The Economic Assimilation of Irish Famine Migrants to the United States

William J. Collins and Ariell Zimran

September 2019

The repeated failure of Ireland's potato crop in the late 1840s led to a major famine and sparked a surge in migration to the US. We build a new 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 in 1880 than their fathers did in 1850. 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 children of the famine Irish indicates that having a more Catholic surname and being born in Ireland were associated with less upward mobility. Our results contribute to literatures on immigrant assimilation, refugee migration, and the Age of Mass Migration.

JEL: F22, J61, J62, N31, O15
Keywords: Migration, Refugees, Assimilation, Intergenerational Mobility, Irish Famine

Contact: william.collins@vanderbilt.edu; ariell.zimran@vanderbilt.edu

For helpful comments, the authors thank Ran Abramitzky (the editor), anonymous reviewers, Zachary Barnett-Howell, Joseph Ferrie, Tim Hatton, Nick Holtkamp, Joel Mokyr, Cormac Ó Gráda, Marianne Wanamaker; seminar participants at Auburn University, UCLA, the University of Minnesota, Vanderbilt University, and William & Mary; and conference participants at the 2019 ASSA Meeting, the 2019 H2D2 Research Day at the University of Michigan, the 2019 Economic History Society Conference, the 2019 Midwest International Economic Development Conference, the 2019 IZA Annual Migration Meeting, and the 2019 Migration and Development Conference. Claire Whittaker provided excellent research assistance. Computations were performed on the Advanced Computing Cluster for Research and Education at Vanderbilt University. Collins is the Terence E. Adderley Jr. Professor of Economics at Vanderbilt University and a Research Associate of the NBER. Zimran is an Assistant Professor of Economics at Vanderbilt University and a Faculty Research Fellow of the NBER.

1. Introduction

Ireland's Great Famine in the late 1840s marked a turning point in the country's demographic and economic history. From a population that numbered just over eight million in 1841, it is estimated that the famine caused the death of about one million Irish and drove another million to emigrate by the early 1850s (Ó Gráda 1999, ch. 3). Most of the Irish emigrants settled in the United States, where virtually open borders gave sanctuary from the imminent threat of starvation and disease (Ó Gráda 2019). The US absorbed over 500,000 new arrivals from Ireland between the famine's onset in 1846 and 1850 (Ferenczi and Wilcox 1929; Barde, Carter, and Sutch 2006). This paper picks up the Irish migrants' story on American shores. We build new datasets to study the economic status and labor market assimilation of the famine-era Irish immigrants, with a particular focus on the adult labor market outcomes of the immigrants' children.

Irish migration to the US 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), but the sharp increase in arrivals during the famine dwarfed previous arrival cohorts. This large wave of arrivals marked the start of the "Age of Mass Migration," and likely comprised the largest group of refugees from a single source that the US has ever absorbed relative to the size of its population.[1] The increased volume of Irish immigration coincided with an apparent change in the migrants' characteristics in comparison to earlier, more prosperous arrivals (Handlin 1991 [1941] p. 51; Miller 1985 p. 295; Anbinder 1992, p. 7). This intensified the view of many Americans at the time that the Irish migrants' relative poverty, tendency to live near one another, and predominantly Catholic religion were barriers to their assimilation and, therefore, justified restrictions on immigration and immigrants' rights (Higham 1955; Anbinder 1992; Hirota 2017). Such concerns have resurfaced throughout US history in response to the arrival of poor or culturally "different" immigrants. In this case, the nativist response was severe, including the political ascent of the "Know Nothings" (Anbinder 1992; Alsan, Eriksson, and Niemesh 2019), but US policymakers did not curtail European immigration until the 1920s, bringing an end to the Age of Mass Migration.

Although the size and historical prominence of the famine-era cohort of Irish immigrants makes them a particularly interesting group to study, data constraints have made it difficult to do so. 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 censuses. The

[1] For perspective, refugees and asylees granted permanent resident status under the Displaced Persons Act (1948) averaged about 118,000 per year from 1950-52, at the program's peak (Barde, Carter, and Sutch 2006, p. 1-632); the Mariel Boatlift entailed approximately 125,000 Cuban migrants in 1980 (Card 1990, p. 245).

linked dataset is large and national in scope. It includes household heads and their sons, enabling 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 individual-level 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 tend to confound inferences from cross-sectional census data (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 inclusion of German and British immigrants, who comprised the next largest groups of immigrants in this period, allows for useful comparisons across arrival cohorts and across immigrant groups, which in turn helps to illustrate the distinctiveness of the famine-era Irish.

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 economics literature on migration (e.g., Borjas 1992, 1993; Card, DiNardo, and Estes 2000; Card 2005; Caponi 2011; Abramitzky, Boustan, and Eriksson 2014; Alexander and Ward 2018; Abramitzky et al. 2019b). 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, at the start of the Age of Mass Migration. 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).[2]

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 Irish famine's onset and, in general, to any study of immigrant assimilation in this early period.[3] Nonetheless, in 1850, we can determine the arrival cohort for many household heads by examining their children's birth-year and country-of-birth information.[4] This approach definitively categorizes the household head's arrival cohort ("pre-famine" or "famine-era") for over two-thirds of the sons of

[2] We use the term "refugee" because many of the Irish were permanently driven from home by an ecological disaster and ensuing economic and social dislocation. The modern literature focuses on refugees that meet a specific legal definition established in the postwar period; see Hatton (2012) for a brief review.
[3] 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 16).
[4] For example, a family with an Irish-born child aged 2 in 1850 immigrated to the US some time in or after 1848, which would make them famine-era migrants. A family with an American-born child aged 7 in 1850 immigrated to the US some time before or during 1843, making them pre-famine migrants.

Irish immigrants. 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. In addition, we show that the "classified" set of children is fairly representative of all immigrants' children circa 1850 in terms of observable characteristics and that the omission of "unclassified" children from the baseline analysis is unlikely to confound our conclusions.

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, concurrent migrants from Germany and Britain, and US natives? In particular, is there evidence of a differential 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 influenced the average household characteristics and economic resources of immigrants' children, though documenting the patterns of migrant selection at the time of the famine is also of independent interest (e.g., Mokyr and Ó Gráda 1982; Cohn 1995). 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-famine and famine-era arrivals.[5] 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 US 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 Irish children's advancement. We find that, on average, the sons of the famine-era Irish immigrants fared poorly in the labor market as

[5] This contribution is qualified in that our method to classify individuals' year of arrival requires that they have children; this group may have differed from childless immigrants. We discuss this issue further in footnote 37.

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 US natives' sons. The gap is even smaller when considering their adverse childhood environment (i.e., when 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 generation.

