THE ECONOMIC ASSIMILATION OF IRISH FAMINE MIGRANTS TO THE UNITED
STATES

William J. Collins
Ariell Zimran

The Economic Assimilation of Irish Famine Migrants to the United States
William J. Collins and Ariell Zimran
NBER Working Paper No. 25287
November 2018
JEL No. F22,J61,J62,N31,O15

ABSTRACT

The repeated failure of Ireland's potato crop in the late 1840s led to a major famine and a surge in migration to the US. We build a dataset of Irish immigrants and their sons by linking males from 1850 to 1880 US census records. For comparison, we also link German and British immigrants, their sons, and males from US native-headed households. We document a decline in the observable human capital of famine-era Irish migrants compared to pre-famine Irish migrants and to other groups in the 1850 census, as well as worse labor market outcomes. The disparity in labor market outcomes persists into the next generation when immigrants' and natives' sons are compared in 1880. Nonetheless, we find strong evidence of intergenerational convergence in that famine-era Irish sons experienced a much smaller gap in occupational status than their fathers. The disparities are even smaller when the Irish children are compared to those from observationally similar native white households. A descriptive analysis of mobility for the famine-era Irish sons indicates that more Catholic surnames and birth in Ireland were associated with less upward mobility. Our results contribute to literatures on immigrant assimilation, refugee migration, and the Age of Mass Migration.

William J. Collins
Department of Economics
Vanderbilt University
VU Station B #351819
2301 Vanderbilt Place
Nashville, TN 37235-1819
and NBER
william.collins@vanderbilt.edu

Ariell Zimran
Department of Economics
Vanderbilt University
2301 Vanderbilt Place
Nashville, TN 37235
and NBER
ariell.zimran@vanderbilt.edu

1. Introduction

Driven from their home countries by violence, natural disasters, or other life-threatening shocks, large numbers of refugees have sought residence in the United States and Europe in recent years (United Nations High Commissioner for Refugees 2017). In the United States, this has fueled a debate in which concerns about immigrants' economic and cultural assimilation have been cited as justification for more restrictive immigration and refugee policies (e.g., Kelly 2018), such as recent proposals that potential entrants be screened to favor those with a greater "likelihood of successful assimilation and contribution to the United States" (US Department of State et al. 2017, p. 8). Concerns with immigrant assimilation are also commonly cited in European immigration policy debates (OECD/EU 2015; Federal Chancellor 2018), which have heightened with the influx of asylum-seeking migrants since 2014.

This focus on immigrants' assimilation and contribution echoes concerns that have surfaced repeatedly in US history in response to the arrival of poor or culturally "different" immigrants. One of the earliest appearances of these concerns coincided with the first large wave of desperate and poor immigrants to arrive in the US—refugees fleeing Ireland's Great Famine.[1] From the onset of the Great Famine in 1846 to 1850, over 500,000 Irish immigrants arrived in the United States (Ferenczi and Wilcox 1929; Barde, Carter, and Sutch 2006).[2] Although Irish immigration had been growing since the 1830s as part of a general rise in transatlantic migration (Mokyr and Ó Gráda 1982; Ó Gráda 1983; Cohn 2009), the arrival of famine refugees marked a sharp increase in the rate of immigration and an apparent change in the character of Irish immigrants (Handlin 1991 [1941] p. 51; Miller 1985 p. 295; Anbinder 1992, p. 7). Many Americans viewed the Irish migrants' poverty, tendency to live near one another, and predominantly Catholic religion as barriers to their assimilation and, therefore, as justification for restrictions on immigration and immigrants' rights (Anbinder 1992; Hirota 2017).

In this paper, we study the economic status and labor market assimilation of famine-era Irish immigrants, with a particular focus on the adult labor market outcomes of the immigrants' children.

---

[1] We refer to migrants driven by famine conditions as "refugees." This is consistent with the definition in the Oxford English Dictionary and usage by scholars such as Ó Gráda and O'Rourke (1997). It is important to note, however, that in the postwar period, the UN High Commissioner for Refugees has limited the term to individuals driven from their homes by violence or persecution. The combined group of migrants, including those fleeing ecological and natural disasters, is sometimes referred to as "forced migrants." See Hatton (2012) for a brief review of refugee and asylum migration to OECD countries since the 1940s.

[2] See Ó Gráda (1999, ch. 3) for discussion of the dating of the Famine and the challenges of estimating excess deaths and emigrants. See Cohn (2009, pp. 26-27) for difficulties in determining the number of Irish immigrants from official US immigration statistics before 1850.

Famine-era Irish immigrants likely comprised the largest wave of refugees from a single source that the US has ever absorbed relative to the size of its population, which makes them a particularly interesting group to study.[3] Unfortunately, data constraints have made it difficult to study the long-term outcomes of these migrants and their children. To overcome these constraints, we constructed a new micro-level dataset by linking males born in Ireland, Germany, Britain, and the US from the 1850 to the 1880 complete-count US census records.[4] The linked dataset is large, national in scope, and conservatively constructed to reduce false matches. It includes household heads and their sons; therefore, it allows us to compare labor market outcomes for immigrant and native groups over two generations, and to consider the second generation's labor market outcomes in light of variation in their early life circumstances. Moreover, the new dataset's panel structure helps us to avoid biases from cohort quality changes and selective return migration that confound inferences from cross-sectional census data (e.g., Lubotsky 2007; Abramitzky, Boustan, and Eriksson 2014). That is, because we follow a fixed set of men over time, there are no changes in sample composition.

The children's economic outcomes are especially interesting in the context of concerns about the long-run assimilation of new immigrant groups. Immigrants' children's outcomes have garnered considerable attention in the literature on the economics of migration (e.g., Borjas 1992, 1993; Card, DiNardo, and Estes 2000; Card 2005; Caponi 2011; Abramitzky, Boustan, and Eriksson 2014; Alexander and Ward 2018). But to our knowledge, this is the first paper to study the children of the large cohorts of immigrants who arrived in the mid-nineteenth century US. Their arrival marked the beginning of the "Age of Mass Migration," which saw nearly 30 million Europeans migrate to the US by 1913. This is also one of a small number of papers that address long-run patterns of economic assimilation by refugee immigrants or their children, whose experiences and outcomes may differ from those of other migrant groups (Edin, Fredriksson, and Åslund 2003; Cortes 2004; Beaman 2012; Evans and Fitzgerald 2017).

Our focus on the immigrants' children is also a practical consequence of the historical census data's limitations. The censuses of this era did not inquire directly about each immigrant's year of arrival. This poses a major challenge to discerning between those who arrived before or after the

---

[3] Arguably, it was the largest influx in absolute terms as well, but changes in US immigration policy make it difficult to form precise comparisons between modern and historical immigrant flows. For perspective, refugees and asylees granted permanent resident status under the Displaced Persons Act (1948) averaged about 118,000 per year from 1950-52, the program's peak years (Barde, Carter, and Sutch 2006, p. 1-632); the Mariel Boatlift entailed approximately 125,000 Cuban migrants in 1980 (Card 1990, p. 245).
[4] Linking females from childhood to adulthood in census records is extremely difficult due to name changes at the time of marriage.

Irish famine's onset and, in general, to any study of immigrant assimilation in this early period.[5] Nonetheless, in 1850, we can determine the arrival cohort for many household heads by examining their children's birth year and birth place information.[6] This approach definitively categorizes the household head's arrival cohort ("pre-famine" or "famine-era") for about two-thirds of the sons of Irish immigrants, and it proves to be very accurate in cases where the classification can be verified with outside information (i.e., searching passenger lists).[7] The relatively high classification rate makes studying the children's outcomes attractive—we know whether their household head arrived before or during the famine, observe their childhood household characteristics in 1850 in detail, and then see their labor market outcomes as prime-aged adult workers in 1880.[8] In addition, the "classified" set of children is fairly representative of all immigrants' children circa 1850 in terms of observable characteristics.

We use the new dataset to address three main sets of questions. First, in 1850, how different were the households headed by famine-era Irish migrants from those headed by earlier Irish migrants, by concurrent migrants from Germany and Britain, and by US natives? In particular, is there evidence of a change in Irish household heads' human capital (reflecting changing selection) and labor market outcomes (reflecting both selection and labor market conditions for newly arrived migrants) between the pre-famine and famine-era migrants? Given this paper's motivation, we care about changing migrant selection primarily because it directly influences the average household characteristics and resources of immigrants' children, though documenting the patterns of migrant selection at the time of the famine is also of independent interest. Consistent with the predictions of a simple Roy (1951)-Borjas (1987) model, we find clear evidence of deterioration in the human capital of Irish immigrant household heads with the onset of the famine, as measured by literacy and age heaping in 1850. This is the clearest evidence to date on human capital differences between pre-

---

[5] One reason much of the recent literature on the Age of Mass Migration focuses on the early twentieth century is that the census microdata report year of immigration starting in 1900. Before 1850, the census did not inquire about place of birth; so, there is limited scope for backward linkage. Extending links to the passenger lists in this early period is difficult due to the sparseness of records and the feasibility of making multiple accurate links across various datasets (see footnote 31).

[6] For example, a family with an Irish-born child aged 2 in 1850 must have immigrated to the US some time in or after 1848, which would make them famine migrants. Conversely, a family with an American-born child aged 7 in 1850 must have immigrated to the US some time before or during 1843, making them pre-famine migrants. We make the same distinctions for the British and German migrants (despite there not actually having been a famine in these countries) in order to have comparison groups of the same arrival cohorts.

[7] In contrast, this approach would enable us to classify the arrival cohort of fewer than half of adults in the 1850 census.

[8] At the time of writing, the 1860 and 1870 complete count data have not been cleaned and coded by IPUMS and, therefore, are not included in this paper's analysis.

famine and famine-era arrivals. We also document a decline in occupational status between the pre-famine and famine-era Irish, both absolutely and relative to differences over arrival cohorts for other immigrant groups. Thus, by changing both the composition and volume of Irish migration, the Great Famine resulted in a large cohort of relatively poor immigrant children in the US circa 1850.

Second, in 1880, how did the adult labor market outcomes of Irish famine migrants' sons compare to those of other immigrants' and natives' sons? Studying these outcomes provides perspective on the potential for long-term assimilation in a setting where a variety of forces— including inauspicious early life conditions, an environment rife with anti-Irish sentiment, and a continuing influx of new immigrants—may have hindered the advancement of the famine Irish children. We find that, on average, the sons of the famine-era Irish immigrants fared poorly in the labor market as adults in comparison with other groups in 1880. Nonetheless, in comparison to their fathers' starting point, they greatly narrowed the gap in occupational status relative to natives. The gap is even smaller when considering their adverse childhood environment (i.e., controlling for 1850 household characteristics). In this sense, fairly strong economic assimilation occurred over generations. The labor market outcomes of Irish immigrants circa 1850 were thus a poor guide to inferring the group's ability to assimilate over a longer period.

Finally, we examine heterogeneity in the upward mobility of the children of famine-era Irish immigrants according to observable characteristics, including measures of their father's human capital, residence in an Irish enclave, social distance (e.g., having a surname that we determine to be predominantly Catholic), and geographic mobility. The data and setting do not allow clear identification of causal relationships; they do, however, provide novel evidence on factors that may have facilitated or impeded immigrants' intergenerational gains. In particular, we find that having a more Catholic surname and being born in Ireland were associated with less upward mobility, conditional on other observables. This pattern is consistent with the presence of discrimination against Irish Catholics, lower levels of human capital for Irish Catholics, and long-term negative consequences from exposure to the famine, all of which merit closer examination in future research.

This paper contributes to the literature on the economics of immigration in several dimensions. First, its emphasis on economic assimilation in the long run speaks to a core theme of the international migration literature, both historical and contemporary.[9] By studying a large wave of

_____

[9] On immigrant assimilation, both historical and contemporary, see inter alia Ferenczi and Wilcox (1929), Chiswick (1978), Borjas (1992, 1993), Hatton (1997), Hatton and Williamson (1998), Card, DiNardo, and Estes (2000), Minns (2000), Cortes (2004), Card (2005), Abramitzky, Boustan, and Eriksson (2014, 2016), and Alexander and Ward (2018).

migrants and their offspring long before the implementation of restrictive immigration policies, the paper yields insight into the process and potential for assimilation in a setting where economic forces predominated.[10] In addition, the experience of Irish famine-era migrants may yield broader insights into the economics of large-scale migration due to natural or man-made disasters.[11] In this case, we find that despite a notable deterioration in migrant selection and a significant political backlash, the migrants' children converged strongly, albeit incompletely, on native outcomes by 1880. In this sense, they showed clear evidence of "assimilation and contribution" to the American economy. Finally, the paper advances our knowledge of the Age of Mass Migration and is thus complementary to research by Hatton and Williamson (1998), Ferrie (1999), Cohn (2009), and Abramitzky, Boustan, and Eriksson (2012, 2014), among others. By studying immigrants' experiences in the early portion of this period, in contrast to a recent scholarship that has focused on the post-1900 period, the paper brings the original cohorts of the Age of Mass Migration into much sharper focus.

2. Background on Ireland's Great Famine and Migration to the US

On the eve of the Great Famine, two-thirds of Irish families were employed primarily in agriculture (Commissioners 1843, p. xviii). Most owned little or no land (Ó Gráda 1999, p. 25) and had few financial resources. Widespread poverty and heavy reliance on the potato left the Irish vulnerable to large and repeated failures of the potato crop, as occurred in 1845 and 1846 due to the spread of a microorganism that causes blight (*Phytophthora infestans*). By late 1846, Ireland was in the grips of a historic famine that was further exacerbated by crop failures in 1848 and, less severely, in 1849 and 1850.[12]

Even before the famine, the migration flow from Ireland to the US was substantial (Mokyr and Ó Gráda 1982). Historians suggest that pre-famine migrants were likely drawn from above the Irish median in terms of skill and socioeconomic status, but perhaps less so over time as costs of migration declined (Miller 1985, pp. 193-201). Consistent with the impression of positive selection, whereas approximately 40 percent of men (ages 16-45) in Ireland could neither read nor write in 1841 (Commissioners 1843, p. xxxvi), only 15 percent of Irish-born men (ages 25-54) were classified

---

[10] See, however, Hirota (2017) on efforts by some states to remove destitute immigrants.

