

6. Conditional gains in occupational status in the 19th and 20th centuries

The changing pattern of unconditional upgrading between the early and later cohorts in the Age of Mass Migration raises the question of whether the results primarily reflect changes in natives' or immigrants' occupational mobility or changes in their initial distribution of occupations, each of which had a different likelihood of leading to upward mobility. To better understand the empirical underpinnings of the unconditional results, we estimate equation (3), presenting results in Table 7.

Columns (1) and (2) of Table 7, which omit the initial rank variable (R_{it}) from the right-hand side, simply reiterate the findings from section 5—that immigrants achieved greater occupational upgrading than natives in the period 1850-80 but somewhat slower upgrading in the period 1900-30. This provides a benchmark to which subsequent estimates can be compared. Columns (3) and (4) add the control for initial rank. From this perspective immigrants' upgrading relative to natives was remarkably similar—*and negative*—over the two periods within the Age of Mass Migration. If anything, immigrants in the later period (-0.03) fared slightly better than in the earlier period (-0.04).

---

[16] Although this changes only the 1870-based scores (and not in the occupational category definition), using these scores might lead to different results for other measures of status, such as occupational category, because these scores factor into the creation of weights to correct for selection into linkage. It is notable that not only the results based on the average occupational rank (which includes the non-age-varying 1900-based rank), but also the ranks based only on 1870 show very similar results to those of the main text.

In both periods immigrants upgraded less than natives who were comparably positioned in the base year. This runs contrary to the notion that immigrants were initially under-placed in the labor market but gradually erased that disadvantage with the acquisition of US-specific experience. It also implies that the overall immigrant-native convergence observed in the mid 19th century occurred despite natives' advantage in upgrading conditional on initial occupation.

Figure 5 provides a simple visual summary of the immigrant-native differences in upgrading conditional on initial occupational status. In this figure, we group men by broad categories for their base-year job, but changes in status are calculated using the detailed occupational information. The figure plots the average change in rank for men in each (broad) occupational category, expressed relative to that for natives who held white-collar occupations in the base year; that is, zero corresponds to the average upgrading for natives who were initially in a white-collar occupation. Three patterns stand out. First, average upgrading from the "unskilled" category tends to be larger than from other categories. Second, immigrants experienced less occupational upgrading than natives in almost every case, especially in the 1850-80 cohort. Third, relative to white-collar natives, upgrading throughout the occupational distribution was greater in 1850-80 than in 1900-30.

We can also make comparisons across the early and later periods within the Age of Mass Migration. In column (5) of Table 7, we re-weight the 1850-80 sample so that its initial occupational distribution matches that of the 1900-30 sample for natives and immigrants separately, but retaining the 1850-80-specific occupational upgrading. Column (6) weights the 1900-30 sample to match the occupational distribution of the 1850-80 sample. Between these periods, young native workers experienced a large change in their occupational distribution: the age-adjusted dissimilarity index between natives' 1850 and 1900 occupational distributions is 0.26, a difference almost as large as that between natives and immigrants in 1850 (in Table 4). Immigrants, on the other hand, had a relatively small change in occupational distributions between 1850 and 1900: the age-adjusted dissimilarity index is only 0.08. The dramatic change in the initial occupational distribution of natives was driven primarily by a shift of native labor out of farming and into unskilled labor. Whereas 41.8 percent of natives were farmers in 1850's linked sample, only 18.2 percent were farmers in 1900's linked sample (a shift of 23.6 percentage points); conversely, whereas only 15.2 percent of natives were in unskilled occupations in 1850's sample, 38.9 percent were by 1900 (an almost identical shift of 23.7 percentage points). As we discuss below, this occupational shift is underpinned by the structural transformation of the US economy over the 19th century.

Column (5) of Table 7 shows that, had 1850-80 immigrants and natives had the 1900-30 occupational distributions, the immigrant-native gap would have *widened* by 4.4 percentiles rather

than narrowing by 6.9 percentiles between 1850 and 1880. Similarly, if the 1900-30 sample is weighted by the 1850 distributions of occupations (column 6), the immigrant-native gap would have narrowed by 8.2 percentiles after 1900 instead of widening by 1.8 percentiles. In this sense, the change in young men's occupational distributions between 1850 and 1900 strongly influenced whether convergence occurred. The changing initial occupational distributions are thus proximately responsible for the different immigrant assimilation patterns between the two periods.[17]

As with the results of section 5, there is concern that these results might be affected by the continued assimilation of second-generation immigrants, by different time spent in the US, by differences in selective mortality, or by changes in economic status within occupational groups. As with the results of section 5, we verify robustness to these concerns in Appendix B.