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, geographic mobility, and social distance (specifically, having a surname that we determine to be predominantly Catholic). 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 study of the economics and history of immigration in several dimensions. Most directly, the paper advances our knowledge of the Age of Mass Migration, especially of Irish migration during the famine. It is thus complementary to research by Ó Gráda and O'Rourke (1997), Hatton and Williamson (1998), Ferrie (1999), Cohn (2009), Abramitzky, Boustan, and Eriksson (2012, 2014), Alexander and Ward (2018), and Ó Gráda (2019), among others. By studying immigrants' experiences in the early portion of the period, in contrast to recent scholarship that has focused on the post-1900 period, the paper brings the first cohorts of the Age of Mass Migration into sharper focus. More broadly, studying the experience of Irish famine-era migrants may yield insights into the economics of large-scale migration due to natural or man-made disasters.[6] By studying a large wave of migrants and their offspring long before the implementation of restrictive immigration policies, the paper yields insight into the migration and assimilation processes in a setting where economic forces predominated. Finally, our emphasis on economic assimilation in the long run speaks to a core theme of the international migration literature, both historical and

[6] On migration due to environmental disasters or war see recent work by Boustan, Kahn, and Rhode (2012), Hornbeck (2012), Moser, Voena, and Waldinger (2014), Nakamura, Sigurdsson, and Steinsson (2017), Blum and Rei (2018), Long and Siu (2018), and Mahajan and Yang (2019).

contemporary.[7] Issues of migrant assimilation have become particularly salient in recent years, as large numbers of refugees have sought residence in the United States and Europe (United Nations High Commissioner for Refugees 2017). This has fueled a debate in which concerns about immigrants' assimilation have been cited as justification for more restrictive policies (e.g., Kelly 2018), such as 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). In the Irish case, we find that despite a deterioration in migrant selection, poor labor market outcomes soon after arrival, and a political backlash, the migrants' children converged strongly, albeit incompletely, on natives' outcomes by 1880. In this sense, they showed clear evidence of "assimilation and contribution" to the American economy.

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. By late 1846, Ireland was in the grip of a historic famine. The relatively good yields of 1847 were of little avail, as farmers had shifted away from planting potatoes. The famine was then exacerbated by crop failures in 1848 and, less severely, in 1849 and 1850.[8]

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

[7] 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), Alexander and Ward (2018), and Abramitzky et al (2019b).

[8] 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.[9]

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.[10] By June 1850, when the census was taken, there were nearly one million Irish-born residents in the US. To be sure, some of those who migrated during the famine would have left Ireland 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 made a leap of faith in a desperate context characterized by widespread death and social collapse. Like many modern-day refugees, their departure was precipitated by an imminent threat to survival, often undertaken hastily, and generally irreversible. A major difference, of course, is that migrants fleeing to the US during the famine did not face the policy restrictions on entry that are prevalent today.

There are indications that the famine 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 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), where residents were more agricultural, more Catholic, and less literate (Commissioners 1843; Geary 1996). 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.[11] Cohn (1995) finds a slight decrease in the fraction of Irish immigrants who were reported as laborers in passenger lists after 1845, whereas Mokyr and Ó Gráda (1982) find an increase in innumeracy as implied by age heaping.

---

[9] These are our calculations from the 1850 complete count census (Minnesota Population Center 2017; Ruggles et al. 2015). We caution that comparing across the censuses is somewhat tenuous as 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 (2019) for evidence of negative selection in the early twentieth century.

[10] 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.

[11] 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.

In the US, Irish immigrants were over-represented in urban areas of the Northeast. It is worth emphasizing, however, that most, 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.[12] Landmarks in the historical literature tend to focus on these cities (Handlin 1991 [1941], Ernst 1949), but a broader geographic scope is required to describe the experience of Irish immigrants. Scholars have long understood that the Irish 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 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 rapid asset accumulation in the bank account records of Irish immigrants 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 by the sheer volume of mass migration. Hirota (2017, p. 2) writes, "The newcomers' religion triggered an outburst of anti-Irish nativism in these states [New York and Massachusetts], but so too did the immigrants' poverty." Hirota (2017, ch. 2) also documents cases of deportation and efforts to prevent destitute Irish passengers from disembarking in Massachusetts. 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, our understanding of the famine migrants' experience is limited. 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. 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

[12] These are our calculations from the complete count census data (Minnesota Population Center 2017; Ruggles et al. 2015).

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 connects 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 see 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

Discerning arrival cohorts of immigrants in the 1850 census

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 others prior to 1890 did not inquire about immigrants' year of arrival, our first task was to develop an approach to distinguish the Irish who arrived during the famine from those who arrived earlier. We did so for each Irish-born male head of household based on the ages and places of birth of all children in the household in the 1850 census.[13] Household heads with a child (son or daughter) born in the US in 1845 or earlier are categorized in the pre-famine group, whereas those with a child born in Ireland in 1846 or later are in the famine-era group. We performed a similar arrival-cohort classification for British- and German-born heads of household in 1850.[14] After the Irish, the Germans and British were the largest groups of immigrants before 1850 (Barde, Carter, and Sutch 2006, p. 1-560). The 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 included households headed by US-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." Not all children in Irish, British, or German-headed households were foreign-born. It follows that not all children who are classified as

[13] We use an indicator for relationship to the head of household provided 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. The census did not record relationship to the head of household until 1880. We focus on male heads due to the lack of coded occupations for females in the 1850 census data.
[14] Collins and Zimran (2019) provide code to replicate this classification and all subsequent analyses.

being "famine-era Irish" would have experienced the famine in Ireland (i.e., if they were born in the US after their family's arrival).

Table 1 divides each immigrant group using the categorization described above, reporting for male 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 71 percent for the Irish sons, but only 44 percent for Irish heads of household (including those without children). 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.

The remaining 29 percent of Irish children lived in households where the children's birthplaces and ages do not clearly indicate the head's arrival cohort. For instance, 5 percent of Irish children (the "Maybe Famine" group in Table 1) were members of a household with a child born in the US in 1846 or later and a child born in Ireland between 1841 and 1845. Although it seems likely that these households migrated during the famine, it is impossible to be certain (e.g., an Irish-born child in 1842 and a US-born child in 1848 would be consistent with migration before or after the famine's onset). Another 24 percent of Irish children (the "Unknown" group in Table 1) were in households where even less insight on migration timing can be gleaned from the children's birthplaces and ages, such as those with only US-born children born in 1846 or later, or only Irish-born children born in 1845 or earlier.

A natural concern is that the 71 percent of Irish children whose father's arrival cohort is definitely classified might not be representative of the population of all immigrant children in 1850, biasing our characterization of assimilation. To consider whether selection into definite arrival-cohort classification is likely to bias our results, Figure 1 compares the means of 1850 observables for the classified Irish children to those of the whole sample of Irish children with male heads of household. Though selection into classification appears 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.[15] This is consistent with the classified children being fairly representative of the population of immigrants' children circa 1850.

[15] Figure 1 presents the linked sample and is weighted to correct for selection into linkage. Similar results (Online Appendix C) are obtained for children of British and German immigrants. Results are qualitatively unchanged if we re-weight to correct for selection into both linkage and classification. However, we prefer not to re-weight for selection into classification because it is unclear whether differences in observables between the classified and unclassified groups reflect selection into classification (which one would like to reverse) or actual differences in the probability that classified and unclassified individuals arrived during the famine.

For the omission of the unclassified children to severely bias our characterization of assimilation, the unclassified children would have to have markedly different 1880 outcomes than classified children despite being similar on 1850 household observables. We examine this possibility in Online Appendix C, where we bring the unclassified children into the analysis, either as a separate group or by randomly assigning them to either pre-famine or famine-era cohorts.[16] The baseline results are robust.

One might also be concerned that arrival-cohort classification is confounded by incorrect reporting of children's ages. We expect age heaping to be less severe for children than for adults due to the greater developmental differences by age among children. To provide direct evidence, we randomly selected 50 classified households headed by Irish men from our linked data, stratifying to include 25 of each categorization group (famine and pre-famine) and searched carefully for their passenger list records using Ancestry.com. Although we were able to locate only 10 of these households definitively, the age- and birthplace-based algorithm correctly categorized all of them.[17]

Linking census records

The next, but separate, step in building the dataset entailed linking males who were born in Ireland, Britain, Germany, and the US from the 1850 to the 1880 full count US census records (Ruggles et al. 2015; Minnesota Population Center 2017). We linked the records without regard for whether arrival cohort could be determined, but in the subsequent analyses, we focus on males with linked records and a clear arrival-cohort designation for the household head.