[11] On migration due to environmental disasters or war see recent work by Boustan, Kahn, and Rhode (2012), Hornbeck (2012), Moser, Voena, and Waldinger (2014), Mahajan and Yang (2017), Nakamura, Sigurdsson, and Steinsson (2017), Blum and Rei (2018), and Long and Siu (2018).

[12] Mokyr (1983) and Ó Gráda (1999) discuss the chronology of agricultural failure and famine in detail. See Solar (1989) on Irish crop failures in historical and international perspective.

as illiterate in the 1850 US census.[13]

The famine marked a turning point in Irish economic and demographic history (Mokyr 1983; Guinnane 1997; Ó Gráda 1999). In 1841, there were over 8 million residents of Ireland; by 1851, the population had declined by about 1.6 million. It is estimated that approximately one million died as a result of the famine, while another million fled Ireland (Ó Gráda 1999, p. 41), mostly to North America. To be sure, some of those who migrated during the famine would have left in any case. But Ó Gráda points out that "… most of the mass emigration of the late 1840s was part of the famine tragedy. It was push migration with a vengeance, and its tragic character has been rightly at the center of historical writings about it" (1999, p. 105). Miller writes of a "mass, indiscriminate rush to leave" Ireland in the wake of the potato blight (1985, p. 292). Many of the migrants were making a leap of faith in a desperate context characterized by widespread death and social collapse. Like modern day refugees, their departure was precipitated by an imminent threat to survival, often undertaken hastily, and generally irreversible.[14] The parallel between the famine migrants and modern refugees fleeing war and instability has been noted in the Irish press (Delaney 2015). A major difference, of course, is that migrants fleeing to the US during the famine did not face the legal restrictions on mobility that are prevalent today.

The bulk of the Irish famine emigrants, including many of those who originally landed in Canada, settled in the US (McInnis 2000; Cohn 2009). Passenger records from US ports of entry suggest that the annual rate of Irish immigration nearly doubled with the famine's onset.[15] By June 1850, when the US population census was taken, there were nearly one million Irish-born residents.

In addition to generating a surge in the volume of migration, there are indications that the famine may have led to a change in the composition of Irish migration to the US. Some clues come from fragmentary information on the migrants' places of origin. It appears that the main sources shifted somewhat from the North and East (Ulster and Leinster) toward the South and the West (Munster and Connacht) (Ó Gráda 1999, p. 113; Anbinder and McCaffrey 2015; Anbinder, Ó Gráda, and Wegge 2017). This suggests that the characteristics of migrants may have changed because residents of the South and West were more agricultural and less literate than elsewhere

---

[13] Authors' calculations from 1850 complete count census. We caution that comparing across the censuses this way is somewhat tenuous; it is unclear how comparable literacy measures are. The differences are sufficiently large in this case that we think the information is useful. See Connor (2018) for evidence of negative selection in the early twentieth century.

[14] The post-famine transformation of the Irish economy, which included consolidation of small landholdings, ensured that emigration continued (Ó Gráda 1999). The flow did not reverse once famine conditions passed.

[15] Cohn (2009, pp. 26-27) warns that the time-series data on Irish immigration in the early to mid-nineteenth century allocates a fixed share of UK immigrants to Ireland (0.70) rather than providing a true count.

(Commissioners 1843, Geary 1996); they were also more likely to be Catholic and to speak Irish. Passenger lists of ships entering US ports are also potentially informative about changing selection. The lists include self-reported occupation and age, typically recorded at the European port of embarkation. They provide a window on the pre-departure activity of migrants, but the evidence is ambiguous and much caution is necessary (Cohn 2009, ch. 5), which is one factor that motivates our census-based approach.[16] Cohn (1995) finds a slight decrease in the fraction of Irish immigrants who were reported as laborers after 1845, whereas Mokyr and Ó Gráda (1982) find an increase in age heaping in passenger lists after 1845, which is viewed as a sign of innumeracy in historical populations.

In the US, Irish immigrants were over-represented in urban areas of the Northeast. It is worth emphasizing, however, that most Irish immigrants, including those we identify as famine-era arrivals, did not reside in New York City (13.8 percent of all Irish-born) or Boston (3.5 percent) in 1850.[17] Landmarks in the historical literature tend to focus on these cities (Handlin 1991 [1941], Ernst 1949), but a much broader geographic scope is required to describe the experience of the Irish immigrants. Scholars have long understood that Irish immigrants were disproportionately employed in relatively low-skill, low-paying lines of work (Handlin 1991 [1941]; Ernst 1949; Miller 1985). Ferrie (1994, 1995, 1997a, 1999) provides a more dynamic view by drawing a sample of immigrant men from New York City's passenger lists in the 1840s and linking them to census records in 1850 or 1860. He finds a great deal of mobility between pre-migration occupational categories, coded from passenger lists, and those reported in the 1850 or 1860 US census (1999, Table 5-2). Irish men had lower pre-migration occupational status than the British and Germans based on the passenger lists (i.e., more unskilled laborers), had lower rates of upward mobility from the unskilled category, and accumulated less real estate wealth (1999, Table 6-4). Even so, Anbinder, Ó Gráda, and Wegge (2017) find evidence of upward mobility in the account records of Irish immigrants at the Emigrant Industrial Savings Bank in New York.

In addition to their economic struggles, mid-century Irish immigrants were subject to a sharp nativist backlash that was rooted in long-standing anti-Catholic sentiment in the US and heightened

---

[16] The lists are potentially incomplete in coverage, may have deteriorated in quality during the famine, and may have varied in enumeration practices across ports of embarkation. Even within a given list, clerks may or may not have accurately recorded each passenger's primary past occupation.
[17] Authors' calculations with complete count census data (Minnesota Population Center 2017; Ruggles et al. 2015).

by the sheer volume of mass migration (Hirota 2017).[18] A salient manifestation of this backlash was the success of the "Know Nothing" party in the early 1850s, which advanced the ideas that "Protestantism defined American society" and that "Catholicism was not compatible with the basic values Americans cherished most" (Anbinder 1992, p. 104).

Despite much valuable scholarship on Irish immigrants in the US, a number of important gaps remain in our understanding of the famine migrants' experience. First, there is no clear picture of differences between the Irish who arrived before the famine and those who arrived during the famine. To date, the evidence is somewhat fragmentary and inconclusive, and it is unknown whether changes over Irish arrival cohorts differed from changes for other immigrant groups. Moreover, although scholars have explored the economic assimilation of first-generation Irish immigrants (e.g., Ferrie 1999), there is no information available on the assimilation experience of the immigrants' children. This absence stands out relative to studies of modern migrant assimilation (e.g., Borjas 1993; Card 2005), in which children's outcomes are of fundamental interest. Finally, the economic mobility of the famine Irish has not been closely compared to that of natives, let alone in an intergenerational framework that allows us to tie children's outcomes to their early life circumstances at an individual level. The mid-nineteenth century was a period of high intergenerational mobility among US natives (Long and Ferrie 2013). It is thus essential to compare immigrants and natives to determine whether immigrants closed the gap in economic status and, in that sense, assimilated over generations. This paper develops new evidence on all of these fronts.

3. Data sources and construction

To construct the dataset, we first obtained the complete count US census data for 1850 (Ruggles et al. 2015; Minnesota Population Center 2017). Because this census (and all those prior to 1890) did not inquire about immigrants' year of arrival in the US, our first task was to develop an alternative approach to distinguish the Irish who arrived during the famine from those who arrived prior to it. We determined the arrival cohort for each Irish-born head of household based on the ages and places of birth of children in the household in the 1850 census.[19] Household heads with a child born in the US in 1845 or earlier are categorized in the pre-famine group, while those with a child

---

[18] By this time, Irish immigrants were overwhelmingly Catholic. Many German immigrants were also Catholic, but they were better off and less concentrated in cities of the Northeast than the Irish.

[19] We use an indicator for relationship to the head of household provided in the original data by Ruggles et al. (2015). Our understanding is that individuals are assumed to be children of the head of household if they are of appropriate age and if they share a surname with the head of household. We must rely on this imputed indicator because the census did not inquire regarding relationship to the head of household until 1880.

born in Ireland in 1846 or later are in the famine-era group.[20] We performed a similar classification for British- and German-born heads of household in 1850.[21] We include the Germans and British in this study because, after the Irish, they were the largest groups of immigrants before 1850 (Barde, Carter, and Sutch 2006, p. 1-560) and they provide a useful point of comparison. This classification scheme leads to six categories of immigrant household heads: pre-famine Irish, famine-era Irish, and similar cohorts for the British and Germans. We also include households headed by native-born white men for comparison.

We assigned the household heads' arrival cohort classification and ethnicity to their sons aged 18 or younger. For example, sons of an Irish father who arrived in the US between 1846 and 1850 are referred to as members of the "famine-era Irish." Importantly, not all children in Irish, British, or German-headed households were foreign-born. It follows that not all children classified as being famine-era Irish would have experienced the famine in Ireland (though 40 percent of our benchmark sample of famine Irish sons were born in Ireland during the famine). Given that the paper's main goal is to characterize long-term assimilation as revealed in the outcomes of immigrants' children, partitioning the data this way provides simple but new insights.

Table 1 divides each immigrant group using the categorization described above, reporting for heads of household in columns (1) to (3) and for sons in columns (4) to (6). The fraction of individuals categorized definitely into the pre-famine or famine-era cohorts was approximately 65 percent for the Irish sons, but only about 44 percent for all Irish heads of household (including those without children).[22] The Germans and British exhibit similar patterns. As mentioned earlier, the greater ability to categorize children than heads of households or other adults is part of our motivation for focusing on the children's long-term outcomes. (We discuss representativeness below.)

In the next stage, we linked individuals observed in 1850 to the full count 1880 census (Ruggles et al. 2015; Minnesota Population Center 2017). We briefly describe the linking procedure here and provide additional details in Appendix A. We began by extracting all males born in Ireland,

---

[20] It is not possible to place all household heads into one of these two categories, and therefore we create two more categories: those with a child born in the United States in 1846 or later and a child born in Ireland between 1841 and 1845 ("probable famine-era"), and all others ("uncategorized"). We do not use these latter two groups in our main analysis. The outcomes of the "probable" group resemble those of the "famine" group, while those of the "uncategorized" group fall between those of the "famine" and "pre-famine" groups.

[21] The data obtained from Ruggles et al. (2015) combine England, Scotland, and Wales into a single birthplace of Great Britain. We retain this combination.

[22] The main group that we cannot classify by our approach (and who are not included in Table 1) is boarders, who were 27 percent of all Irish-born adult males in the 1850 US census.

Britain, Germany, and the United States from the 1850 census.[23] We retained only individuals who were unique in terms of place of birth, age-implied year of birth (plus or minus four years), and first and last name (with allowances for orthographic differences).[24] Finally, we linked the remaining 1850 men to the 1880 census, according to the same identifying characteristics. We retained only matches in which both the 1850 and the 1880 record had a unique match in the opposite census.

Table 2 presents rates of successful linkage. Column (1) shows the number of observations with which the linkage process began for each 1850 group. Column (2) shows the fraction of individuals who remained after removing those who were not unique on name, birth place, and age. The fraction of these individuals who were successfully located in 1880 is shown in column (3). Among the Irish children, we linked about 15 percent of the unique famine-era Irish and about 20 percent of the unique pre-famine Irish. Among Germans, the rates of linkage are comparable to the Irish, whereas they are between 23 and 26 percent for the British and Americans. These match rates for the British and Americans are comparable to those of Ferrie (1996, Table 4). The match rates for the Irish and Germans are somewhat lower, and there are reasons to expect this. Besides the lower levels of literacy and education of the Irish, which may tend to reduce match rates, the arrivals of large cohorts of Irish and German immigrants after 1850 would have potentially confounded links even of individuals who were unique in 1850.[25]

In sum, our benchmark dataset consists of males who were aged 0 to 18 in 1850, sons of the household head in 1850 (according to the Ruggles et al. 2015 classification), successfully linked from the 1850 to the 1880 US census, born in the US, Ireland, Britain, or Germany, and with a father born in the US, Ireland, Britain, or Germany. We also require that, if the child's head of household in 1850 was an immigrant, it was possible to use the method described above to determine whether the head of household was in the famine-era or the pre-famine cohort.

Potential pitfalls in data construction

Our data construction process is subject to four main potential pitfalls. First, only a fraction of 1850 individuals are linked to the 1880 census, and selection into linkage is likely to be non-

---

[23] We do not constrain the sample for linking to those whose arrival cohort could be determined, as this might result in false matches. Instead, we attempt to link the 1850 census without regard for whether arrival cohort could be determined, and then retain only successful links whose arrival cohorts could also be classified.
[24] That is, we linked the 1850 census to itself and retained individuals who linked only to themselves.
[25] For instance, a Patrick Kelly born in 1845 in Ireland may have been unique in the United States in 1850, but other Patrick Kellys born in 1845 in Ireland may have arrived after 1850, preventing the identification of the correct Patrick Kelly in 1880.

random. In Appendix A, we present linear probability models that relate the probability of being linked to a variety of characteristics in 1850. There are several statistically significant predictors of selection into linkage, such as higher property ownership and literacy; this is typical in studies that link census data (e.g., Abramitzky, Boustan, and Eriksson 2014; Beach et al. 2016). However, most differences are small, as indicated in Figure 1, which compares means of variables for linked children and all children of Irish migrants in 1850. To address selection into linkage, we weight all of the analyses by the inverse of the estimated conditional probability of successful linkage.[26]

The second issue is that automated linkage methods may yield a high rate of false positives in matching (Bailey et al. 2017). Our linkage approach mitigates this concern. We eliminate individuals who are not unique in terms of the linkage characteristics from the sample that we attempt to link; this reduces the probability of a making false match. Moreover, we do not make links on the basis of a name standardization algorithm, a practice that Bailey et al. (2017) highlight as a source of false positives.[27] To verify that false positives in linkage are not responsible for our results, we have also examined a sample of exact unique matches only.[28] This further reduces the probability of false matches and does not qualitatively affect our main results.