7. Discussion

Taken together, the results of sections 5 and 6 indicate that prevalent views of European immigrants' labor market assimilation during the Age of Mass Migration are incomplete. In section 5, we showed that immigrants' assimilation patterns changed dramatically, with immigrants in the 19th century experiencing substantial convergence in occupational status relative to natives, unlike immigrants in the early 20th century. But while the "catching up" of immigrants in the 19th century was real and contrasts with the conclusions for the Age of Mass Migration based solely on 20th-century data, it was clearly not the product of immigrants' superior performance relative to comparably placed natives, nor was it the product of "old source" immigrants faring better than "new source" immigrants, as we show below. Rather, as shown in section 6, the 19th-century immigrants were disproportionately concentrated in jobs from which many workers were likely to upgrade during their lifetime, whether native or immigrant, new source or old. That is, immigrants' initial concentration in low-status high-upgrading occupations led to greater average gains.[18] The shift of young natives into these higher-upgrading occupations by 1900, combined with natives' greater upgrades conditional on initial status, led to a *divergence* in occupational status for the later cohort.

The change in young natives' occupational distribution from the 19th to the 20th centuries is

---

[17] The results of columns (5) and (6) also imply that, if the initial occupation-specific upgrading were interchanged between periods, this would be insufficient to reverse the results of assimilation in the 19th century and slight divergence in the 20th century.

[18] Although it might be tempting to suppose that being concentrated toward the bottom of the occupational distribution left 19th-century immigrants with upward mobility relative to white natives as the only possibility, the findings are not the result of processes that are mechanical or inevitable. In particular, the experience of black Americans after the Civil War shows that upward mobility from the bottom of the occupational distribution could be very slow (Collins and Wanamaker 2020).

fundamentally linked to the structural transformation of the US economy in the 19th century. Three properties of modern economic growth underpinned this change, driving the sharp decline in natives' agricultural employment—the low income elasticity of demand for farm products, relatively rapid productivity growth in agriculture, and the declining cost of acquiring education with the extension of public schooling (inter alia, Clark 1940; Kuznets 1966; Caselli and Coleman 2001). Even if it was difficult and unlikely for a given farmer to switch to non-farm employment in mid-career, these secular forces ensured that each new generation of workers faced a different opportunity set, and many responded by investing in human capital and migration to pursue jobs outside of farming.

Identifying the precise mechanisms of natives' upgrading is beyond this paper's scope, but the patterns we uncover are consistent with natives responding to both the secular forces mentioned above (by moving out of farming) and to the large influx of less-skilled immigrants (by upgrading within the non-farm sector). In particular, Foged and Peri (2016) and Tabellini (2020) emphasize that immigrants may be complementary factors of production for natives, leading natives to upgrade in terms of skill, occupations, and wages. This mechanism would contribute to our finding of natives' advantage in upgrading relative to similarly placed immigrants, especially for natives who started in the unskilled labor category. Our interpretation of US economic history is that such complementarities would have been especially strong in the latter part of our study, when more natives were starting out in less-skilled jobs in urban areas, which were simultaneously attracting large numbers of European immigrants with limited English skills.

8. The (small) role of changing source countries

There are, of course, other possible interpretations of the finding of immigrants' labor market assimilation in the mid-19th century versus their lack of assimilation in the early 20th century. For instance, influential commentators and policymakers in the early 20th century lamented the arrival of "new" immigrants and questioned their ability to adapt to life in the US. This view framed the "immigration problem" that the literacy test and then the quota system were intended to solve, but at the time (and until now) data did not exist to study the issue carefully.