We first retained individuals from the 1850 census records who were unique in their combination of birthplace, age-implied birth year (plus or minus four years), and first and last names with some allowance for spelling differences (e.g., Patrick Kavanagh and Patrick Kavanaugh would not be considered unique names). This step drops from further consideration individuals who are

[16] First, we repeated the main results including the unclassified as a separate group in the analysis; the unclassified Irish children also display strong intergenerational assimilation. Second, we randomly assigned unclassified children to the pre- or famine-era cohorts, repeating the assignment 500 times, and analyzed the range of resulting point estimates. We repeated this for several possible probabilities that an unclassified child's head arrived during the famine. In all cases, the results were qualitatively unchanged.

[17] This exercise also illustrates the difficulties of linking from the passenger lists to the census. 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; in some cases, individuals could be located only when searching for their household members. The use of household information in linkage is generally avoided due to the bias it introduces. The census-based approach to determining the arrival year enables us to avoid these problems of linkage from the passenger lists and to maintain a larger sample.

"too similar" to others in 1850 because such close proximity could lead to mismatching in 1880's records. We linked the remaining 1850 records forward to the 1880 census records according to the same set of identifying characteristics (birthplace [exactly], birth year [+/- 4 years], and names, including middle initials when they were listed in both records). To do so, we used the soundex algorithm to narrow the set of possible name matches within birthplace and birth-year bands; then, we used the SAS function SPEDIS to evaluate candidate matches based on spelling distance between names.[18] This step might yield one or several candidate matches in 1880, or none at all. We kept links in which both the 1850 and the 1880 record had a single, unique match in the opposite census. The precise linkage procedure is provided in Appendix A.

This linkage method, like many others in the literature, descends from Ferrie's (1996) pioneering effort to link records from a public use sample drawn from 1850's census to the 1860 census manuscripts according to fixed, pre-determined rules. Relative to Ferrie (1996), our approach (and those of several other recent studies) incorporates methodological advances made possible by new data sources and more computational power, such as using complete-count digitized census records in both years, measuring spelling distance between names to evaluate potential matches, and eliminating non-unique men before attempting to make a link. Our approach most closely resembles the method used by Ferrie and his coauthors in Beach et al. (2016), as well as variants of this method used in Aizer et al. (2016) and Long and Ferrie (2018). For instance, Beach et al. (2016) also uses the soundex algorithm to narrow the set of possible name matches and SPEDIS to evaluate candidate matches based on spelling distance; however, our method differs in the precise rules for making a match and in the elimination of non-unique men in the base year dataset. Our method also has some similarities to Abramitzy, Boustan, and Eriksson's (2014), as it too descends from Ferrie's original work.[19]

We followed this approach because of the long-standing use and acceptance of linkage methods that derive from Ferrie (1996); we discuss below how similar our results are when using

---

[18] We require the first three characters of the soundex of the last name to match. After this restriction, linking is based on orthographic distance using the SAS SPEDIS function. Details on the SPEDIS function are available from Gershteyn (2000). SPEDIS is similar to Jaro-Winkler distance but assigns different penalties for different operations needed to convert one string into another. For example, deleting or adding one of a double letter is considered a smaller change than adding a letter to the end of a string, which is in turn considered a smaller change than adding a letter to the middle of a string. The code to carry out the linkage procedure is provided by Collins and Zimran (2019).

[19] Our approach is closest to their method that uses only unique names within a five-year age band (described on p. 484, fn 27). We use a wider age band for name uniqueness and a different approach to comparing names to select matches. See Abramitzky et al. (2019a) for a detailed discussion of that method and variations on it.

different methods. Our approach has features that should result in relatively conservative matches and that correspond to suggestions in the literature that studies match quality (Abramitzky et al. 2019a; Bailey et al. 2019). In particular, the elimination of individuals who are not unique in a particular band, the use of orthographic distance measures, and the use of middle names or initials when they are available may tend to reduce false matches.

Table 2 presents rates of successful linkage for classified individuals.[20] Column (1) reports the total number of males, age 0 to 18, who were sons of male household heads born in Ireland, Britain, Germany, or the US in the 1850 complete count census data. Column (2) shows the number and fraction (in parentheses) that remained after removing those who were not unique on name, birthplace, and birth year. The number and fraction of men who were successfully located in 1880 are shown in column (3), with the fraction linked relative to column (1) shown in square brackets. 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. The match rates for the British and Americans are comparable to those reported by Ferrie (1996, p. 145 and Table 4) for unique individuals.[21] The match rates for the Irish and Germans are somewhat lower, and there are reasons to expect this. The arrival of large cohorts of Irish and German immigrants *after* 1850 would tend to confound links even of individuals who were unique in 1850.[22] The lower level of literacy for the Irish might also tend to reduce match rates.

Expressed relative to the full population (column (1)), the linkage rates range from 6.4 percent for the famine-era Irish children to 13.1 percent for the famine-era British children. As shown in Online Appendix D, these rates are similar to those obtained when using alternative linkage methods based on Abramitzky, Boustan, and Eriksson (2014) and described in detail by Abramitzky et al. (2019a).[23] They are somewhat lower than the rates achieved by other similar linkage methods (e.g., Beach et al. 2016 have a match rate of 15 percent among native-born children in the early 20th century), but as discussed above, this is to be expected due to ongoing mass migration in the study period.

---

[20] Appendix Table A.1 presents rates of successful linkage for the full sample, dividing by own birthplace rather than the ethnicity of the head of household in 1850.
[21] Note, however, that Ferrie (1996) defines uniqueness relative to the target census, not the base census.
[22] For instance, a Patrick Kavanagh born in 1845 in Ireland may have been unique in the US in 1850, but other Patrick Kavanaghs born in 1845 in Ireland may have arrived after 1850, preventing successful linkage.
[23] Using code from Abramitzky et al. (2019a), we linked 1850 to 1880 using ABE-NYSIIS with 5-year band uniqueness (without middle names) and the ABE-Exact with 5-year band uniqueness (without middle names).

Selection of men into the linked dataset is likely to be non-random. Appendix A reports linear probability models that relate the probability of being linked to a variety of 1850 characteristics. There are several statistically significant predictors of selection into linkage, such as property ownership and literacy; this is typical in studies that link US census data (e.g., Abramitzky, Boustan, and Eriksson 2014; Beach et al. 2016). However, most differences are small, as shown in Figure 2, 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.[24]

It is straightforward to test whether our results are sensitive to changes in the record linkage method. First, we can restrict the baseline dataset to the subset of our matches that, in addition to satisfying our matching criteria, are also exact matches between the two censuses.[25] Doing so may help to alleviate concerns over the potential of bias from false matches (Abramitzky et al. 2019a; Bailey et al. 2019). When we do so, the results are similar, as reported in Online Appendix E. We can also examine alternative datasets obtained by applying two linkage methods based on Abramitzky, Boustan, and Eriksson (2014) and described by Abramitzky et al. (2019a). This serves as another substantial check on the robustness of our baseline results. As shown in Online Appendix D, we find that our results (and linkage rates, as mentioned above) are similar across the three differently linked datasets.

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," "farm 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 or South) cell, we calculate a "wealth score" that equals the average total wealth of men (ages 30-65) in the one-percent sample of the 1870 census

[24] This probability is computed from separate probit regressions of a linkage indicator on 1850 covariates for each of the seven ethnicity-cohort groups.