A third concern is that the roughly two-thirds of children whose father's arrival cohort is definitively classified might not be representative of the population of children in 1850. To determine whether selection into successful classification is likely to bias our results, we compare the means of 1850 observables of the classified Irish children to those of the whole sample of Irish children in Figure 2.[29] Though the classification does appear to favor individuals with better socioeconomic status, such as those whose household heads owned more property in 1850, the differences between the classified group and all children are small.[30]

---

[26] This probability is computed from separate probit regressions of a linkage indicator on a variety of covariates for each of the seven ethnicity-cohort groups.

[27] We use the soundex algorithm to reduce the set of candidate matches, but all of the linking is ultimately based on orthographic distance, which is a score derived from the number and type of spelling changes that would be required to move from one name spelling to another.

[28] We define an exact match as one in which there is no more than a one-year difference in age-implied birthyear and the first and last names of both records are identical after removing double letters and common first name abbreviations (e.g., changing "Wm" to "William").

[29] Figure 2 covers the linked sample and is weighted to correct for selection into linkage.

[30] Any concern that non-random selection into successful classification may be responsible for our results can be addressed by noting that our results are qualitatively unchanged when we re-weight our sample to correct for selection into both linkage and classification on the basis of observables. These results are available on request. However, we prefer not to re-weight to correct for selection into classification in our benchmark results because it is not clear whether differences in observables between the classified and unclassified groups

Finally, there is concern that the categorization of arrival cohort is confounded by incorrect reporting of children's ages. We view the danger of age heaping as less severe for children than it would be for adults due to the greater developmental differences by age among children. To provide direct evidence, we randomly selected 50 classified households from our linked data, stratifying to include 25 of each categorization group (famine and pre-famine) and searched for their passenger list records using Ancestry.com. Although we were able to locate only 10 of these households definitively,[31] all 10 were correctly categorized by the age- and birthplace-based algorithm.

Additional variables of interest

The US census did not collect information on income before 1940; therefore, we rely heavily on occupation to indicate men's relative economic status. The simplest approach is to define and study a categorical variable for "unskilled labor," which combines "laborers" and other relatively low-skilled occupations, such as "porter" and "miner" (following Ferrie 1995, 1999). In addition, we create a variable that exploits the 1870 census's information on real and personal wealth. For each occupation-by-region (North/South) cell, we calculate a "wealth score" that equals the average total wealth of men (age 30-65).[32] This is akin to the occupation score variable based on 1950 income data that is often used in studies of twentieth-century US labor markets.[33] We believe the 1870-based occupation score is better suited for studying the mid-nineteenth century because it is, of course, closer in time to our period of study, and it is designed to allow for differences between the North and South. We also use this wealth score to define an occupational rank variable, which represents

---

reflect selection into classification (which we would like to reverse) or actual differences in the probability that a classified and unclassified individual arrived during the famine.

[31] This exercise also exhibits the difficulties of linkage from the passenger lists to the census. One clear concern is the rate of linkage. Combined with the number of observations lost in the link from 1850 to 1880, requiring an additional link to the passenger lists would result in a very small sample. Even achieving this low rate of linkage to the passenger lists required using information on individuals' household composition, and in some cases, individuals could be located only when searching for their household members. The use of household information in linkage is generally to be avoided due to the bias it introduces. Using our household composition-based approach to determine the arrival year enables us to avoid these limitations of linkage from the passenger lists and to maintain a larger sample.

[32] We have also experimented with wealth scores at the source country-occupation-region level. Unfortunately, this cuts the data thin in many instances, and it is not very well suited to our study of immigrants' children (many of whom were US-born). The 1850 census did not ascertain personal wealth, and the 1880 census did not inquire about wealth at all.

[33] The correlation between our measure in 1880 and the 1950-based occupational income score used by studies such as Abramitzky, Boustan, and Eriksson (2014) is only 0.28 in our linked data. However, this is largely due to the changing occupational status of farmers in the intervening period. When farmers are excluded, the correlation is 0.75.

the percentile of the wealth score in the sample.[34]

For analysis of heterogeneity in outcomes among the Irish (Section 5), we constructed three additional variables. The Irish immigrants' predominantly Catholic religion was salient and controversial at the time, but of course there was variation in religious affiliation among the Irish. The US population census has never inquired about individuals' religious affiliation, and no such information is available from passenger lists. To gain some insight on the association between religion and economic outcomes for the Irish, we construct a measure of "Catholicness" based on surnames. In particular, we use full-count micro-level data from Ireland's 1901 census to calculate the fraction of individuals that were Catholic by surname; then, we merge that surname-based information with our linked dataset from the US census.[35] The details of the procedure are provided in Appendix B. This adds a useful dimension to the description of Irish immigrants and within-Irish heterogeneity in labor market outcomes. To our knowledge, it is the first analysis that uses information in Irish immigrants' surnames to study the connection between religion and economic status in the mid-nineteenth century US.

We also create a variable that measures the fraction of the 100 closest households that were headed by Irish men in 1850, as a way to gauge residence in Irish immigrant enclaves. In this, we follow Logan and Parman (2017) and rely on the idea that individuals close to one another on census manuscript pages were geographically proximate to one another. Finally, we created an indicator for having moved states between 1850 and 1880, which may shed light on whether geographic mobility was associated with upward occupational mobility for the Irish. Geographic mobility is a major theme in American economic development, especially in the nineteenth century, but whether it was a direct conduit for Irish immigrants' economic gains is an open question.

4. Results describing the selection and economic assimilation of Irish immigrants

We first describe the 1850 household characteristics of the children of famine-era immigrants. The 1850 data are informative regarding three aspects of the famine-era migration. First, differences in the human capital characteristics of fathers over arrival cohorts (i.e., in their literacy and numeracy) are informative about changes in migrant selection during the famine.[36]

---

[34] In the interest of brevity, we do not report the results for occupational rank in all specifications below. Where it is not shown (and where it is), it is qualitatively similar to the results with the wealth score.
[35] Ideally, we would have used data from an earlier Irish census. However, the 1901 census is the earliest with micro data available; the 1841 records were destroyed by fire in the 1920s (National Archives of Ireland n.d.).
[36] There is a concern that the census data do not reflect migrants' characteristics at time of arrival. Our focus on human capital mitigates this concern because it is less likely that immigrants' literacy and numeracy

Second, fathers' labor market outcomes are informative about the combination of changing migrant selection and difficult labor market conditions experienced by immigrants arriving during the Great Famine. Finally, and most importantly given our focus on intergenerational assimilation, all of the 1850 variables are informative about the early-life conditions experienced by the children of immigrants.

Then, we examine the children's outcomes in the labor market in 1880. By this time, the children are between 30 and 48 years old, a useful range for observing adult labor market outcomes in an intergenerational framework (Haider and Solon 2006, Long and Ferrie 2013, Feigenbaum 2018). Individual-level information on income or wealth is not available in 1880, and so we study occupational status as defined above. We are particularly interested in whether the famine-era Irish children narrowed gaps in status relative to the children of US natives and other immigrant groups. We are also interested in how they fared in a conditional sense—whether they did as well as observationally similar children from other groups. The ability to connect the children's adult outcomes to their childhood households is a useful feature of the dataset, and it is rare in studies of US immigrant assimilation.

Human capital and occupational status of household heads in 1850

Table 3 summarizes the 1850 variables for our main sample of linked and classified children, reported separately by ethnicity and arrival cohort. The first set of variables describes the individual's father in 1850. We collect literacy directly from the census, and we construct a measure of numeracy, which is an indicator for reporting an age that is not divisible by five.[37] We characterize occupational status with the wealth score, as described above, and with broad occupational groups. We also observe directly the household head's real property ownership in 1850. For the children, we observe whether they attended school in the year prior to the census (ages 5 to 15). We also create indicator variables for whether the child was US-born, and in the case of Irish children, whether the child was in utero during the famine (i.e., Irish-born 1846-1850).

Differences between Irish and non-Irish households are immediately apparent from these summary statistics, as are differences between pre-famine and famine-era arrival cohorts of Irish

---

changed after arrival than did other measures, such as occupation. However, because we do not have information on the distribution of these characteristics in Ireland, we are able to study only changes in migrant selection to the US, rather than levels of selection relative to Ireland's population.

[37] This variable does not indicate the numeracy of any individual in the sample, but differences between groups in the share reporting an age that ends in zero or five is informative.

immigrants. Relative to natives and to other immigrant groups, both arrival cohorts of the Irish were more likely to be illiterate, innumerate, and to hold unskilled occupations; as a result of the latter, they also had lower occupational wealth scores and ranks. The Irish also owned less property than other similarly tenured immigrants (or natives). Within the Irish, there was a considerable disadvantage for the famine-era arrivals compared to the pre-famine arrivals along these same dimensions. The famine-era Irish household heads were less literate and numerate, were more likely to hold an unskilled occupation, and had more Catholic surnames than earlier Irish immigrants. They were also more likely to reside in an urban area and in proximity to other Irish than were earlier Irish immigrants.

To characterize these differences more precisely, we estimate a regression specification of the form,

$$y_{ijk} = \alpha + \sum_j \sum_k \delta_{jk} + \mathbf{x}'_{ijk}\beta + \varepsilon_{ijk}, \tag{1}$$

where y_{ijk} is some outcome for individual i from ethnicity j and arrival cohort k. The coefficients of interest are δ_{jk}, which reflect differences in means between each immigrant-cohort group relative to white native-headed households (the excluded category), conditional on $\mathbf{x}_{ijk}$. The controls $\mathbf{x}_{ijk}$ in the baseline case are simply a quartic in the age of the head of household, or indicators for the individual's age for the analysis of children's school attendance.

We also test the null hypothesis that the difference in δ_{jk} coefficients across Irish arrival cohorts is equal to the difference for the British and Germans.[38] This provides a kind of difference-in-differences perspective on the pre- and post-1846 Irish arrivals compared to the pre- and post-1846 patterns for other groups. The idea is simply to see whether the differences over arrival cohorts are common to all immigrant source countries (e.g., reflecting the duration of time in the US or common trends in transport costs), or if there is something different about the change in Irish outcomes relative to the contemporaneous changes for the British and Germans. We do not interpret such differences as strong evidence of a causal impact of the Irish famine per se, but they would be consistent with such an effect.

Table 4 presents estimates of δ_{jk} for the 1850 characteristics. Results in columns (1) and (2) indicate that Irish household heads had significantly less human capital than US-native heads of

---

[38] Formally, this is a joint test of the null hypotheses that $\delta_{I1} - \delta_{I0} = \delta_{B1} - \delta_{B0}$ and $\delta_{I1} - \delta_{I0} = \delta_{G1} - \delta_{G0}$, where δ_{j1} represents the indicator for the famine arrival cohort of ethnicity j and δ_{j0} represents the indicator for the pre-famine arrival cohort of ethnicity j.

household, who were in turn less likely to be literate but more likely to be numerate than British and German heads. Conditional on age, there is evidence of a large decrease in literacy and numeracy between the pre-famine and famine-era Irish compared to natives (-0.05 to -0.11 for literacy; -0.18 to -0.25 for numeracy), with no similar changes evident among the other nationalities. The difference-in-differences test across ethnicities is statistically significant in both columns, indicating that the worsening observable quality of Irish migrants during the famine was not an artifact of cohort-specific shocks or trends affecting all nationalities. Instead, it is plausible that famine conditions caused a marked deterioration in migrants' quality, as gauged by literacy and numeracy.

This worsening selection is consistent with Mokyr and Ó Gráda's (1982) findings of increasing age heaping among Irish immigrants in passenger lists after 1845. It is also consistent with the predictions of a canonical Roy (1951)-Borjas (1987) model of migrant selection. As discussed above, positive selection from Ireland is believed to have prevailed before the famine. In this setting, a widespread and persistent negative shock to expected income at home, such as the potato blight, would then tend to draw *lower*-skilled workers into the migrant flow at the margin, leading to a decline in average migrant quality.[39]

Similar patterns are evident for the heads' labor market outcomes as shown in columns (3)-(6). Both arrival cohorts of Irish household heads were more likely to hold unskilled occupations and had lower (log) occupational wealth scores and lower occupational ranks than natives and other immigrant groups, conditional on age. They were also less likely than natives to be farmers. We find strong evidence of a deterioration in labor market status between the pre-famine and famine-era Irish, which is unmatched by the other immigrant groups. For example, while pre-famine Irish household heads were about 29 percentage points more likely to hold unskilled occupations than native household heads, the famine Irish household heads were over 56 percentage points more likely. Such large differences are not present in the other ethnicities (i.e., the difference-in-differences *F*-tests indicate a significantly greater deterioration over arrival cohorts for the Irish than

---

[39] Richer models of selection that allow migration costs to vary by level of skill (Chiquiar and Hanson 2005, McKenzie and Rapoport 2010) suggest that high migration costs may prevent the migration of the poorest workers. Financial constraints on migration were relevant for the Irish (Hatton and Williamson 1994), and the poorest Irish could not afford to migrate to the US during the famine (Ó Gráda 1999, pp. 104-114). Recent work has also emphasized that networks shape migrant selection and migration responses to shocks (Wegge 1998, McKenzie and Rapoport 2010; Mahajan and Yang 2017; Spitzer 2018; Spitzer and Zimran 2018). If Irish migrants were positively selected before the famine, it is possible that relatively well-off families or places had stronger migrant networks to facilitate emigration (as well as more resources to finance the move). In theory, this might skew migrant selection in a more positive direction, but on net we observe selection become more negative.

for other groups).[40] This suggests that the Irish pattern is unlikely to be solely the result of the famine Irish having been in the US for a shorter time than the pre-famine Irish and thus had less time to upgrade occupations.