With linked census data, a simple way to test whether "old-source" and "new-source" immigrants fared differently in terms of occupational upgrading in the early 20th century is to include an indicator variable for "old source," which we define as immigrants from Britain, Germany, and Ireland, in the upgrading regression, as shown in column (7) of Table 7.[19] The

---

[19] Table 2 shows that "old-source" immigrants made up 42 percent of all immigrants in our 1900-30 sample.

coefficient is small, negative, and statistically insignificant. This implies that on average there was no meaningful difference between the groups in average upgrading and, therefore, that changes in immigrants' source countries cannot account for the lack of upgrading relative to natives in the 1900-30 period. A subtler question is whether "old" immigrants fared better than "new" immigrants conditional on initial occupational rank. In column (8) of Table 7, there is some evidence of better performance for the "old" immigrants, but it is a small difference (approximately 1 percentile), and the sum of the coefficients on *immigrant* and *old source* still implies that on average "old" source immigrants fell behind natives during the 1900-30 period.

We provide more detailed information on the relative outcomes for immigrants from different source countries in Figures 6 (occupational rank) and 7 (unskilled occupations), which present estimates of equation (1) with separate β_t coefficients for selected countries. The Irish are of particular interest given the magnitude of their migration and their relatively low initial occupational status (see also Collins and Zimran 2019). Importantly, whereas the Irish in 1850-80 showed strong evidence of "catching up" to natives, the Irish in 1900-30 did not. The patterns for German and British immigrants, though less stark, are consistent: their occupational status converged relative to that of natives in 1850-80, but they fell behind slightly in 1900-30. Again, this undermines the view that the main difference between the early and later period was the source of immigrants.

Given the emphasis in sections 6 and 7 on the quantitative importance of upgrading by men who started in the unskilled category, it may seem surprising that the Irish and Italians, who were relatively concentrated in unskilled work in 1900, failed to narrow the gap in average occupational status by 1930. Instead, the gap widened by 2.2 percentiles for the Irish and by 5.5 percentiles for the Italians (Appendix Table C.1). In part, this is because natives upgraded more than the Irish and Italians from all initial job categories after 1900 (Appendix Figure C.1), but it is also because the Irish and Italian concentration in unskilled work relative to natives was small by 1900 (around 6 percentage points more than natives). The importance of the changes in occupational distributions between 1850 and 1900 can be seen in a counterfactual that is specific to Irish and Italian immigrants (Appendix Table C.1): if the 1900-30 Irish and Italian immigrants are re-weighted to match the 1850 occupational distribution of all immigrants, while the 1900-30 natives are re-weighted to match the 1850 distribution of natives, the Irish would have closed the gap relative to natives by 5.6 percentiles and the Italians by 2.2 percentiles.

In Appendix G, we show that none of changing patterns of return migration over the Age of Mass Migration, changing patterns of selection, or changes in the correlates of upward mobility can

account for the changes in labor market assimilation observed, leaving the structural transformation of the economy as the best and most parsimonious interpretation of the results.

9. Conclusions

The availability of complete-count census datasets and the ability to link records across those datasets have enabled scholars to revisit fundamental questions about the first era of mass immigration to the US. One of those questions pertains to immigrants' labor market outcomes, specifically whether immigrants improved their economic status relative to natives as they gained experience in the US labor market. Although a clearer picture of migrants' assimilation in the 20th century has recently emerged as a result of these data advances, the experience of earlier migrant cohorts is comparatively uncharted territory. In this paper, we provide a more complete view of the Age of Mass Migration by encompassing both the era's early and later decades, providing consistent comparisons of immigrants who arrived at different times and from different sources, and delving into the factors driving changing patterns of assimilation over time. We do this by creating new datasets of linked census records for men between 1850 and 1880 and between 1900 and 1930.

From a contemporary perspective, at a time when immigration policy is again at the forefront of political discourse, this paper's investigation is valuable because many Americans have strongly held views about the experiences of European immigrants in the Age of Mass Migration and what those experiences imply about appropriate policy today (Smith et al. 2018; Martin 2019). Anecdotes, family lore, and wishful thinking have tended to predominate over systematic evidence in forming and sustaining those views of American history. We provide here new data and analyses that illuminate the Age of Mass Migration and confront the idealized notion that European immigrants simply "worked their way up" the economic ladder in America. The more nuanced reality is that the likelihood of immigrants catching up to natives in labor markets changed over time and was strongly influenced by the context of US economic development.