[25] We define an exact match as one in which there is no more than a one-year difference in age-implied birth year 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").

(Ruggles et al. 2015; Minnesota Population Center 2017).[26] 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.[27] We believe the 1870-based occupation score is better suited for studying the mid-nineteenth century because it is closer in time to our period of study, and it is designed to allow for differences between the North and South. We also use the wealth score to define an occupational rank variable, which represents the percentile of the wealth score in the sample.[28] In Online Appendix F, we verify that our results are robust to a variety of definitions of the occupational wealth score, including the use of medians rather than means of wealth in a given occupation-by-region cell.

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 insight on the association between religion and economic outcomes for the Irish, we construct a measure of "Catholicity" based on surnames.[29] 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.[30] The details of the procedure are provided in Online Appendix G. This adds a useful dimension to the description of Irish immigrants and within-Irish heterogeneity in labor market outcomes.

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

---

[26] The 1850 census did not ascertain personal wealth, and the 1880 census did not inquire about wealth at all, hence our reliance on 1870.

[27] The correlation between our measure in 1880 and the 1950-based occupational income score 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.

[28] 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.

[29] The use of information contained in surnames to glean additional information on immigrant origins is increasingly common (e.g., Spitzer 2018; Spitzer and Zimran 2018), including in the study of the origins of the Irish in Ireland and abroad (Ferrie 1997c; Ó Gráda 2016; Connor 2019).

[30] 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.).

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 immigrants' children. The 1850 data are informative regarding three aspects of the famine-era migration. First, differences in the human capital characteristics (literacy and numeracy) of fathers over arrival cohorts are informative about changes in migrant selection during the famine.[31] Second, fathers' labor market outcomes are informative about the combination of changing migrant selection and 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 described 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 children from other groups whose households were observationally similar in 1850.

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.[32] The first set of variables describes the individual's father in 1850.[33] We collect literacy directly from the census, and we construct a

[31] 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 changed after arrival than did other measures, such as occupation.

[32] We have also performed the analysis of 1850 outcomes in a sample that does not condition on successful linkage. Because we have re-weighted the linked sample to match the observables of the population (of classified children) in 1850, the results are very similar. Detailed results are available on request.

[33] To be included in this analysis, a head of household had to be the father of a son in our linked sample. This does not imply that the father had a son before immigration because daughters' birthplaces and ages are also

measure of numeracy, which is an indicator for reporting an age that is not divisible by five. This variable does not indicate the numeracy of any individual, but differences between groups in the share reporting an age that ends in zero or five are informative. 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 in 1850, 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 immigrants. Relative to natives and other immigrant groups, the Irish (both arrival cohorts) 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). Between the Irish arrival cohorts, 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

$$y_{ijk} = \alpha + \sum_{j}\sum_{k} \delta_{jk} + \mathbf{x}'_{ijk}\beta + \varepsilon_{ijk}, \quad (1)$$

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 baseline controls $\mathbf{x}_{ijk}$ 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 (henceforth referred to as "the *F*-test"

---

used to determine immigration cohort; but the father must have had a son either before or after immigration. It also does not imply that the father also had to be successfully linked to 1880.

for brevity).[34] This test enables us to determine 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 evidence of a causal impact of the Irish famine (for instance, the German revolutions of 1848 might have led to a contemporaneous change in the selection and assimilation patterns of German immigrants, confounding efforts to draw such a causal conclusion).[35]

Table 4 presents estimates of δ_{jk}. Results in columns (1) and (2) indicate that Irish household heads had significantly less human capital than US-native heads of 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 (from -0.048 to -0.113 for literacy and -0.177 to -0.248 for numeracy), with no similar changes evident for the other nationalities; for ease of comparison, we report the changes over arrival cohorts in the table's lower panel. The *F*-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.

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.

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

[34] 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*.

[35] A causal conclusion would require the assumption that Irish immigrants would have followed the same trend over the pre- and post-1846 cohorts as the British and Germans in the absence of the famine shock. We do not view this as a tenable assumption. All that we claim is that if the famine Irish fared especially poorly relative to prior Irish immigrants in comparison to arrival cohort patterns within other immigrant groups, then their poor outcome is unlikely to be due solely to their later arrival or a common shock to transatlantic migration.

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, and the difference between the two cohorts is strongly statistically significant. Such large differences are not present in the other ethnicities (i.e., the between-cohort differences for the other ethnicities are smaller, and the *F*-tests indicate a significantly greater deterioration over arrival cohorts for the Irish than for other groups).[36] 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 having 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 the decline in property holdings over cohorts was actually greatest among the British. There was 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.[37] 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

---

[36] If the comparison were in the relative difference in the probability of holding an unskilled occupation rather than the absolute difference, then 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.

[37] 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.

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. Moreover, there may have been some mismatch of the new arrivals and the pre-existing network, as the geographic sources of emigration within Ireland shifted during the famine.

In sum, the new dataset of linked census records reveals that the famine-era Irish household heads were less educated and attained lower occupational status than other immigrant groups, earlier immigrants from Ireland, and US natives. There was nothing about the famine migrants that would favorably impress policymakers concerned primarily with immigrant "quality."[38] Thus, their children began life at a severe disadvantage in terms of economic resources relative to others in the US. 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 our focus to the 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, there are clear differences between the Irish and the non-Irish, as well as between the famine-era and pre-famine Irish, all in similar directions as observed for 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 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 0.48 log point or 11 percentile rank disadvantage in terms

[38] As unfortunate as the famine-era Irish household heads appear, our focus on them may suggest an overly optimistic view of famine-era Irish immigrants' average occupational status. Comparing Irish household heads ages 18-30 with children (the group included in our analysis) to all other Irish men ages 18-30 in the 1850 census (many of whom were boarders) shows that the heads with children were 10.3 percentage points less likely to hold an unskilled occupation and had a 0.25 log point advantage in the occupational wealth score. The excluded men, however, were 4.0 percentage points more likely to be literate than heads of household with children.

of occupational score. For all immigrant groups, occupational outcomes were worse for the children of later (famine-era) arrivals than for children of earlier arrivals. 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., 15.9 percentage points more likely to have an unskilled occupation). The *F*-tests reveal that the famine-era Irish disadvantage relative to earlier Irish arrivals is statistically 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 sons' outcomes in 1880. For both unskilled occupations in Figure 3(a) and the occupational wealth 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.

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 adding control variables that reflect the adverse conditions faced by the famine-era Irish children. When studying immigrants' intergenerational mobility, scholars often rely on comparisons of cross-sectional data observed roughly a generation apart; that is, they compare foreign-born workers' outcomes in one census to outcomes for US-born workers who report having foreign-born parents in a later census (see Borjas 1993; Card 2005; Abramitzky, Boustan, Eriksson 2014). By linking micro-level data from childhood to adulthood, we can take account of each child's early life conditions in studying their labor market outcomes as adults.

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. The ethnicity-cohort indicators (δ_{jk}) can thus be interpreted as the difference in intergenerational mobility between immigrants and observationally similar natives.

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., the same as in Table 6). The second, labeled "Controls," are the estimates obtained when controlling for the 1850 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.[39] 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; Alexander and Ward 2018).[40]

Figure 4 shows that 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 included, the gaps decline to 14 percentage points and 27 log points, respectively. Additionally controlling for native birth further reduces the gaps to just 8 percentage points and 12 log points. Separately, it is notable that the "controls plus nativity" results for the English and Germans are small and indicate slightly *better* performance for the immigrants' children than for observationally similar natives.