Results for property ownership and the school attendance of children are less stark, as shown in columns (7) and (8). We do observe a decline in the value of real property owned between pre-famine and famine-era Irish; this is not surprising given that the famine-era Irish had been in the US for a shorter period. Similar cross-cohort declines are apparent for the Germans and British, and although it is still possible to reject the difference-in-differences null hypothesis, the decline in property holdings over cohorts is actually greatest among the British. There is also a decline in school attendance over arrival cohorts, conditional on children's age, for each immigrant group. The children of the famine-era Irish were substantially less likely to attend school than similarly aged children of US natives (by over 14 p.p.), but in this respect, they fared better than German children.[41] It is not possible to reject the null hypothesis that the change in schooling across the Irish cohorts was the same as that of the other ethnicities.

As with the deterioration in migrant quality, the worsening labor market outcomes of the famine Irish relative to other groups can be rationalized in a cogent theoretical framework. In addition to the deterioration in Irish migrants' observable human capital, the sheer volume of the immigrants' arrival cohort may have affected their labor market outcomes. The size of the famine-era arrival cohort may have hindered their economic progress by congesting networks that would ordinarily facilitate their transition. Beaman (2012) shows, in theory and in a dataset of recent refugees to the US, that pre-existing stocks of immigrants can facilitate labor market assimilation for newly arrived countrymen by providing information about job opportunities. Relatively large cohorts of new immigrants, however, exacerbate competition for information in the network and may lead to worse labor market outcomes than would otherwise occur, at least in the short term. This mechanism is potentially applicable to the case of the famine-era Irish migrants, who, upon arriving in the US, would have entered labor markets that were populated by many previous Irish migrants, but also glutted with a large cohort of new arrivals.

---

[40] It is possible that if the comparison were in the relative difference in the probability of holding an unskilled occupation rather than the absolute difference, that results might differ. However, given that our goal is to describe the childhood household characteristics of the sons of the famine migrants, we are more concerned with the absolute differences, as they reflect the number of children whose head of household was unskilled.
[41] The greater school attendance of the famine Irish relative to the famine Germans is largely explained by differences in the geographic distribution of the two groups. That is, controlling for state fixed effects reduces the gap between the two groups from 9.5 to 1.4 percentage points. This likely reflects a greater tendency for Germans to live in the Midwest, where schooling would have been less readily available.

In sum, the new dataset of linked census records reveals that the famine-era Irish household heads were notably less educated and attained extremely low occupational status compared to other contemporaneous immigrant groups, earlier immigrants from Ireland, and US natives. There was nothing about the famine refugees that would favorably impress policymakers concerned primarily with immigrant "quality." Thus, their children began life at a severe disadvantage relative to others in the US during an era with minimal public assistance and ongoing mass immigration from Europe. Under these circumstances, how far, if at all, were the children of the famine-era Irish able to narrow the gap in economic status relative to their peers as they moved into adulthood?

Assimilation by immigrants' children

We now shift the focus to the adult labor market outcomes in 1880 of individuals first observed as children in 1850. Table 5 presents summary statistics for various occupational measures. As with the 1850 data, clear differences are evident between the Irish and the non-Irish, as well as between the famine-era and pre-famine Irish immigrants, all in similar directions as observed for the household heads in 1850.

Table 6 presents regressions of the form of equation (1) for the children's adult labor market outcomes in 1880; again, the only baseline control variable is a quartic in the individual's age. The children of famine-era Irish immigrants were 24 percentage points more likely to hold an unskilled occupation than children of natives, translating into a 48 log point or 11 percentile rank disadvantage in terms of occupational score. For all ethnicities, occupational outcomes were worse for the children of famine-era arrivals than for children of earlier arrivals. However, whereas these differences were small for the British and Germans, the famine-era Irish children fared much worse than the children of the pre-famine Irish (e.g., 16 percentage points more likely to have an unskilled occupation). The difference-in-differences tests reveal that the famine-era Irish disadvantage relative to earlier Irish arrivals is significantly different from the pattern observed in other immigrant groups; it does not merely reflect worse outcomes for the children of later arriving immigrants. Thus, it is clear that the distinct disadvantage of the famine-era Irish arrivals persisted to the adult outcomes of the children, 30 years after the famine.

The results of Tables 5 and 6 taken in isolation paint a fairly negative picture of the economic assimilation of the famine-era Irish. However, when considered in comparison to Tables 3 and 4, the results indicate that considerable assimilation occurred from one generation to the next. Figure 3 directly compares the coefficients from the estimation of equation (1) for the heads' outcomes in 1850 to those for the (adult) children's outcomes in 1880. For both the probability of holding an

unskilled occupation in Figure 3(a) and the occupational status score in Figure 3(b), the famine-era Irish disadvantage relative to natives in 1880 is much smaller than the gap observed for their fathers in 1850. In this sense, there is strong evidence of convergence in labor market outcomes between immigrants and natives, even for the most impoverished and desperate group of mid-nineteenth century immigrants. Below, we discuss this finding in light of other studies of immigrants' economic assimilation.

Outcomes conditional on 1850 characteristics

Given their relatively poor starting point, it is unsurprising that Irish immigrants' children had worse labor market outcomes than the average child of natives. A different perspective on assimilation is afforded by controlling for the adverse childhood conditions faced by the famine-era Irish children. In studying intergenerational mobility for immigrants, scholars often rely on comparisons of cross-sectional data observed roughly a generation apart (e.g., comparing foreign-born workers' outcomes in one census to outcomes for US-born workers who report having foreign-born parents decades later) (see Borjas 1993; Card 2005; Abramitzky, Boustan, Eriksson 2014). An advantage of the linked census dataset studied here is that we observe at the individual-level both the childhood household characteristics and the adult outcomes. That is, we can study the immigrants' children's outcomes conditional on a rich set of early life circumstances. To compare the immigrants' children to observationally similar natives' children, we repeat the estimation of equation (1) but expand the vector of control variables in $\mathbf{x}_{ijk}$ to include quartic polynomials in the log of the occupational wealth score of the head of household (in 1850), the head's real property holdings, and the head's age; indicators for head's literacy and numeracy; indicators for head's broad occupational category; and indicators for urban residence, school attendance, and state of residence.

Each panel in Figure 4 presents the coefficient estimates for the ethnicity-cohort indicators (δ_{jk}) for three specifications. The first set of estimates, labeled "Baseline," are the coefficients obtained when controlling only for the quartic in age (i.e., they are the same as in Table 6). The second, labeled "Controls," are the coefficient estimates obtained when controlling for the 1850 household characteristics listed above. The last set, labeled "Controls plus nativity," are the estimates obtained with all the controls mentioned above and an indicator for US birth.[42] US birth might be advantageous to later life outcomes, as it implies no exposure to a long trans-Atlantic voyage and full exposure to American educational institutions, language, and society (Hatton 1997;

---

[42] The coefficients of the regressions with the expanded sets of controls are presented in Appendix Table C.1.

Alexander and Ward 2018).

The famine-era Irish children's disadvantage in occupational status relative to other groups is greatly diminished, but still visible, when the regression includes background controls. In the baseline results, recall that the sons of the famine-era Irish were 24 percentage points more likely than the sons of natives to have an unskilled occupation, and their occupations scored 48 log points lower than natives' sons on average. Once the background control variables are added to the specification, the gaps decline to 14 percentage points and 27 log points for unskilled occupations and occupation scores, respectively. Additionally controlling for native birth further reduces the gaps to just 8 percentage points and 12 log points. The difference-in-differences tests indicate that the change over Irish cohorts was statistically different than the changes for other immigrant groups.

A Gelbach decomposition indicates that the native birth variable and the measures of the father's occupational status in 1850 (the indicators for occupational category and the quartic in the occupational wealth score) are responsible for the greatest portion of the decline in the magnitude of the famine Irish coefficient relative to the more parsimonious baseline (Gelbach 2016).[43] In this sense, the relatively poor early life economic circumstances of the children of the famine-era Irish immigrants were quantitatively important to their subsequent outcomes. Separately, it is notable that the "controls plus nativity" results for the English and Germans are small but indicate slightly better performance for the immigrants' children than for observationally similar natives.

Interpreting the residual gap for the famine-era Irish is, of course, difficult in this kind of analysis. It is consistent with both unobservably worse background conditions for the Irish children—circumstances that are not fully captured in $\mathbf{x}_{ijk}$—and discrimination that affects labor market outcomes. The main point here is simply that the children of famine-era Irish fared only slightly worse than children of natives with observationally similar backgrounds, despite facing obstacles that might have been particularly relevant to the children of Irish immigrants. In addition to hindrances associated with relatively low human capital parents, they might also have experienced overcrowding of schools, housing, and public goods; a nativist backlash against immigrants, especially Irish Catholics; intense labor market competition from new waves of immigrants; and, for some, the physical and psychological trauma of the famine itself.

---

[43] Gelbach (2016) provides a method to decompose the change in a coefficient (in our case, the coefficient for the famine Irish indicator) that results from the addition of a vector of controls $\mathbf{x}_{ijk}$ into the separate contribution of each component of that vector. Unlike approaches that address coefficient stability with the sequential addition of controls, this decomposition does not depend on the order in which controls are added.

Discussion

The results described in this section answer simple but fundamental questions about immigrants and their children in the early Age of Mass Migration. The answers are made possible by advances in data resources and techniques that allow us to create linked census records at large scale. Our interpretation emphasizes that the famine-era Irish were more negatively selected than prior Irish immigrants and arrived in a large wave on the northeastern seaboard. Their poverty and relative lack of human capital square with historical accounts; but the new dataset offers much more detail and precision than previously available. It can be no surprise that the Irish children did not completely close the gap in average economic status relative to natives' children by 1880. Yet under the circumstances, the strong evidence of convergence is notable. They may have benefited from arriving at a time when whites' intergenerational mobility was at a historically high level compared to later periods (Long and Ferrie 2013). Movement toward the mean from the bottom of the occupational distribution was fairly common for low-status whites, but such strong convergence is neither mechanical nor to be taken for granted, as evidenced by the history of black men's intergenerational mobility (Collins and Wanamaker 2017).

Our finding of substantial but incomplete intergenerational convergence by the famine Irish is consistent with important studies of mid- to late-twentieth century immigrants, which typically rely on cross-sectional data sources (as described above). For instance, our finding that the children of famine-era Irish migrants fared worse than other groups is consistent with Borjas (1993, pp. 122-128), which finds evidence of substantial but incomplete income convergence by the children of relatively poor immigrant groups. Our convergence results are also roughly consistent with Card (2005), which finds that second-generation immigrants tend to converge on or even overtake the educational attainment of US natives (pp. F318-F319). Second-generation Mexican immigrants, for instance, eliminated about 80 percent of the education gap faced by first-generation fathers (Card 2005, p. F319).[44] For an earlier period, Hatton and Williamson (1998, pp. 128-129) describe intergenerational gains in occupational status based on the reports of the US Immigration Commission (1911).

Our findings also complement evidence provided in Abramitzky, Boustan, and Eriksson (2014, pp. 498-500), who study immigrant assimilation in the early to mid-twentieth century. Their paper is primarily focused on describing assimilation within a generation using linked data, but it

---

[44] Educational attainment cannot be studied with the US census in the nineteenth century. Even the literacy variable is unavailable for the 1880 complete count data.

also provides some insight regarding the second generation's outcomes using cross-sectional data. They find persistence in outcomes across generations, such that groups that were initially disadvantaged (or advantaged) relative to natives in terms of occupational status tended to maintain that position in the second generation. Our results for the Irish are consistent with this finding, but our results for the pre-famine cohorts of British and German immigrants (Table 6) are somewhat different; their children appear to have surpassed those of natives by 1880, though only slightly. Abramitzky, Boustan, and Eriksson (2014, p. 499) also find that the second generation fared better than the first for about half of the 16 sending countries in their dataset, with second-generation German, English, Scottish, and Welsh men faring worse, and second-generation Irish men faring better. These mixed results contrast with the fairly consistent patterns of improvement in the relative status of the immigrant groups that we document in Figure 3.

There are several possible reasons for differences between our findings and those of Abramitzky, Boustan, and Eriksson (2014). The most obvious is that we study different cohorts and periods of time, with our focus being on the earliest cohorts of the Age of Mass Migration and their focus being on the later cohorts. The later cohorts were drawn from a wider range of sending countries, faced lower transport costs, and arrived in an economy that had been transformed by rapid industrialization and urbanization. It is possible that features of the mid-nineteenth century US economy facilitated the occupational advancement of poor immigrants relative to later periods. For instance, the frontier remained open, and both native-born workers and immigrants moved westward in large numbers, possibly in response to continuing immigrant inflows (Ferrie 1995). Although we cannot observe whether the men in our sample participated in the US Civil War, it is possible that the tumult of the period provided opportunities for economic mobility. Finally, return migration was far less common in the period that we study than in the later years of the Age of Mass Migration, which may have had implications for immigrants' investment in US-specific human capital (Cortes 2004).

5. Correlates of economic mobility among the famine Irish

In this section, we investigate heterogeneity in economic mobility among the famine-era Irish immigrants, offering the first such analysis of this group. The analysis aims to highlight characteristics that were strongly associated with differences in economic mobility for the sons of famine-era immigrants. We begin with a focus on variables that from the child's perspective were pre-determined, such as the father's literacy, whether the son was born in the US or born in Ireland, whether an Irish-born child was born during or before the famine, the share of neighbors who were Irish, and the "Catholicness" of his surname. These correspond to prominent themes in the literature

on immigrant assimilation regarding human capital, enclave residence, and social distance. Variation in these characteristics is not quasi-random, and our interpretations are therefore descriptive not causal; even so, the results provide novel insights into the group's pattern of economic mobility.