Our analysis arrives at two main conclusions, each with important implications for understanding the long history of immigration in the US. The first is that European immigrants substantially upgraded their occupational status relative to natives in the 19th century, but not in the early 20th century. The second is that the difference in immigrants' upgrading between the early and later decades of the Age of Mass Migration was not rooted in the changing composition of immigrant source countries, which was widely remarked upon and lamented by social scientists and policymakers at the time. Rather, we show that the answer lies in the collapsing differences in occupational distributions between immigrants and natives over the Age of Mass Migration, which in

turn reflects the structural transformation of the US economy. In 1900, many more young native men held unskilled occupations than in 1850. Upgrading from unskilled occupations was more common than from other occupations for both immigrants and natives; moreover, natives had a consistent advantage in upgrading conditional on initial occupation. Thus, as young immigrants' and natives' initial occupational distributions became more similar between 1850 and 1900, immigrants' likelihood of advancing relative to natives eroded. In this sense, the mid-19th-century immigrants' "catching up" to natives was due primarily to the early timing of their arrival, not to productive advantages associated with their home countries nor to their exceptional personal characteristics.

Figure 4: Occupational distributions
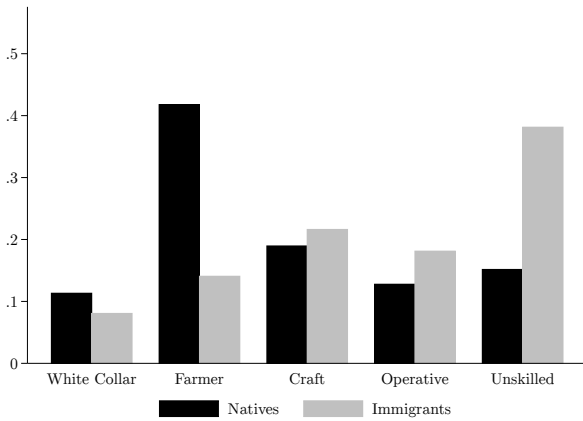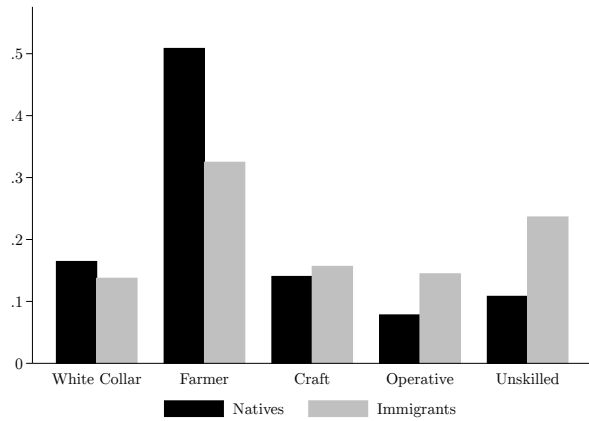
Figure 4(a): 1850
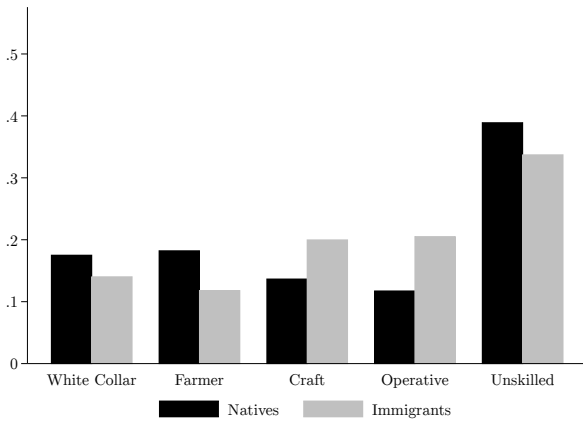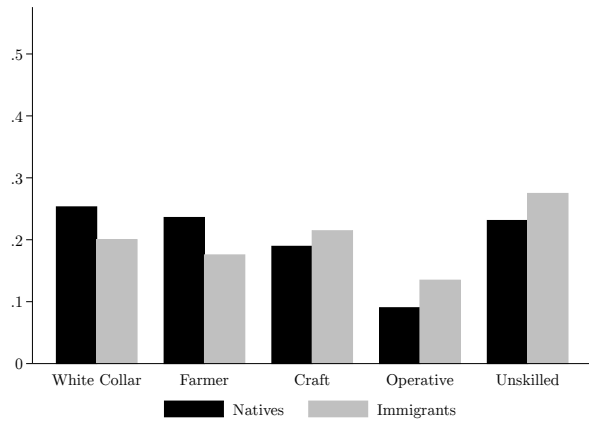
Figure 4(b): 1880