A Gelbach (2016) 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. 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.

---

[39] The coefficients of the regressions with the expanded sets of controls are presented in Appendix Table B.1.
[40] Consistent with this view, repeating the main analysis including only the US-born children of famine-era immigrants shows a greater degree of assimilation relative to the children of US-born heads. The estimates are far less precise, however, since over 85 percent of the famine-era children in the full sample are foreign born (as reported in Table 3). These results are available upon request.

Interpreting the residual gap for the famine-era Irish is, of course, difficult in this kind of analysis. It is consistent with both 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. For instance, in addition to hindrances associated with relatively low human capital parents, they might also have experienced more 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.

Discussion

The results described thus far 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. Their poverty and relative lack of human capital square with historical accounts, but the new dataset offers much more detail 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 later in the nineteenth century (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. 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, pp. F318-F319), which finds that second-generation immigrants tend to converge on or even overtake the educational attainment of US natives. Second-generation Mexican

immigrants, for instance, eliminated about 80 percent of the education gap faced by first-generation fathers (Card 2005, p. F319).[41] For an earlier period, White and Mullen (2016) describe intergenerational gains in occupational status using IPUMS microdata samples, and Hatton and Williamson (1998, pp. 128-129) do so based on the reports of the US Immigration Commission (1911). Our results also relate to recent findings by Abramitzky et al. (2019b), who study intergenerational assimilation of immigrants, with the first cohort of adults observed in 1880 and the first cohort of children observed as adults in 1910. They, too, find substantial closing of the gap in occupational status between immigrants and natives over one generation.

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 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. Our results for the pre-famine cohorts of British and German immigrants (Table 6) are somewhat different, as their children appear to have surpassed those of natives by 1880, but only slightly. Abramitzky, Boustan, and Eriksson (2014, p. 499) also find that the second generation fared better than the first (relative to US natives) 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, albeit for a small number of groups.

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 (a factor cited by Pérez 2019), and both native-born workers and

[41] 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.

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 in Ireland, whether an Irish-born child was born during or before the famine, the share of neighbors who were Irish, and the "Catholicity" 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

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

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 $\mathbf{z}_i$. The sample is restricted to sons of the famine-era Irish immigrants; within this group, some were born in Ireland before the famine (the implicit comparison group), some were born in Ireland during the famine (*famine in utero* =1), and others were born in the US (*native* =1).

Table 7 reports the results.[42] Native (US) birth was associated with greater upward mobility

[42] The number of observations in this table is reduced relative to the number in Tables 3 and 5. The difference is due to individuals whose surnames were not sufficiently common in the 1901 Irish census to determine their Catholicity measure (we require at least 200 individuals with similar names in the 1901 Irish census). Columns

in terms of the occupational wealth score relative to children born in Ireland before the famine. A similar advantage for the native born is evident in columns (1) and (3) in terms of probability of holding an unskilled occupation, but the differences are not statistically significant. 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.[43] 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 interesting since both groups spent nearly their entire lives in the US, but of course the Irish-born children experienced the famine directly.

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 controversial at the time. Their religion was seen as a barrier to assimilation in a predominantly Protestant society, and some US natives were openly hostile (Anbinder 1992; Hirota 2017). Having a "more Catholic" surname is negatively correlated with sons' labor market outcomes, even after conditioning on fathers' economic status and literacy and the Irishness of the child's neighborhood. We caution that this result is not straightforward 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 that are not sufficiently controlled for in the regression. That said, the statistical association is strong. A typical "Catholic" surname had a Catholicity of 0.95 whereas a

---

(1) and (3) 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.

[43] We recognize that some born in the US would have been in utero in Ireland during the famine. Unfortunately, because we do not observe the actual year of arrival (only a range) we cannot identify these individuals. Therefore, the coefficient on Famine In Utero may understate the correlation in question.

typical "Protestant" surname had a Catholicity of 0.18.[44] 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.

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 suggests that geographic mobility was an important channel for economic mobility by the urban poor (Turner 1920). 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 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 new dataset of linked US census records to shed light on the assimilation of a large group of desperate, poor, and culturally distinct immigrants—the Irish fleeing the Great Famine. The dataset provides several advantages relative to the existing literature. It 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. It also allows us to study a fixed group of individuals over time, which avoids spurious results driven by changes in sample composition, as emphasized 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 labor market assimilation patterns for Irish, British, and German immigrants 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

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

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 Catholicity 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, they add 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 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 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 migrant selection and assimilation in the absence of federal policy barriers to entry. 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, micro-level 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. 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 many 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, Katherine Eriksson, James J. Feigenbaum, and Santiago Pérez (2019a). "Automated Linking of Historical Data." NBER Working Paper 25825.

Abramitzky, Ran, Leah Platt Boustan, Elisa Jácome, and Santiago Pérez (2019b). "Intergenerational Mobility of Immigrants in the US over Two Centuries." Mimeo., Stanford University.

Aizer, Anna, Shari Eli, Joseph Ferrie, and Adriana Lleras-Muney (2016). "The Long-Run Impacts of Cash Transfers to Poor Families." *American Economic Review* 106:4, pp. 935-971.

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.

Alsan, Marcella, Katherine Eriksson, and Gregory Niemesh (2019). "The Rise and Fall of the Know-Nothing Party." Mimeo., Miami University.

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 (2019). "How Well Do Automated Methods Perform in Historical Samples? Evidence from New Ground Truth." *Journal of Economic Literature* Forthcoming.

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.

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.

Collins, William J. and Ariell Zimran (2019). "Replication: The Economic Assimilation of Irish Famine Migrants to the United States." Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2019-09-05. https://doi.org/10.3886/E111708V1.

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. (2019). "The Cream of the Crop? Geography, Networks, and Irish Migrant Selection in the Age of Mass Migration." *Journal of Economic History* 79:1, pp. 139-175.

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.

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.

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. (1997c). "A New View of the Irish in America: Economic Performance and the Impact of Place of Origin, 1850-1920." 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.

Gershteyn, Yefim (2000). "Use of SPEDIS Function in Finding Specific Values." Proceedings of the Twenty-fifth Annual SAS Users Group International Conference, Indianapolis.

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 (1998). *The Age of Mass Migration: Causes and Economic Impact*. New York: Oxford University Press.

Higham, John (1955). *Strangers in the Land: Patterns of American Nativism 1860-1925*. New Brunswick, NJ: Rutgers 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 Joseph P. Ferrie (2018). "Grandfathers Matter(ed): Occupational Mobility across Three Generations in the US and Britain, 1850-1911." *Economic Journal* 128, pp. F422-F445.

Long, Jason and Henry E. Siu (2018). "Refugees from Dust and Shrinking Land: Tracking the Dust Bowl Migrants." *Journal of Economic History* 78:4, pp. 1001-1033.

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 (2019). "Taken By Storm: Hurricanes, Migrant Networks, and US Immigration." *American Economic Journal: Applied Economics* Forthcoming.

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.

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.

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 (2016). "What's in an Irish Surname? Connollys and Others a Century Ago." UCD Centre for Economic Research Working Paper WP 2016/05.

Ó Gráda, Cormac (2019). "The Next World and the New World: Relief, Migration, and the Great Irish Famine." *Journal of Economic History* 79:2, pp. 319-355.

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.

Pérez, Santiago (2019). "Intergenerational Occupational Mobility Across Three Continents." *Journal of Economic History* 79:2, pp. 383-416.

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.

Turner, Frederick Jackson (1920). *The Frontier in American History*. New York: H. Holt.