We estimate an equation of the form

$$y_i^c = \eta y_i^h + \mathbf{z}_i'\gamma + f(a_i^c, a_i^h) + \varepsilon_i$$

where y_i^c is an occupational outcome for the child in 1880 (an indicator for unskilled occupation or the log of the occupational wealth score), y_i^h is the same occupational outcome variable for the child's father in 1850, $\mathbf{z}_i$ is the vector of regressors of interest, and $f(a_i^c, a_i^h)$ represents quartic polynomials in the age of the head of household in 1850 and the age of the child in 1880. The γ coefficients measure the conditional correlation of the sons' economic status with elements of $\mathbf{z}_i$, taking account of differences in fathers' occupational status, age variables, and other covariates. The sample is restricted to sons of the famine-era Irish immigrants.

Table 7 reports the results.[45] Native (US) birth was associated with greater upward mobility relative to children born in Ireland before the famine (the implicit comparison group). This is consistent with findings regarding age at arrival and assimilation later in the Age of Mass Migration (Hatton 1997, Alexander and Ward 2018). In this context, native birth implies no exposure to famine conditions in Ireland or a long trans-Atlantic voyage, and full exposure to US educational and other institutions. These children would have acquired US-specific human capital from the start. Those born in Ireland during the famine did not have statistically different outcomes from those born earlier in Ireland, who would still have been young at the famine's onset.[46] They did, however, fare worse than similarly aged US-born children of famine-era immigrants (i.e., comparing to the coefficient on native-born). This difference is particularly interesting since both groups spent nearly their entire lives in the US, but of course the Irish-born children experienced the famine directly.

---

[45] The number of observations in this Table is reduced relative to the number of famine Irish described in Tables 3 and 5. The difference is accounted for by individuals whose surnames were not sufficiently common in the 1901 Irish census to determine their Catholicness measure (we require at least 200 individuals with similar names in the 1901 Irish census). Columns (1) and (3) of this Table include controls for the 1850 head of household's occupational category. Columns (2) and (4) control for the log of the 1850 head of household's occupational wealth score. The coefficient on this variable, which is the intergenerational elasticity of the occupational wealth score, is 0.087 (se=0.038) in both columns. As the occupational wealth measure is novel to this paper, there is no reason to expect that this elasticity would resemble intergenerational elasticities estimated in other contexts.

[46] We recognize that there are individuals who would have been born in the US but who would have been in utero in Ireland during the famine. Unfortunately, the fact that we do not observe the actual year of arrival (only a range) prevents us from identifying these individuals. We thus consider the coefficient on Famine In Utero to be an understatement of the correlation in question.

In other settings, research has shown that immigrants tend to negatively select into enclaves, but that conditional on this sorting, immigrant enclaves may be beneficial (Cutler, Glaeser, and Vigdor 2008; Damm 2009; Edin, Fredriksson, and Åslund 2003) or detrimental (Battisti, Peri, Romiti 2018; Eriksson 2018) for employment outcomes. Table 7 shows that residing in a home with a high share of Irish neighbors was associated with worse outcomes for children, but the estimates are very imprecise. Thus, there is no clear evidence that links residence in an Irish enclave to worse (or better) outcomes, conditional on 1850 observables.

The predominantly Catholic character of the famine immigrants was highly controversial at the time. Their religion was seen as a barrier to assimilation in a predominantly Protestant society, and it led to open hostility (Anbinder 1992; Hirota 2017). As described above, we used full-count micro-level data from Ireland's 1901 census to calculate the fraction of individuals who were Catholic by surname; we then merged that information with the linked dataset. Having a "more Catholic" surname is strongly negatively correlated with sons' labor market outcomes, even after conditioning on fathers' status and literacy and the Irishness of the child's neighborhood. We caution that this result is difficult to interpret. It is plausible that it reflects poor treatment of Irish Catholics in labor markets; it is also plausible that it reflects unobserved differences in background characteristics and human capital accumulation that are not sufficiently controlled for in the regression. That said, this statistical association is strong. A typical "Catholic" surname had a Catholicness of 0.95 whereas a typical "Protestant" surname had a Catholicness of 0.18.[47] This difference translates to Catholics being about 11 percentage points more likely to hold an unskilled occupation (relative to a group mean of 42 percent) and having a roughly 20 percent lower occupational wealth score. To our knowledge, this is the first quantitative analysis of nineteenth-century US immigrants' religious affiliation and labor market outcomes. Given the strength of the association revealed here and the salience of religious differences in nineteenth-century America, we suggest that this is an important area for future research.

Finally, in a separate specification we add an indicator variable for interstate migrants. The connection between economic and geographic mobility is salient to mid-nineteenth century economic history; in the US context, the prominent "safety valve" theory argues that geographic mobility was an important channel for economic mobility by the urban poor (e.g., Ferrie 1997b; Stewart 2006). In recent work, Pérez (2018) has shown that geographic mobility may have been an important contributor to economic mobility in nineteenth century Argentina, another major destination for

---

[47] That is, plotting the density of the Catholicness variable reveals local modes at 0.95 and at 0.18.

European migrants. It is interesting, and perhaps surprising, that the results in Table 7 show that there is no evidence that geographic mobility, as measured by moving states, was an important avenue for occupational advance among the famine-era Irish children. Most of the famine-era migrants' children did relocate between 1850 and 1880 (Table 5), but their relocation was not associated with occupational upgrades relative to their father. It remains possible, however, that the Irish benefited as a group from the geographic mobility of others, or that negative selection into internal migration may have obscured gains (Ferrie 1997b; Stewart 2006).

6. Conclusions

This paper builds a large dataset of linked US census records to shed new light on the assimilation of a large group of desperate, poor, and culturally distinct immigrants—the Irish fleeing the Great Famine. The complete count census-based approach to linking and studying the Irish and other immigrants provides several advantages relative to the existing literature. The linked dataset includes immigrants regardless of their port of arrival (including arrivals via Canada), is nationally representative, distinguishes clearly between those arriving before and during the Irish famine, and relies on measures of human capital and occupation that are consistently enumerated across groups. The linked data ensure that we study a fixed group of individuals over time, which avoids spurious results driven by changes in sample composition, as discussed by Lubotsky (2007) and Abramitzky, Boustan, and Eriksson (2014). The new dataset forms the basis for a more definitive picture of migrant selection during the Irish famine and immigrant assimilation patterns in the early Age of Mass Migration.

We find evidence of a decline in human capital among the famine-era Irish immigrants relative to previous Irish migrants and relative to other immigrant and native groups. As a result of this deteriorating selection, the migrants' poor labor market outcomes, and the backlash against Irish immigrants, the children of the famine Irish faced long odds. But as adults they significantly narrowed the gap in occupational status relative to natives in comparison to their fathers' starting point, and they nearly kept pace with sons of US natives whose fathers were similarly situated in 1850. In this sense, there is strong evidence of economic assimilation by the famine Irish. This is the paper's main finding. Among the children of the famine-era Irish immigrants, conditional on observables in 1850, we find differences in occupational outcomes depending on the Catholicness of their surname; we also find that children born in Ireland, even those who spent nearly their entire life in the US, fared worse that those born in the US, potentially reflecting exposure to famine conditions.

The paper's findings advance the literature on the economics of migration in several ways.

First and most directly, it adds to our knowledge of the early decades of the Age of Mass Migration (Hatton and Williamson 1998, Cohn 2009), especially Irish migration to the US during the Great Famine (Mokyr and Ó Gráda 1982; Miller 1985; Ó Gráda and O'Rourke 1997; Ferrie 1999). Prior research on this period has been hampered by data limitations that are inherent to mid-nineteenth century records. The census did not inquire about the year of immigration in this period, and passenger lists from ships are incomplete in their coverage and difficult to link. Consequently, much of the recent surge in research on the Age of Mass Migration has focused on the early twentieth century, by which time the US economy had passed through a remarkable period of economic growth and transformation. The complete count census files allow us to form relatively large, conservatively linked samples and to classify arrival cohorts for many households in 1850 using the age and birthplace of children. We can trace labor market assimilation at the micro level by connecting fathers and sons over a period of 30 years. This distinguishes our work from that of Ferrie (1999), as do our detailed comparisons of the famine-era arrivals to those who arrived earlier.

The paper also contributes to the extensive economics literature on immigrants' labor market assimilation. The influx of immigrants circa 1850 was very large relative to the size of the US population; it is, therefore, a particularly important event in the annals of mass migration, albeit understudied for reasons discussed above. Apart from its sheer magnitude, this historical setting is appealing because it allows scholars to study assimilation in the absence of federal policy barriers to entry, legal curbs on discrimination, or a robust social safety net. The observed assimilation patterns are primarily the product of the economic and social forces, rather than of screening or support policy. Moreover, unlike much of the literature on immigrant children's outcomes that relies on cross-sectional data sources, the linked dataset built in this paper allows a direct and detailed connection to be made between childhood circumstances and adult labor market outcomes.

The paper's final contribution is to the economics of migration in cases of natural or man-made disasters. In addition to research on the quantity of migration in response to large negative shocks, the selection of refugees into migration and their subsequent labor market assimilation are important concerns for receiving countries. This paper speaks to both selection and assimilation questions from the perspective of perhaps the largest wave of disaster refugees that the US has ever absorbed. Despite the desperate conditions in which the famine-era migrants arrived and their relatively low levels of human capital, their children showed strong evidence of "assimilation and contribution" in the labor market. Thus, the paper provides useful historical perspective at a time when refugees are once again viewed by many with skepticism and scorn.

References

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2012). "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration." *American Economic Review* 102:5, pp. 1832-1856.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2014). "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration." *Journal of Political Economy* 122:3, pp. 467–506.

Abramitzky, Ran, Leah Platt Boustan, and Katherine Eriksson (2016). "Cultural Assimilation During the Age of Mass Migration." NBER Working Paper 22381.

Alexander, Rohan and Zachary Ward (2018). "Age at Arrival and Assimilation during the Age of Mass Migration." *Journal of Economic History* 78:3, pp. 904-937.

Anbinder, Tyler (1992). *Nativism and Slavery: The Northern Know Nothings and the Politics of the 1850s.* New York: Oxford University Press.

Anbinder, Tyler and Hope McCaffrey (2015). "Which Irish Men and Women Immigrated to the United States during the Great Famine Migration of 1846–54?" *Irish Historical Studies* 39:156, pp. 620–642.

Anbinder, Tyler, Cormac Ó Gráda and Simone A. Wegge (2017). "Networks of Opportunity: How Emigrants from the Great Irish Famine Survived, and Eventually Thrived, in New York and Beyond." Mimeo., University College, Dublin.

Bailey, Martha, Connor Cole, Morgan Henderson, and Catherine Massey (2017). "How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth." NBER Working Paper 24019.

Barde, Robert, Susan B. Carter, and Richard Sutch (2006). "Table Ad106-120: Immigrants, by country of last residence—Europe, 1820-1997." In *Historical Statistics of the United States*. Susan B. Carter, Scott Sigmund Gardner, Michael R. Haines, Alan L. Olmstead, Richard Sutch, and Gavin Wright (ed.). Cambridge: Cambridge University Press, pp. 1.560-1.563.

Battisti, Michele, Giovanni Peri, and Agnese Romiti (2018). "Dynamic Effects of Co-Ethnic Networks on Immigrants' Economic Success." NBER Working Paper 22389.

Beach, Brian, Joseph P. Ferrie, Martin Saavedra, and Werner Troesken (2016). "Typhoid Fever, Water Quality, and Human Capital Formation." *Journal of Economic History* 76:1, pp. 41-75.

Beaman, Lori A. (2012). "Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the US." *Review of Economic Studies* 79:1, pp. 128–161.

Blum, Matthias and Claudia Rei (2018). "Escaping Europe: Health and Human Capital of Holocaust Refugees." *European Review of Economic History* 22:1, pp. 1-27.

Borjas, George J. (1987). "Self-Selection and the Earnings of Immigrants." *American Economic Review* 77:4, pp. 531–553.

Borjas, George J. (1992). "Ethnic Capital and Intergenerational Mobility." *Quarterly Journal of Economics* 107:1, pp. 123-150.

Borjas, George J. (1993). "The Intergenerational Mobility of Immigrants." *Journal of Labor Economics* 11:1, pp. 113-135.

Boustan, Leah Platt, Matthew E. Kahn, and Paul W. Rhode (2012). "Moving to Higher Ground: Migration Response to Natural Disasters in the Early Twentieth Century." *American Economic Review, Papers and Proceedings* 102:3, pp. 238–244.

Caponi, Vincenzo (2011). "Intergenerational Transmission of Abilities and Self-Selection of Mexican Immigrants." *International Economic Review* 52:2, pp. 523-547.

Card, David (1990). "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* 43:2, pp. 245-257.

Card, David (2005). "Is the New Immigration Really So Bad?" *Economic Journal* 115, pp. F300-F323.

Card, David, John DiNardo, and Eugena Estes (2000). "The More Things Change: Immigrants and the Children of Immigrants in the 1940s, the 1970s, and the 1990s." In *Issues in the Economics of Immigration.* George J. Borjas (ed.). Chicago: University of Chicago Press, pp. 227-270.

Chiquiar, Daniel and Gordon H. Hanson (2005). "International Migration, Self-Selection, and the Distribution of Wages: Evidence from Mexico and the United States." *Journal of Political Economy* 113:2, pp. 239-281.

Chiswick, Barry R. (1978). "The Effect of Americanization on the Earnings of Foreign-Born Men." *Journal of Political Economy* 86:5, pp. 897-921.

Cohn, Raymond L. (1995). "Occupational Evidence on the Causes of Immigration to the United States, 1836-1853." *Explorations in Economic History* 32, pp. 383-408.

Cohn, Raymond L. (2009). *Mass Migration Under Sail: European Immigration to the Antebellum United States*. Cambridge: Cambridge University Press.

Collins, William J. and Marianne H. Wanamaker (2017). "Up from Slavery? African American Intergenerational Mobility since 1880." NBER Working Paper 23395.

Commissioners Appointed to Take the Census of Ireland (1843). *Report of the Commissioners Appointed to Take the Census of Ireland for the Year 1841*. Dublin: Alexander Thom.