Figure 4(c): 1900

Figure 4(d): 1930

Notes: These graphs show the occupational distribution of natives and immigrants in each year using inverse-probability weights to correct for selection into linkage, but making no other corrections. Occupational categories are defined as in text. Sample limited to individuals with occupations in both years.

Figure 5: Change in occupational rank by initial occupation
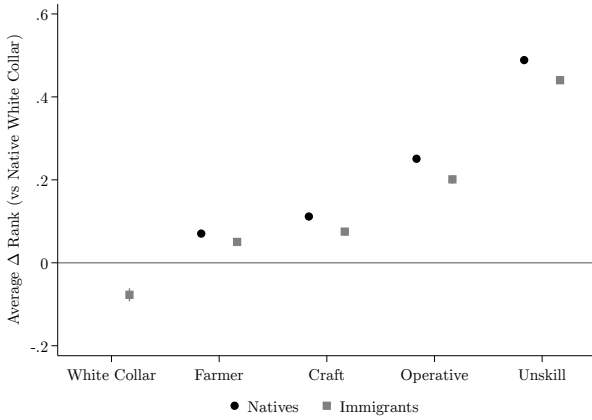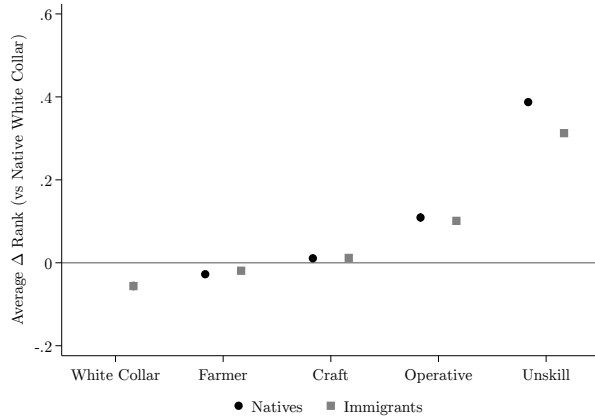
Figure 5(a): 1850-1880

Figure 5(b): 1900-1930

Notes: Each figure presents coefficients from regressing the change in the average occupational rank on nativity-initial occupational category indicators, with native-white collar as the excluded group and controlling for a quartic in age. Robust 95 percent confidence intervals reported (but are so small that they do not exceed width of the point estimate markers). Observations weighted to correct for selection into linkage. Sample limited to individuals with an occupation on both years.

Figure 6: Average occupational rank by country of origin

Figure 6(a): 1850-1880

Figure 6(b): 1900-1930

Notes: Each bar represents a coefficient from estimating equation (1) with separate β_t coefficients for each ethnicity and with the average occupational rank measure as the dependent variable. The excluded group is natives. Robust 95 percent confidence intervals reported. Observations weighted to correct for selection into linkage. Sample limited to individuals with occupations in both years