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

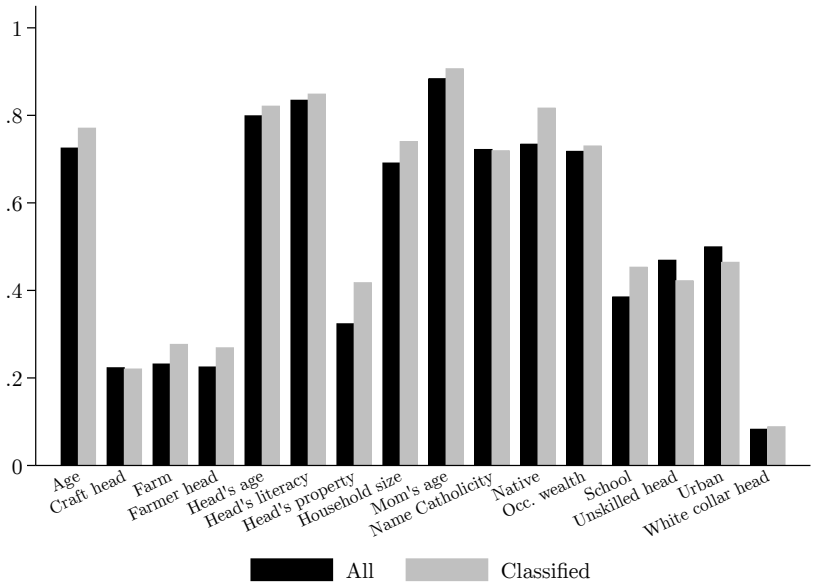
US Department of State, US Department of Homeland Security, and US Department of Health and Human Services (2017). *Proposed Refugee Admissions for Fiscal Year 2018*.

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

White, Michael J. and Erica Jade Mullen (2016). "Socioeconomic Attainment in the Ellis Island Era." *Social Science History* 40:1, pp. 147-181.

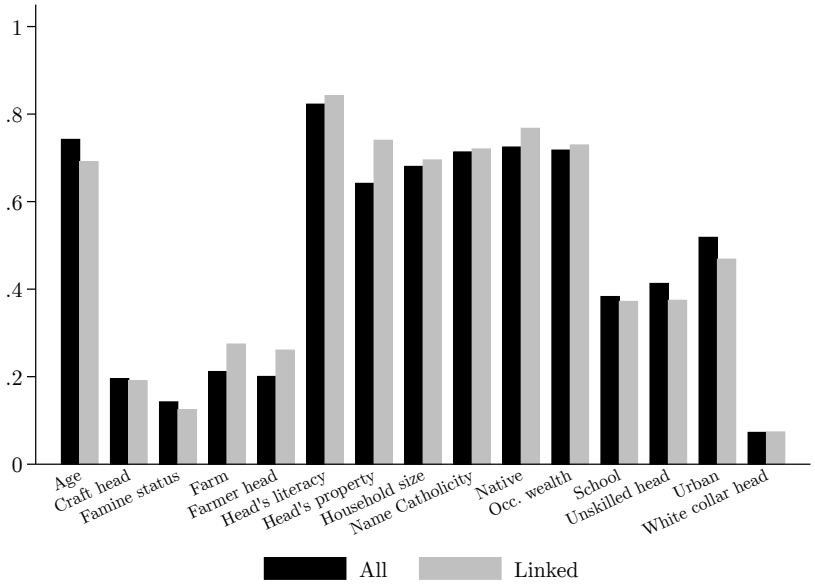
Figures

Figure 1: Selection into classification



Note: Sample covers sons of 1850 Irish male 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 (e.g., head's age is divided by 50).

Figure 2: Selection into linkage



Note: Sample covers sons of 1850 Irish male 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 for household size, age, and occupational wealth are divided by 10 needed to generate a mean below 1.

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

Figure 3(a): Unskilled Occupation

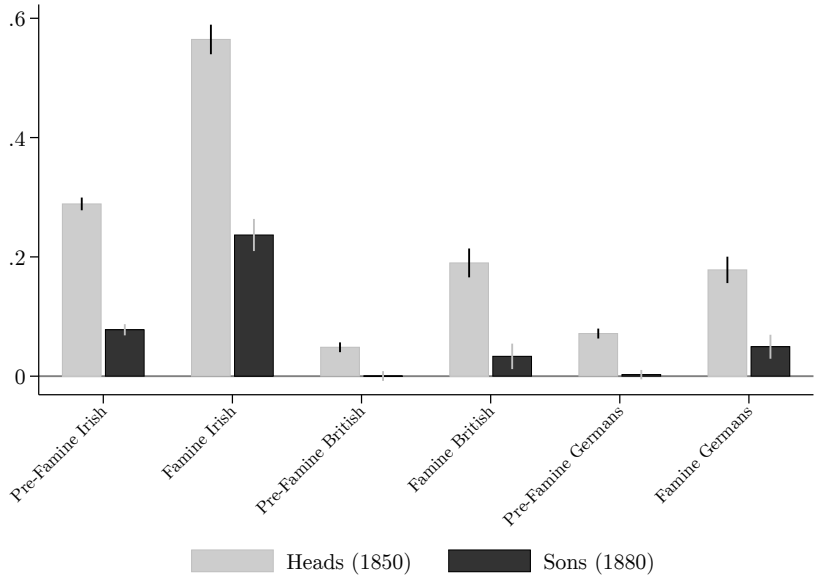
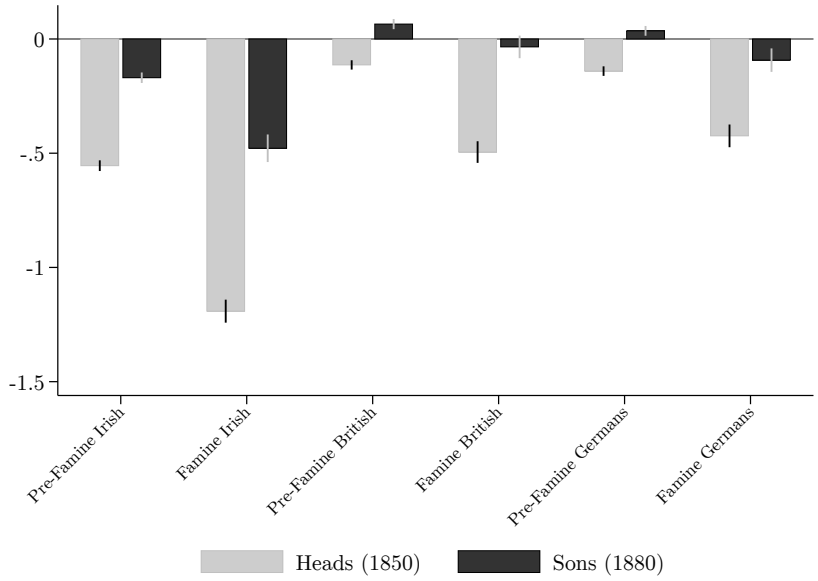