Connor, Dylan S. (2018). "The Cream of the Crop? Geography, Networks, and Irish Migrant Selection in the Age of Mass Migration." *Journal of Economic History* Forthcoming.

Cortes, Kalena E. (2004). "Are Refugees Different from Economic Immigrants? Some Empirical Evidence on the Heterogeneity of Immigrant Groups in the United States." *Review of Economics and Statistics* 86:2, pp. 465-480.

Cutler, David M., Edward L. Glaeser, and Jacob L. Vigdor (2008). "When Are Ghettos Bad? Lessons from Immigrant Segregation in the United States." *Journal of Urban Economics* 63, pp. 759-774.

Damm, Anna Piil (2009). "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence." *Journal of Labor Economics* 27:2, pp. 281-314.

Delaney, Enda (2015). "Refugee Crisis Has Chilling Parallels with the Famine Years." *Irish Times*.

Edin, Per-Anders, Peter Fredriksson, and Olof Åslund (2003). "Ethnic Enclaves and the Economic Success of Immigrants—Evidence from a Natural Experiment." *Quarterly Journal of Economics* 118:1, pp. 329-357.

Eriksson, Katherine (2018). "Ethnic Enclaves and Immigrant Outcomes: Norwegian Immigrants during the Age of Mass Migration." NBER Working Paper 24763.

Ernst, Robert (1949). *Immigrant Life in New York City, 1825-1863*. Syracuse, NY: Syracuse University Press.

Evans, William N. and Daniel Fitzgerald (2017). "The Economic and Social Outcomes of Refugees: Evidence from the ACS." NBER Working Paper 23498.

Federal Chancellor (2018). "Towards a Cosmopolitan and Diverse Germany." https://www.bundeskanzlerin.de/bkin-en/news/towards-a-cosmopolitan-and-diverse-germany-1141998. Accessed November 13, 2018.

Feigenbaum, James J. (2018). "Multiple Measures of Historical Intergenerational Mobility: Iowa 1915 to 1940." *Economic Journal* 128, pp. F446-F481.

Ferenczi, Imre and Walter F. Wilcox (1929). *International Migrations*. New York: National Bureau of Economic Research.

Ferrie, Joseph P. (1994). "The Wealth Accumulation of Antebellum European Immigrants to the US, 1840–60." *Journal of Economic History* 54:1, pp. 1–33.

Ferrie, Joseph P. (1995). "Up and Out or Down and Out? Immigrant Mobility in the Antebellum United States." *Journal of Interdisciplinary History* 26:1, pp. 33–55.

Ferrie, Joseph P. (1996). "A New Sample of Males Linked from the Public Use Microdata Sample of the 1850 US Federal Census of Population to the 1860 US Federal Census Manuscript Schedules." *Historical Methods* 29:4, pp. 141-156.

Ferrie, Joseph P. (1997a). "The Entry into the US Labor Market of Antebellum European Immigrants, 1840–1860." *Explorations in Economic History* 34, pp. 295–330.

Ferrie, Joseph P. (1997b). "Migration to the Frontier in Mid-Nineteenth Century America: A Re-Examination of Turner's 'Safety Valve.'" Mimeo., Northwestern University.

Ferrie, Joseph P. (1999). *Yankeys Now: Immigrants in the Antebellum United States, 1840–1860*. New York: Oxford University Press.

Geary, Frank (1996). "Regional Industrial Structure and Labour Force Decline in Ireland between 1841 and 1851." *Irish Historical Studies* 30:118, pp. 167-194.

Gelbach, Jonah B. (2016). "When Do Covariates Matter? And Which Ones, and How Much?" *Journal of Labor Economics* 34:2, pp. 509-543.

Guinnane, Timothy W. (1997). *The Vanishing Irish: Households, Migration, and the Rural Economy in Ireland, 1850-1914*. Princeton: Princeton University Press.

Haider, Steven and Gary Solon. 2006. "Life-Cycle Variation in the Association between Current and Lifetime Earnings." *American Economic Review* 96:4, pp. 1308-1320.

Handlin, Oscar (1991 [1941]). *Boston's Immigrants, 1790-1880: A Study of Acculturation, Revised and Enlarged Edition*. Cambridge, MA: Belknap Press.

Hatton, Timothy J. (1997). "The Immigrant Assimilation Puzzle in Late Nineteenth-Century America." *Journal of Economic History* 57:1, pp. 34-62.

Hatton, Timothy J. (2012). "Refugee and Asylum Migration to the OECD: A Short Overview." IZA Discussion Paper 7004.

Hatton, Timothy J. and Jeffrey G. Williamson (1994). "What Drove the Mass Migrations from Europe in the Late Nineteenth Century?" *Population and Development Review* 20, 3, pp. 533-559.

Hatton, Timothy J. and Jeffrey G. Williamson (1998). *The Age of Mass Migration: Causes and Economic Impact*. New York: Oxford University Press.

Hirota, Hidetaka (2017). *Expelling the Poor: Atlantic Seaboard States and the Nineteenth-Century Origins of American Immigration Policy*. New York: Oxford University Press.

Hornbeck, Richard (2012). "The Enduring Impact of the American Dust Bowl: Short- and Long-Run Adjustments to Environmental Catastrophe." *American Economic Review* 102:4, pp. 1477–1507.

Kelly, John (2018). "Transcript: White House Chief of Staff John Kelly's Interview with NPR." *NPR.org.*

Logan, Trevon D. and John M. Parman (2017). "The National Rise in Residential Segregation." *Journal of Economic History* 77:1, pp. 127-170.

Long, Jason and Joseph P. Ferrie (2013). "Intergenerational Occupational Mobility in Great Britain and the United States since 1850." *American Economic Review* 103:4, pp. 1109-1137.

Long, Jason and Henry E. Siu (2018). "Refugees from Dust and Shrinking Land: Tracking the Dust Bowl Migrants." *Journal of Economic History* Forthcoming.

Lubotsky, Darren (2007). "Chutes or Ladders? A Longitudinal Analysis of Immigrant Earnings." *Journal of Political Economy* 115: 5, pp. 820-867.

Mahajan, Parag and Dean Yang (2017). "Taken By Storm: Hurricanes, Migrant Networks, and US Immigration." NBER Working Paper 23756.

McInnis, Marvin (2000). "The Population of Canada in the Nineteenth Century." In *A Population History of North America*. Michael R. Haines and Richard H. Steckel (eds.) Cambridge University Press. Chapter 12, pp. 529-599.

McKenzie, David and Hillel Rapoport (2010). "Self-Selection Patterns in Mexico-US Migration: The Role of Migration Networks." *Review of Economics and Statistics* 92:4, pp. 811-821.

Miller, Kirby A. (1985). *Emigrants and Exiles: Ireland and the Irish Exodus to North America*. New York: Oxford University Press.

Minnesota Population Center (2017). *North Atlantic Population Project: Complete Count Microdata: Version 2.3* [Machine-readable dataset]. Minneapolis, MN: University of Minnesota.

Minns, Chris (2000). "Income, Cohort Effects, and Occupational Mobility: A New Look at Immigration to the United States at the Turn of the Twentieth Century." *Explorations in Economic History* 37:4, pp. 326-50.

Mokyr, Joel (1983). *Why Ireland Starved: An Analytical and Quantitative Study of Irish Poverty, 1800–1851*. Boston: George Allen and Unwin.

Mokyr, Joel and Cormac Ó Gráda (1982). "Emigration and Poverty in Prefamine Ireland." *Explorations in Economic History* 19, pp. 360-384.

Moser, Petra, Alessandra Voena, and Fabian Waldinger (2014). "German Jewish Émigrés and US Invention." *American Economic Review* 104:10, pp. 3222-3255.

Nakamura, Emi, Jósef Sigurdsson, and Jón Steinsson (2017). The Gift of Moving: Intergenerational Consequences of a Mobility Shock." NBER Working Paper 22392.

National Archives of Ireland. n.d. History of Irish Census Records. http://www.census.nationalarchives.ie/help/history.html. Accessed July 20, 2018.

OECD/EU (2015). *Indictors of Immigrant Integration 2015: Settling In.* Paris: OECD Publishing.

Ó Gráda, Cormac (1983). "Across the Briny Ocean: Some Thoughts on Irish Emigration to America 1800–1850." In *Ireland and Scotland, 1600–1850: Parallels and Contrasts in Economic and Social Development*. T. M. Devine and David Dickson (ed.). John Donald. Chap. 9, pp. 118–130.

Ó Gráda, Cormac (1999). *Black '47 and Beyond: The Great Irish Famine in History, Economy, and Memory*. Princeton: Princeton University Press.

Ó Gráda, Cormac and Kevin H. O'Rourke (1997). "Migration as Disaster Relief: Lessons from the Great Irish Famine." *European Review of Economic History* 1, pp. 3–25.

Pérez, Santiago (2018). "Railroads and the Rural to Urban Transition: Evidence from 19th Century Argentina." Mimeo., UC Davis.

Roy, A. D. (1951). "Some Thoughts on the Distribution of Earnings." *Oxford Economic Papers* 3:2, pp. 135–146.

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek (2015). *Integrated Public Use Microdata Series: Version 6.0* [Machine-readable database]. Minneapolis: University of Minnesota.

Solar, Peter M. 1989. "The Great Famine Was No Ordinary Subsistence Crisis." In E.M. Crawford (ed.), *Famine: The Irish Experience, 900-1900*, Edinburgh: John Donald, pp. 112-131.

Spitzer, Yannay (2018). "Pogroms, Networks, and Migration: The Jewish Migration from the Russian Empire to the United States 1881–1914." Mimeo. Hebrew University of Jerusalem.

Spitzer, Yannay and Ariell Zimran (2018). "Migrant Self-Selection: Anthropometric Evidence from the Mass Migration of Italians to the United States, 1907–1925." *Journal of Development Economics* 134, pp. 226-247.

Stewart, James I. (2006). "Migration to the Agricultural Frontier and Wealth Accumulation, 1860-1870." *Explorations in Economic History* 43, pp. 547-577.

United Nations High Commissioner for Refugees (2017). *Global Trends: Forced Displacement in 2017*. United Nations High Commissioner for Refugees.

US Department of State, US Department of Homeland Security, and US Department of Health and Human Services (2017). *Proposed Refugee Admissions for Fiscal Year 2018*.

US Immigration Commission (1911). Reports, 61st Congress, 1st Session. Washington, DC: GPO.

Wegge, Simone A. (1998). "Chain Migration and Information Networks: Evidence from Nineteenth-Century Hesse-Cassel." *Journal of Economic History* 58:4, pp. 957-986.

Figures

Figure 1: Selection into linkage
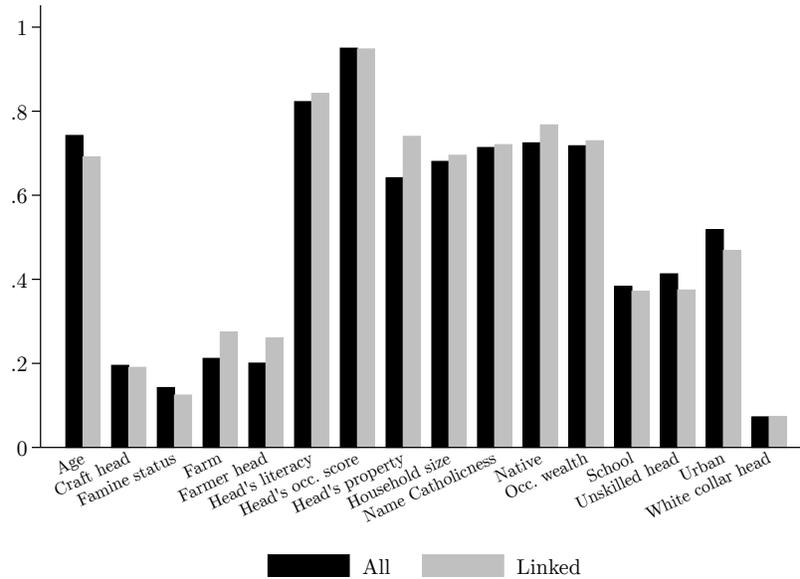
Note: Sample covers sons of 1850 Irish heads of household. The "Linked" group contains 1850 sons who were successfully linked to 1880. The "All" group includes these individuals as well as those whose could not be linked. Values are divided by the smallest multiple of 10 needed to generate a mean below 1.

Figure 2: Selection into classification
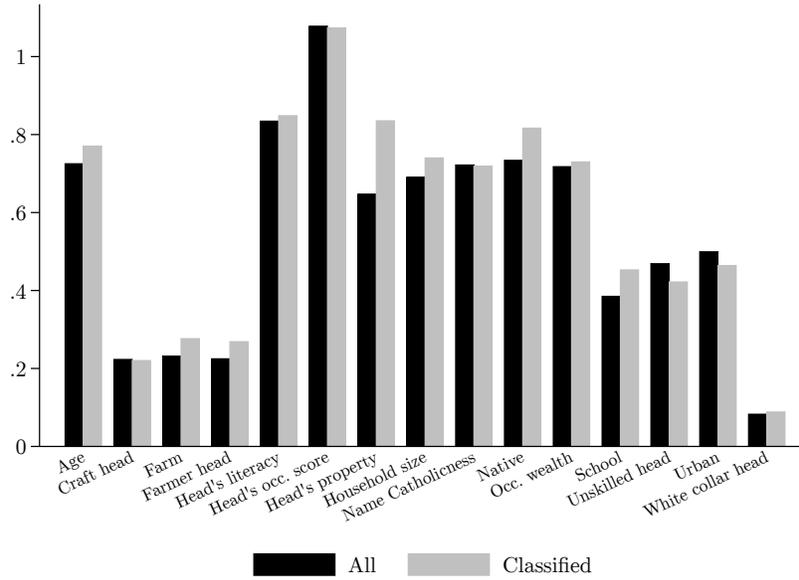
Note: Sample covers sons of 1850 Irish heads of household and is limited to those who could be linked from 1850 to 1880. Observations weighted to correct for selection into linkage. The "Classified" group contains 1850 sons who were linked to 1880 and whose head of household's arrival cohort could be determined. The "All" group includes these individuals as well as those whose head of household's arrival cohort could not be determined. Values are divided by the smallest multiple of 10 needed to generate a mean below 1.