Figure 7: Unskilled occupation by country of origin
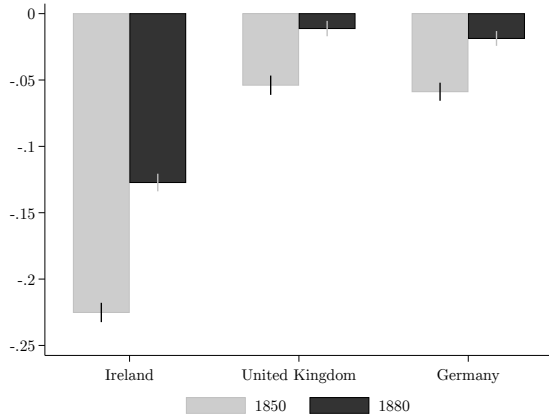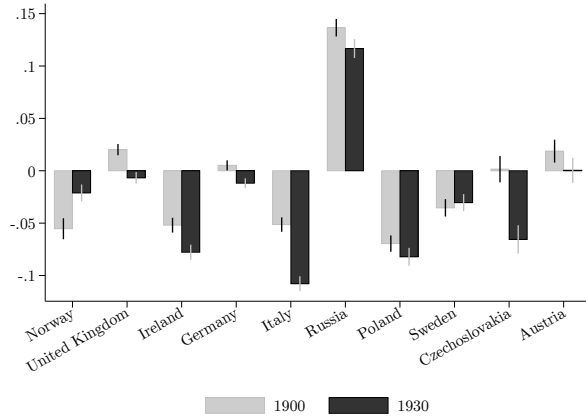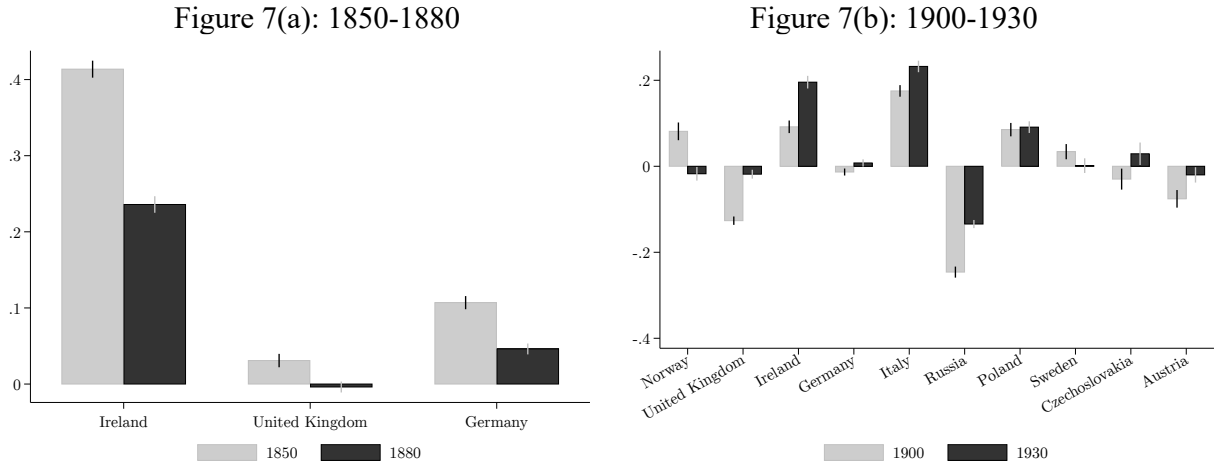
Figure 7(a): 1850-1880

Figure 7(b): 1900-1930

Notes: Each bar represents a coefficient from estimating equation (1) with separate β_t coefficients for each ethnicity and with an indicator for holding an unskilled occupation as the dependent variable. The excluded group is natives. Robust 95 percent confidence intervals reported. Observations weighted to correct for selection into linkage. Sample limited to individuals with occupations in both years.

Table 3: Occupational category differences, immigrants vs. natives

	1850–1880			1900–1930		
	(1)	(2)	(3)	(4)	(5)	(6)
Variable	1850	1880	Diff.	1900	1930	Diff.
Unskilled	0.238^a	0.128^a	-0.111^a	0.005	0.046^a	0.041^a
	(0.004)	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)
Farmer	-0.278^a	-0.187^a	0.091^a	-0.107^a	-0.084^a	0.024^a
	(0.003)	(0.004)	(0.004)	(0.002)	(0.002)	(0.002)
Craft	0.023^a	0.018^a	-0.006^c	0.053^a	0.031^a	-0.022^a
	(0.003)	(0.003)	(0.003)	(0.002)	(0.002)	(0.003)
Operative	0.053^a	0.068^a	0.015^a	0.092^a	0.054^a	-0.038^a
	(0.003)	(0.002)	(0.003)	(0.002)	(0.002)	(0.003)
White Collar	-0.037^a	-0.026^a	0.011^a	-0.043^a	-0.048^a	-0.005^c
	(0.005)	(0.005)	(0.003)	(0.002)	(0.002)	(0.003)