Figure 3(b): Occupational Wealth Score



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

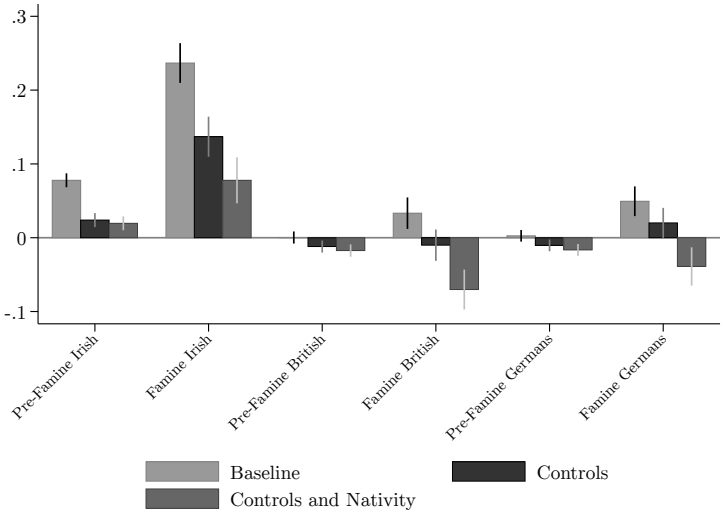
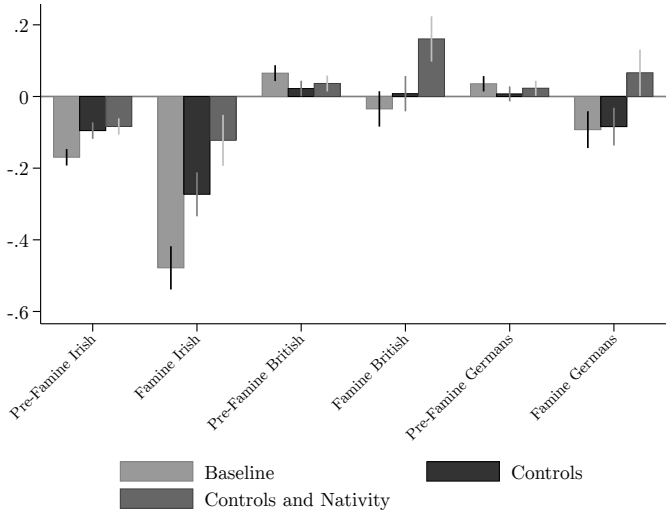


Figure 4(b): Occupational Wealth Score



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 Appendix Table B.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 Appendix Table B.1.

Tables

Table 1: Immigrants and their sons, by source country and arrival cohort

	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)	(0.607)	(0.669)	(0.545)
Famine	15,009	7,422	11,237	23,116	12,127	17,229
	(0.072)	(0.068)	(0.069)	(0.100)	(0.097)	(0.101)
Maybe Famine	6,268	2,833	7,452	11,508	5,787	14,472
	(0.030)	(0.026)	(0.046)	(0.050)	(0.046)	(0.085)
Unknown	109,846	51,511	93,517	55,865	23,591	46,014
	(0.527)	(0.471)	(0.578)	(0.242)	(0.188)	(0.269)
Total	208,361	109,360	161,844	230,457	125,541	170,778

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: 1850 to 1880 linkage rates for immigrants' children, by heads' arrival cohort, and for children of US-born heads

Famine Status	(1) Start	(2) Searched	(3) Linked
Panel A: Irish			
Pre-Famine	139,968	52,189	10,338
		(0.373)	(0.198)
			[0.074]
Famine	23,116	9,665	1,473
		(0.418)	(0.152)
			[0.064]
Panel B: British			
Pre-Famine	84,036	35,051	9,182
		(0.417)	(0.262)
			[0.109]
Famine	12,127	7,066	1,589
		(0.583)	(0.225)
			[0.131]
Panel C: Germans			
Pre-Famine	93,063	57,398	10,071
		(0.617)	(0.175)
			[0.108]
Famine	17,229	13,194	1,834
		(0.766)	(0.139)
			[0.106]
Panel D: Americans			
All	3,725,019	1,686,500	424,500
		(0.453)	(0.252)
			[0.114]

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. Numbers in square brackets indicate the fraction of observations for which a successful link was made.

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 Catholicity	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	0.116	0.163	0.079	0.093	0.034
	(0.225)	(0.229)	(0.139)	(0.172)	(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 Catholicity, Famine in utero, and Fraction Irish in 100.

Table 4: Regressions of 1850 characteristics on immigrant and arrival group

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
Δ Irish	-0.065^a	-0.071^a	0.276^a	-0.166^a	-0.637^a	-13.771^a	-0.500^a	-0.184^a
	(0.013)	(0.015)	(0.014)	(0.009)	(0.028)	(0.661)	(0.065)	(0.024)
Δ British	-0.010	0.020	0.141^a	-0.126^a	-0.381^a	-10.249^a	-0.678^a	-0.156^a
	(0.007)	(0.013)	(0.013)	(0.012)	(0.026)	(0.699)	(0.052)	(0.022)
Δ German	0.005	-0.031^b	0.107^a	-0.044^a	-0.283^a	-7.548^a	-0.471^a	-0.214^a
	(0.004)	(0.012)	(0.012)	(0.012)	(0.027)	(0.652)	(0.066)	(0.020)

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. Δ rows indicate differences between the famine and pre-famine cohorts of each ethnicity.

Table 5: Summary statistics for linked children, 1880

	Irish		British		Germans		Americans
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Pre-Famine	Famine	Pre-Famine	Famine	Pre-Famine	Famine	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 on immigrant and arrival group

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
Δ Irish	0.159^a	-0.309^a	-6.837^a
	(0.015)	(0.033)	(0.813)
Δ British	0.033^a	-0.100^a	-3.784^a
	(0.012)	(0.027)	(0.752)
Δ German	0.047^a	-0.128^a	-3.378^a
	(0.011)	(0.028)	(0.740)

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. Δ rows indicate differences between the famine and pre-famine cohorts of each ethnicity.

Table 7: Correlates of upward mobility among the famine Irish

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Head Literacy (1850)	−0.074	0.064	−0.072	0.064
	(0.045)	(0.098)	(0.045)	(0.098)
Native	−0.075	0.326[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 Catholicity	0.166[a]	−0.294[b]	0.163[a]	−0.294[b]
	(0.051)	(0.117)	(0.051)	(0.117)
Fraction Irish in 100	0.044	−0.124	0.047	−0.124
	(0.074)	(0.168)	(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 occupation 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.

Appendix A: Census linkage

This appendix describes the record linkage procedure used to construct our main sample in detail. The code used to execute the linkage is provided by Collins and Zimran (2019). The detailed procedure is as follows.

1. We extracted males with non-empty first and last names from the full count 1850 census and removed punctuation.
2. We divided the listed first name into a given name and middle initial, when one was present.
3. We replaced standard first name abbreviations (e.g., "Wm" was replaced with "William").
4. We removed any remaining spaces removed from the names.
5. After steps 1-4, we linked the 1850 dataset to itself, meaning that we linked one complete count 1850 file to another just like it. The point of this step is to find and eliminate males whose characteristics are "too similar" to those of other males in the base year. The males we attempt to link forward to 1880 are those who do not have a close match in 1850's population. The following characteristics and criteria were used:
    a. The birthplace (country for the foreign-born or US state for the native-born) matches exactly.
    b. The absolute difference in birth years is less than or equal to 4.
    c. The first three characters of last name soundex match, and either of the following two conditions are true (note that we are not making soundex matches; rather, we use soundex to identify a set of possible matches, which are evaluated based on spelling distance):
        i. The last name soundexes are identical and the SAS spelling distance (using the SAS function SPEDIS) is less than or equal to 20.[45]
        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.