Figure 3: Comparison of adults' and children's outcomes

Figure 3(a): Unskilled Occupation

Heads (1850)   Sons (1880)
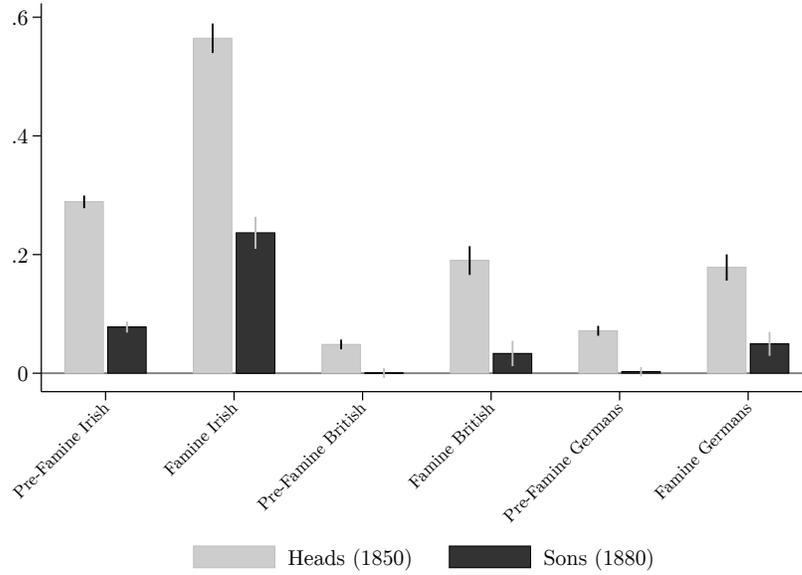
Figure 3(b): Occupational Wealth Score

Heads (1850)   Sons (1880)
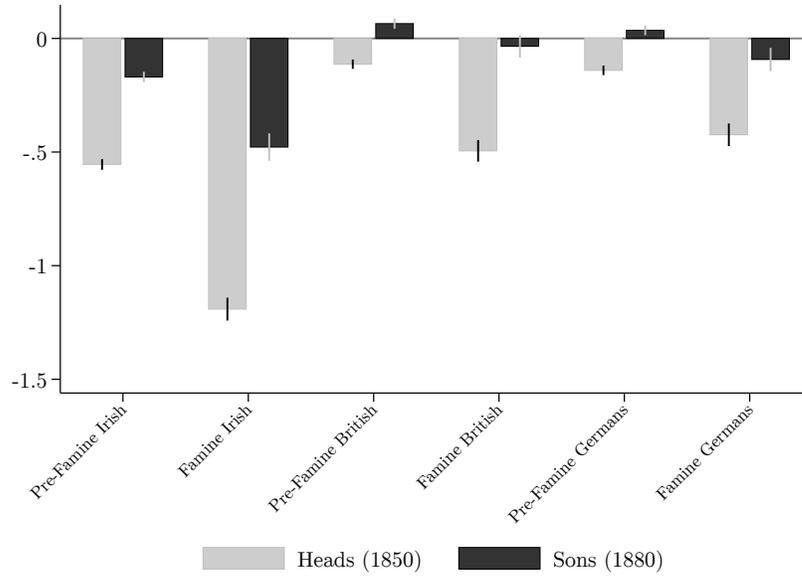
Note: Bars in panel 3(a) for 1850 heads of household are coefficients from column (3) of Table 4. Bars in panel 3(a) for 1880 sons are coefficients from column (1) of Table 6. Bars in panel 3(b) for 1850 heads of household are coefficients from column (5) of Table 4. Bars in panel 3(b) for 1880 sons are coefficients from column (2) of Table 6.

Figure 4: Unconditional and conditional coefficients for 1880 characteristics

Figure 4(a): Unskilled Occupation
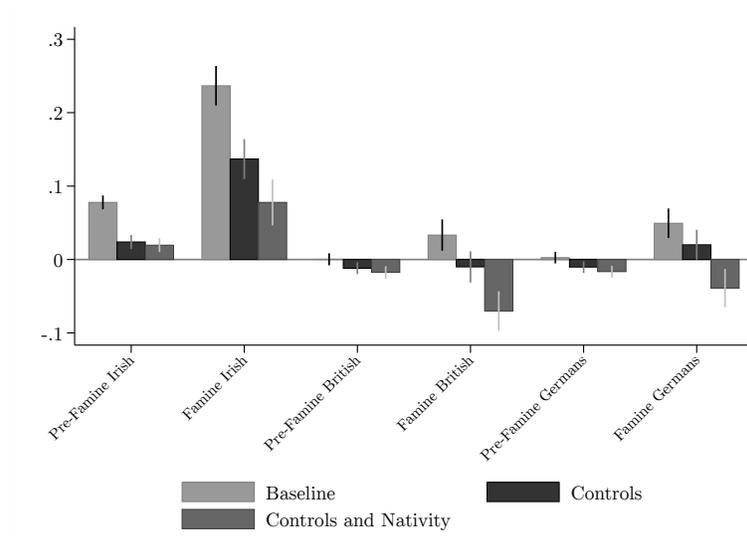
Figure 4(b): Occupational Wealth Score
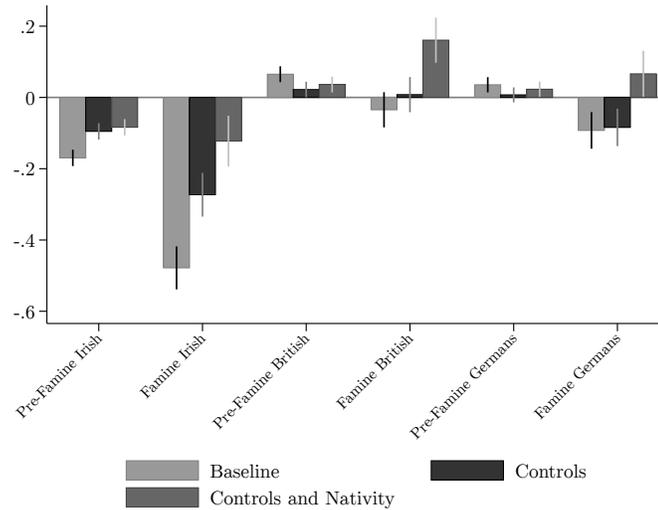
Note: Coefficients portrayed by the "Raw" bars are reached by conditioning only on a quartic in individual's age; they are from columns (1) and (2) of Table 6. Coefficients portrayed by the "Controls" bars come from regressions that control for the 1850 household characteristics, as described in text. They are presented in columns (1) and (2) of Table C.1. Coefficients portrayed by the "Native" bars come from regressions including all 1850 controls and indicators for native birth of sons. They are presented in columns (3) and (4) of Table C.1.

Tables

Table 1: Sample divisions

	Household Heads			Sons		
	(1)	(2)	(3)	(4)	(5)	(6)
Famine Status	Irish	British	Germans	Irish	British	Germans
Not Famine	77,238	47,594	49,638	139,968	84,036	93,063
	(0.371)	(0.435)	(0.307)	(0.561)	(0.640)	(0.524)
Famine	15,009	7,422	11,237	23,116	12,127	17,229
	(0.072)	(0.068)	(0.069)	(0.093)	(0.092)	(0.097)
Maybe Famine	6,268	2,833	7,452	11,508	5,787	14,472
	(0.030)	(0.026)	(0.046)	(0.046)	(0.044)	(0.081)
Unknown	109,846	51,511	93,517	55,865	23,591	46,014
	(0.527)	(0.471)	(0.578)	(0.224)	(0.180)	(0.259)
Total	208,361	109,360	161,844	249,583	131,304	177,739

Notes: This Table enumerates individuals observed in 1850 who were either heads of household over 18 years old, or who were aged 0 to 18 years and were sons of men born in Ireland, Britain, or Germany. Famine status is determined as described in text. Divisions in columns (4)–(6) are according to the place of birth of the head of household. Numbers in parentheses are fractions of the column total.

Table 2: Linkage rates for the seven categories

Famine Status	(1) Start	(2) Searched	(3) Linked
Panel A: Irish			
Pre-Famine	139,968	52,189	10,338
		(0.373)	(0.198)
Famine	23,116	9,665	1,473
		(0.418)	(0.152)
Panel B: British			
Pre-Famine	84,036	35,051	9,182
		(0.417)	(0.262)
Famine	12,127	7,066	1,589
		(0.583)	(0.225)
Panel C: Germans			
Pre-Famine	93,063	57,398	10,071
		(0.617)	(0.175)
Famine	17,229	13,194	1,834
		(0.766)	(0.139)
Panel D: Americans			
All	3,725,019	1,686,500	424,500
		(0.453)	(0.252)

Notes: Panel headers refer to ethnicities (i.e., fathers' birthplaces). Column (1) enumerates all males in 1850 who were aged between 0 and 18 years and who were the children of a head of household born in the United States, Ireland, Britain, or Germany who was either native or whose household composition enabled the determination of arrival cohort. Column (2) enumerates those individuals from column (1) for whom an attempt was made to link to 1880. Column (3) enumerates those individuals from column (2) who were successfully linked to the 1880 census. Numbers in parentheses indicate the fraction of observations from the previous column that were advanced to the next column.

Table 3: Summary statistics for linked children, 1850

	Irish		British		Germans		Americans
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Pre-Famine	Famine	Pre-Famine	Famine	Pre-Famine	Famine	All
Head's human capital							
Literacy	0.858	0.790	0.961	0.946	0.972	0.975	0.903
	(0.349)	(0.407)	(0.195)	(0.227)	(0.165)	(0.157)	(0.296)
Numeracy	0.560	0.494	0.695	0.723	0.716	0.688	0.738
	(0.496)	(0.500)	(0.460)	(0.448)	(0.451)	(0.463)	(0.439)
Head's labor market characteristics							
ln(Occ. Wealth)	7.395	6.719	7.843	7.403	7.810	7.498	7.939
	(1.101)	(0.955)	(0.898)	(0.915)	(0.980)	(1.024)	(0.804)
Occ. Rank	29.974	15.238	39.532	27.878	39.442	31.214	40.486
	(27.872)	(22.131)	(26.577)	(24.204)	(26.715)	(24.836)	(22.235)
White collar	0.095	0.050	0.137	0.077	0.133	0.080	0.079
	(0.293)	(0.217)	(0.344)	(0.266)	(0.340)	(0.271)	(0.270)
Craft	0.227	0.178	0.352	0.419	0.299	0.302	0.185
	(0.419)	(0.383)	(0.477)	(0.494)	(0.458)	(0.459)	(0.388)
Farmer	0.296	0.104	0.371	0.208	0.402	0.340	0.640
	(0.456)	(0.306)	(0.483)	(0.406)	(0.490)	(0.474)	(0.480)
Unskilled	0.382	0.668	0.140	0.297	0.165	0.279	0.096
	(0.486)	(0.471)	(0.347)	(0.457)	(0.371)	(0.449)	(0.295)
Real property (1,000)	0.946	0.158	1.315	0.214	0.991	0.312	1.523
	(4.130)	(2.148)	(4.382)	(0.715)	(4.423)	(1.076)	(7.102)
Son's characteristics							
School (ages 5-15)	0.714	0.518	0.742	0.577	0.647	0.427	0.675
	(0.452)	(0.500)	(0.438)	(0.494)	(0.478)	(0.495)	(0.468)
Native	0.930	0.123	0.915	0.110	0.906	0.126	
	(0.256)	(0.329)	(0.278)	(0.313)	(0.292)	(0.331)	
Irish name Catholicness	0.687	0.754					
	(0.308)	(0.272)					
Famine in utero		0.403					
		(0.491)					
Place of residence characteristics							
Urban	0.451	0.546	0.311	0.373	0.362	0.422	0.106
	(0.498)	(0.498)	(0.463)	(0.484)	(0.480)	(0.494)	(0.308)
Fraction Irish in 100	0.254	0.297	0.117	0.163	0.079	0.094	0.034
	(0.225)	(0.229)	(0.139)	(0.171)	(0.111)	(0.119)	(0.068)
Observations	9,977	1,388	8,939	1,536	9,791	1,742	376,905

Notes: Table includes sons who were under 18 years old in 1850 who could be linked to 1880 and whose head of household's arrival cohort could be determined. Observations grouped by head's nativity and arrival cohort. All figures weighted to correct for selection into linkage on observables. Standard deviations in parentheses. Sample sizes are the minimum with data for all variables other than School, Name Catholicness, Famine in utero, and Fraction Irish in 100.

Table 4: Regressions of 1850 characteristics

Variables	(1) Literacy	(2) Numeracy	(3) Unskill	(4) Farmer	(5) Occ. Wealth	(6) Occ. Rank	(7) Property	(8) School
Pre-Famine Irish	-0.048^a	-0.177^a	0.289^a	-0.345^a	-0.555^a	-10.798^a	-0.627^a	0.043^a
	(0.004)	(0.005)	(0.005)	(0.005)	(0.012)	(0.306)	(0.058)	(0.007)
Famine Irish	-0.113^a	-0.248^a	0.565^a	-0.511^a	-1.191^a	-24.568^a	-1.127^a	-0.141^a
	(0.012)	(0.014)	(0.013)	(0.008)	(0.026)	(0.588)	(0.037)	(0.023)
Pre-Famine British	0.054^a	-0.041^a	0.048^a	-0.274^a	-0.113^a	-1.390^a	-0.307^a	0.067^a
	(0.002)	(0.005)	(0.004)	(0.005)	(0.010)	(0.305)	(0.049)	(0.007)
Famine British	0.044^a	-0.021^c	0.190^a	-0.400^a	-0.495^a	-11.639^a	-0.985^a	-0.089^a
	(0.007)	(0.012)	(0.012)	(0.011)	(0.024)	(0.631)	(0.027)	(0.021)
Pre-Famine Germans	0.065^a	-0.021^a	0.071^a	-0.237^a	-0.141^a	-1.351^a	-0.582^a	-0.022^a
	(0.002)	(0.005)	(0.004)	(0.005)	(0.011)	(0.279)	(0.062)	(0.007)
Famine Germans	0.070^a	-0.052^a	0.178^a	-0.281^a	-0.424^a	-8.899^a	-1.054^a	-0.237^a
	(0.004)	(0.011)	(0.011)	(0.011)	(0.025)	(0.592)	(0.033)	(0.019)
Observations	410,278	410,278	410,278	410,278	410,278	410,278	410,278	202,139
R-squared	0.007	0.008	0.055	0.054	0.039	0.024	0.008	0.025
F-Test	14.145^a	10.828^a	45.845^a	31.561^a	42.866^a	22.584^a	3.898^b	1.871

Significance levels: ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1
Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. F-tests are of the joint null that the difference between the famine Irish and pre-famine Irish coefficients is equal to the difference between the famine and pre-famine coefficients for the other two ethnicities. Excluded group in all specifications is members of native-headed households.