Significance levels: [a] p<0.01, [b] p<0.05, [c] p<0.1
Notes: Robust standard errors in parentheses. Sample limited to individuals with occupations in both years. All specifications include a quartic in age and are weighted by inverse linkage probability, as described in text. Columns (1), (2), (4), and (5) present estimates of β_t from equation (1) for each year with the listed variable on the left-hand side. Columns (3) and (6) present estimates of $\beta_t - \beta_{t-30}$.

Table 4: Dissimilarity indices between natives and immigrants

	(1) 1850	(2) 1880	(3) 1900	(4) 1930
Dissimilarity	0.3148 (0.0037)	0.2133 (0.0039)	0.1500 (0.0027)	0.1315 (0.0027)
Observations	237,203	237,203	668,061	668,061

Notes: Dissimilarity indices between natives' and immigrants' occupational distributions in each year, controlling for a quartic in age as described in equation (2). Robust delta method standard errors in parentheses. Observations weighted to correct for selection into linkage. Sample limited to individuals with occupations in both years.

Table 5: Occupational income differences, immigrants vs. natives

	1850–1880			1900–1930		
Variable	(1) 1850	(2) 1880	(3) Diff.	(4) 1900	(5) 1930	(6) Diff.
Average Occ. Rank	-0.138^a (0.003)	-0.069^a (0.002)	0.069^a (0.002)	-0.015^a (0.002)	-0.035^a (0.002)	-0.020^a (0.002)
log(Occ. Wealth)	-0.661^a (0.010)	-0.404^a (0.013)	0.257^a (0.009)	-0.151^a (0.006)	-0.164^a (0.006)	-0.013 (0.008)
log(PH Occ. Score)	-0.107^a (0.004)	-0.049^a (0.002)	0.059^a (0.004)	0.027^a (0.002)	-0.021^a (0.002)	-0.048^a (0.003)
Occ. Wealth Rank	-0.183^a (0.003)	-0.105^a (0.003)	0.078^a (0.002)	-0.032^a (0.002)	-0.046^a (0.002)	-0.014^a (0.002)
PH Score Rank	-0.094^a (0.003)	-0.034^a (0.003)	0.060^a (0.003)	0.002 (0.002)	-0.024^a (0.002)	-0.025^a (0.002)

Significance levels: [a] $p<0.01$, [b] $p<0.05$, [c] $p<0.1$
Notes: Robust standard errors in parentheses. Sample limited to individuals with occupations in both years. All specifications include a quartic in age and are weighted by inverse linkage probability, as described in text. Columns (1), (2), (4), and (5) present estimates of β_t from equation (1) for each year with the listed variable on the left-hand side. Columns (3) and (6) present estimates of $\beta_t - \beta_{t-30}$.

Table 7: Conditional changes in rank

Variables	(1) 1850–1880	(2) 1900–1930	(3) 1850–1880	(4) 1900–1930	(5) 1850–1880	(6) 1900–1930	(7) 1900–1930	(8) 1900–1930
Immigrant	0.069^a (0.002)	-0.018^a (0.002)	-0.040^a (0.002)	-0.029^a (0.002)	-0.044^a (0.004)	0.082^a (0.002)	-0.017^a (0.003)	-0.034^a (0.002)
Initial Avg. Occ. Rank			-0.781^a (0.007)	-0.712^a (0.004)				-0.713^a (0.004)
Old Source							-0.003 (0.003)	0.012^a (0.002)
Observations	237,016	667,717	237,016	667,717	237,011	659,693	667,717	667,717
R-squared	0.028	0.100	0.470	0.427	0.068	0.077	0.100	0.427
Weights					1900	1850		

Significance levels: [a] $p<0.01$, [b] $p<0.05$, [c] $p<0.1$
Notes: Dependent variable is change in average occupational rank. Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. Excluded group in all specifications are natives. Weights indicate that the data are reweighed to match the occupational distribution of the year listed in the last row of the table, for immigrants and natives separately. Sample limited to individuals with occupations in both years.