Table 5: Summary statistics for linked children, 1880

Variable	Irish (1) Pre-Famine	Irish (2) Famine	British (3) Pre-Famine	British (4) Famine	Germans (5) Pre-Famine	Germans (6) Famine	Americans (7) All
Labor market characteristics							
ln(Occ. Wealth)	7.628	7.281	7.877	7.726	7.830	7.667	7.803
	(1.045)	(1.090)	(0.992)	(0.943)	(1.022)	(1.062)	(0.993)
Occ. Rank	39.434	31.758	45.876	40.955	44.622	40.481	43.776
	(27.446)	(26.849)	(27.118)	(25.884)	(27.098)	(27.734)	(24.953)
White collar	0.195	0.145	0.224	0.171	0.207	0.196	0.152
	(0.396)	(0.352)	(0.417)	(0.376)	(0.405)	(0.397)	(0.359)
Craft	0.304	0.289	0.264	0.341	0.266	0.330	0.168
	(0.460)	(0.454)	(0.441)	(0.474)	(0.442)	(0.470)	(0.374)
Farmer	0.226	0.137	0.323	0.257	0.330	0.229	0.487
	(0.418)	(0.344)	(0.468)	(0.437)	(0.470)	(0.420)	(0.500)
Unskilled	0.244	0.415	0.162	0.211	0.170	0.227	0.165
	(0.430)	(0.493)	(0.369)	(0.408)	(0.376)	(0.419)	(0.371)
Place of residence							
Moved state	0.364	0.690	0.379	0.634	0.327	0.699	0.339
	(0.481)	(0.462)	(0.485)	(0.482)	(0.469)	(0.459)	(0.474)
Observations	9,654	1,366	8,683	1,505	9,529	1,710	366,144

Notes: Table includes sons who were under 18 years old in 1850 who could be linked to 1880 and whose head of household's arrival cohort could be determined. Observations grouped by head's nativity and arrival cohort. All figures weighted to correct for selection into linkage on observables. Standard deviations in parentheses. Sample sizes are the minimum with data for all variables.

Table 6: Regressions of 1880 characteristics

Variables	(1) Unskill	(2) Occ. Wealth	(3) Occ. Rank
Pre-Famine Irish	0.078^a	-0.170^a	-4.239^a
	(0.005)	(0.012)	(0.307)
Famine Irish	0.237^a	-0.478^a	-11.075^a
	(0.014)	(0.031)	(0.756)
Pre-Famine British	0.000	0.065^a	1.891^a
	(0.004)	(0.011)	(0.309)
Famine British	0.033^a	-0.035	-1.893^a
	(0.011)	(0.025)	(0.689)
Pre-Famine Germans	0.003	0.036^a	1.022^a
	(0.004)	(0.011)	(0.287)
Famine Germans	0.049^a	-0.093^a	-2.356^a
	(0.010)	(0.026)	(0.685)
Observations	410,278	398,591	398,591
R-squared	0.013	0.017	0.014
F-Test	26.015^a	13.252^a	5.729^a

Significance levels: ^{a} $p<0.01$, ^{b} $p<0.05$, ^{c} $p<0.1$

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and are weighted by inverse linkage probability. F-tests are of the joint null that the difference between the famine Irish and pre-famine Irish coefficients is equal to the difference between the famine and pre-famine coefficients for the other two ethnicities. Excluded group in all specifications is members of native-headed households.

Table 7: Correlates of upward mobility among the famine Irish

	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Variables				
Head Literacy (1850)	−0.074	0.064	−0.072	0.064
	(0.045)	(0.098)	(0.045)	(0.098)
Native	−0.075	0.327^b	−0.058	0.326^b
	(0.063)	(0.135)	(0.064)	(0.135)
Famine in utero	0.063	0.000	0.052	0.001
	(0.049)	(0.113)	(0.049)	(0.114)
Irish Name Catholicness	0.166^a	-0.294^b	0.163^a	-0.294^b
	(0.051)	(0.117)	(0.051)	(0.117)
Fraction Irish in 100	0.045	−0.128	0.048	−0.128
	(0.074)	(0.167)	(0.074)	(0.168)
Moved State			0.065^c	−0.003
			(0.036)	(0.080)
Observations	968	951	968	951
R-squared	0.046	0.032	0.049	0.032

Significance levels: ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1

Notes: Robust standard errors in parentheses. Sample limited to famine Irish children. All specifications include a quartic in age of the individual and the 1850 household head and are weighted by inverse linkage probability. Columns with an indicator for holding an unskilled occuptation as the dependent variable condition on indicators of the father's occupational status. Columns with the log of the occupational wealth score as the outcomes condition on the log of the father's occupational wealth score.

Data Appendix (not for publication)

Appendix A: Census linkage

The linkage procedure was the following.

1. Men with non-empty first and last names were extracted from the 1850 census, and punctuation was removed.

2. The listed first name was divided into a given name and middle initial, when one was present.

3. Standard first name abbreviations were replaced (e.g., "Wm" was replaced with "William").

4. Any remaining spaces were removed from the names.

5. The list of cleaned men's names generated by steps 1-4 was linked to itself on the following characteristics:

   a. Birthplace matches

   b. Absolute difference in birth years less than or equal to 4

   c. First three characters of last name soundex match, and either of the following two conditions are true:[48]

      i. The last name soundexes are identical and the SAS spelling distance is less than or equal to 20.[49]

      ii. The last name soundexes are non-identical, and the SAS spelling distance is less than or equal to 17.

   d. First letter of first name matches, and spelling distance between first names is less than or equal to 20.

6. Any individual with a match in step 5 (other than himself) was removed from the sample.

7. The remaining men from 1850 and all those from 1880 were cleaned according to steps 1, 2, and 4 above (no standardization of name abbreviations was made).

8. Men in 1850 were linked to 1880 according to the criteria listed in step 5, as well as the

---

[48] This is not a soundex match; the soundex is used to limit the set of candidate matches, which are then evaluated based on spelling distance.

[49] The SAS spelling distance is not a symmetric measure. Whenever I use spelling distance, I make a match when $\min\{\text{spedis}_{a,b}, \text{spedis}_{b,a}\} \leq c$, where c is the linkage cutoff.

following additional characteristic:

    a. Where both records report a middle initial, the middle initials must match for a match to occur.

9. Of the candidate matches created in step 8, any match in which more than one 1880 individual matched to an 1850 individual, or in which more than one 1850 individual matched to the same 1880 individual, was removed.

Table A.1 presents the rates of successful linkage by the individual's place of birth (so that native-born sons of immigrants are in column 4). This Table is not restricted to sons; it also includes others, such as heads of household and boarders. Tables A.2 presents the results of a linear probability model relating the probability of being successfully linked between 1850 and 1880 to a variety of characteristics observed in 1850. This Table divides the sample by the ethnicity of the head of household in 1850.

Table A.1: Linkage rates by place of birth

	Birth Place			
	(1)	(2)	(3)	(4)
Stage	Irish	British	German	American
Start	998,625	391,615	602,955	17,667,555
Males	520,096	224,874	345,581	8,963,889
Searched	164,710	108,399	222,648	4,286,869
Linked	18,378	18,684	24,822	892,443

    Notes: All observations enumerated are from the 1850 census. Table is not limited to individuals who were sons in 1850. Observations are divided by birth place and not ethnicity. Individuals enumerated in the row labeled Searched are those for whom an atempt was made to make a match to the 1880 census. Individuals enumerated in the row labeled Linked are those for whom a unique match could be made in the 1880 census.

Table A.2: Selection into linkage, children

Variables	(1) Irish	(2) British	(3) Germans	(4) Americans
HH Size	0.002^a	0.002^a	0.003^a	0.001^a
	(0.000)	(0.000)	(0.000)	(0.000)
Age	-0.001^a	-0.002^a	-0.001^a	-0.001^a
	(0.000)	(0.000)	(0.000)	(0.000)
Urban	-0.006^a	-0.004	-0.005^b	-0.017^a
	(0.001)	(0.002)	(0.002)	(0.001)
Farm	0.019^b	0.008	0.010	0.014^a
	(0.007)	(0.008)	(0.010)	(0.001)
School	-0.000	0.004^c	0.008^a	0.005^a
	(0.001)	(0.002)	(0.002)	(0.000)
Head's Literacy	0.006^a	0.013^a	0.004	0.007^a
	(0.001)	(0.004)	(0.004)	(0.001)
Head's Occ Wealth	-0.001	0.002	0.008^a	0.003^a
	(0.001)	(0.002)	(0.002)	(0.001)
Head's Property	0.000	0.000	0.000	0.000^a
	(0.000)	(0.000)	(0.000)	(0.000)
Native	0.007^a	-0.023^a	-0.004^b	
	(0.001)	(0.002)	(0.002)	
Head WC	0.008	-0.008	0.022^b	-0.004^a
	(0.008)	(0.009)	(0.010)	(0.001)
Head Unskill	-0.003	-0.014	0.017	-0.006^a
	(0.008)	(0.009)	(0.011)	(0.002)
Head Craft	0.004	-0.005	0.020^b	-0.001
	(0.008)	(0.009)	(0.010)	(0.001)
Observations	214,935	118,624	163,421	3,382,996
R-squared	0.010	0.010	0.003	0.011

Significance levels: ^{a} p<0.01, ^{b} p<0.05, ^{c} p<0.1

Notes: These are regressions of an indicator for successful linkage on a variety of 1850 characteristics. Sample includes all males who were under age 18 in 1850 and were sons of a head of household, including those who were omitted for not having unique names. Columns are divided by place of birth of the head of household. All specifications include state FE. Robust standard errors in parentheses.

Appendix B: Surname-Based Catholicism Imputation Algorithm

The algorithm to determine the probability of being Catholic based on surname proceeds as follows. It is based on Spitzer and Zimran's (2018) algorithm to determine the province of origin of Italian migrants to the United States by their surnames.

1. We obtained the complete 1901 Census of Ireland. Religions listed in this source were manually classified as Catholic, Protestant, or other. Using this source, a list was created of all distinct surnames, along with the number of these individuals who were and were not Catholic. For example, there were 12,238 individuals in 1901 with the surname Collins who were Catholic, and 1,076 individuals in 1901 with the surname Collins who were not Catholic.

2. A list was created of all distinct surnames of individuals in the 1850 US Census who were either born in Ireland or lived in a household headed by an Irish-born individual.

3. These two lists were matched to one another if the first letter matched, and on the following criteria.

    a. If the soundexes of the entry in the two sources were the same, a match was made if the minimum SPEDIS value between them was less than or equal to 20.

    b. If the soundexes of the entries in the two sources were not the same, a match was made if the minimum SPEDIS value between them was less than or equal to 17.

    This approach created, for each surname in the 1850 census, a count of Catholic and non-Catholic individuals to whom the individual was matched in the 1901 Irish census.

4. The surname "Catholicness" was calculated as

$$\frac{\sum_{n \in N} C_n}{\sum_{n \in N} P_n}$$

    where *N* is the set of 1901 surnames to which an 1850 individual's surname was matched, C_n is the number of Catholic individuals with surname *n* in 1901 Ireland, and P_n is the number of non-Catholic individuals with the surname *n* in 1901 Ireland. In cases where $C_n + P_n < 200$, no categorization was made.

Appendix C: Additional results

Table C.1: Coefficients underlying Figure 4

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Pre-Famine Irish	0.024^a	-0.095^a	0.020^a	-0.084^a
	(0.005)	(0.012)	(0.005)	(0.012)
Famine Irish	0.137^a	-0.273^a	0.078^a	-0.122^a
	(0.014)	(0.031)	(0.016)	(0.036)
Pre-Famine British	-0.012^a	0.022^b	-0.017^a	0.036^a
	(0.004)	(0.011)	(0.004)	(0.011)
Famine British	-0.010	0.008	-0.070^a	0.161^a
	(0.011)	(0.025)	(0.014)	(0.032)
Pre-Famine Germans	-0.010^b	0.007	-0.017^a	0.023^b
	(0.004)	(0.011)	(0.004)	(0.011)
Famine Germans	0.020^c	-0.084^a	-0.039^a	0.066^b
	(0.010)	(0.027)	(0.013)	(0.033)
Native			-0.068^a	0.173^a
			(0.009)	(0.022)
Observations	410,278	398,591	410,278	398,591
R-squared	0.047	0.091	0.048	0.092
F-Test	18.549^a	7.389^a	18.469^a	7.388^a

Significance levels: ^{a} p$<$0.01, ^{b} p$<$0.05, ^{c} p$<$0.1

Notes: Robust standard errors in parentheses. All specifications include a quartic in age and the controls described in text, and are weighted by inverse linkage probability. F-tests are of the joint null that the difference between the famine Irish and the pre-famine Irish coefficients is equal to the difference between the famine and pre-famine coefficients for the other two ethnicities. Excluded group in all specifications is members of native-headed households.