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

6. We removed from the sample any individual who had a candidate match in step 5 that was someone other than himself. That is, we do not attempt to link males to 1880 when it is known that there were very similar males in the base year. This reduces the scope for false matches.
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. We linked males from 1850 to 1880 according to the same criteria listed in step 5, as well as the following additional criterion: where both records report a middle initial, the middle initials must match for a match to be made. This yields a set of candidate 1880 matches for every 1850 male that we searched for, or it yields no candidate matches.
9. We dropped any case in which more than one 1880 individual matched to an 1850 individual; we also dropped any case in which more than one 1850 individual matched to the same 1880 individual. Thus, the final set of matches includes males from 1850, without close substitutes in that year (in terms of name, birthplace, and birth year), who are located in 1880, again without close substitutes in that year.

This matching algorithm is a descendant of Ferrie (1996) in that it eliminates non-unique individuals in the base year before attempting a match, makes a match based on fixed rules for name similarity and birth year similarity, and accepts matches only when they are unique (i.e., an 1850 individual had only one match in 1880). Our method is most similar to that used by Beach et al. (2016), which is another derivative of the Ferrie (1996) approach and is in turn similar to the methods used by Aizer et al. (2016) and Long and Ferrie (2018).[46] Beach et al.'s (2016) approach also makes matches on the basis of birth year differences and SAS spelling distance comparisons. Our method differs from the Beach et al. (2016) method in the specific criteria that must be met for a match,[47] and in that we require uniqueness of the 1850 record within a nine-year band (and a band

[46] All of these methods make matches based on the same criteria that we do (soundex and SPEDIS of surname and first name), though the precise cutoffs for a match differ.

[47] The specific differences in the match criteria are as follows:

1. Beach et al. (2016) allow matches with up to a two-year difference in birth year. We allow matches with up to a four-year difference.
2. Beach et al. (2016) require an exact match of last name soundex, whereas we require an exact match only of the first three characters of the last name soundex. In both algorithms, however, matches must pass a spelling distance cutoff.
3. Beach et al. (2016) require that the average of the spelling distances in each direction be less than 15. We require that the minimum spelling distance (of the two directions) be less than either 20 or 17, depending on the soundex similarity of the last names.

defined by the name linkage criteria discussed above) in addition to uniqueness of the link to 1880, whereas the Beach et al. (2016) method requires only uniqueness of the link.

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
	(0.317)	(0.482)	(0.644)	(0.478)
Linked	18,378	18,684	24,822	892,443
	(0.112)	(0.172)	(0.111)	(0.208)
	[0.035]	[0.083]	[0.072]	[0.100]

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. Numbers in parentheses indicate the fraction of observations from the previous row that were advanced to the next. Numbers in square brackets indicate the fraction of observations for which a successful link was made.

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: Additional results

Table B.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.

Appendix C: Selection into Arrival Cohort Classification (For Online Publication)

One motivation for focusing on children's long-term outcomes is our greater ability to classify the arrival cohorts of the parents of children, as compared to all heads of household, through the use of the birthplaces and ages of children in the household. But there remains a non-negligible fraction of children (just under 30 percent) who are excluded from our main analysis because we are not able to determine their father's arrival cohort (henceforth "unclassified children"). A natural concern is that these children might be systematically different from those from the classified households, and that their exclusion might bias our results. This appendix shows that the exclusion of these children from our main analysis is unlikely to have meaningfully affected our main results.

Figure 1 in the main text shows that the Irish children who we are able to classify are observationally similar to the entire population of Irish children with male heads of household in 1850. Thus, to bias our results, the unclassified children would have to have very different 1880 outcomes from the classified children despite having very similar 1850 characteristics. Figures C.1(a) and C.1(b) repeat this analysis for the British and German children, finding that in these populations the classified sub-sample is also similar to the population of interest on the basis of 1850 observables. Figure C.1(c) compares the selection into classification for the three ethnicities, plotting the ratio of the average value of each 1850 variable for the classified sub-set of children relative to all children (if the classified groups were identical to the population, all the ratios would be 1). Reassuringly, the selection patterns are similar across the ethnicities—the same variables are over- and under-represented across ethnicities, or the ratios for all three ethnicities are very close to one. Moreover, with the exception of the stronger over-representation of the children of farmers among the classified Irish as compared to the other ethnicities, the Irish are not very different in their selection into classification from the other groups.

We can also rule out major differences in the assimilation patterns of the unclassified children, further supporting the argument that their exclusion from analysis has not biased our results. In Tables C.1, C.2, and C.3, we repeat the main analyses of the paper including the unclassified children of each ethnicity as separate categories, with their outcomes again compared to those of the children of natives. The unclassified children exhibit average outcomes in all regressions that are between the famine and pre-famine groups of their ethnicities. This reflects the fact that the unclassified children are, by definition, a mix of those whose household head arrived before 1846 and those whose household head arrived later. The unclassified Irish children also exhibit considerable intergenerational assimilation and also upgraded almost as fast as observationally similar children of natives, just as the famine-era Irish children did. Thus, if we were

able to classify these children and include them in our analysis, our main results of rapid assimilation by the children of the famine-era Irish would likely not be affected.

To more systematically test how sensitive our results might be to including the unclassified children, we perform a simulation exercise. We randomly assign each linked but unclassified child to either the famine or pre-famine cohort for his ethnicity, and then repeat our main analyses. Doing this 500 times provides a range of point estimates corresponding to various allocations of the unclassified children across arrival cohorts. Note that since the unclassified children are an unknown mix of pre-famine and famine-era cohorts, randomly assigning them to one group or the other will tend to pull the arrival cohorts' point estimates closer together (relative to results using only the definitively classified children); the more interesting question is whether the point estimates under the simulation have a wide range, indicating that some allocations of the unclassified children yield strongly different results than others.

The main difficulty in this analysis is that we do not know the fraction of the unclassified cohort that truly belongs to the famine or pre-famine cohorts. We therefore repeat this exercise with a variety of assumptions regarding the probability with which an unclassified Irish child should be assigned to the famine cohort. Specifically, we repeat the exercise for a probability of 0.1247 (the fraction of classified Irish children whose fathers were famine-era immigrants, as shown in Table 1 of the main text), as well as probabilities of 0.30, 0.50, 0.70, and 0.90. In each exercise, we fix the probability with which unclassified British and German children are assigned to their ethnicity's famine cohort to the fraction of classified children for each ethnicity whose fathers were famine-era immigrants (0.1475 and 0.1541, respectively, as shown in Table 1).

Figures C.2 and C.3 present the results of these exercises for two outcomes—unskilled occupation and the occupational wealth score—for famine assignment probabilities for the Irish of 0.1247 and 0.9.[48] Each panel of these figures presents the point estimates from our main analysis (i.e., results excluding the unclassified children) using the bars, and the complete range of point estimates from the 500 random assignments of the unclassified group using the lines (resembling confidence intervals). We present estimates for the 1850 outcomes of heads of household, for the 1880 outcomes of children, and for the 1880 outcomes of children conditional on 1850 observables and conditional on observables and the child's birth country. For all of the probabilities by which we assign unclassified Irish children to the famine cohort, the *range* of point estimates from the random assignment is small. As expected, assigning a mixture of children of the famine and pre-famine

[48] The results for famine probabilities of 0.3, 0.5, and 0.7 are similar and are omitted for brevity.

cohorts (from the unclassified) to the famine cohort causes it to appear more similar to the pre-famine cohort, and vice versa. In all cases, the range of point estimates from the simulation continues to show strong evidence that the famine Irish children experienced strong intergenerational assimilation relative to natives, and upgraded almost as fast as observationally similar natives. Thus, we are confident that any bias that arises from our inability to classify the arrival cohorts of all children's heads of household is not responsible for any of our qualitative results.

Figure C.1: Selection into classification

Figure C.1(a): British

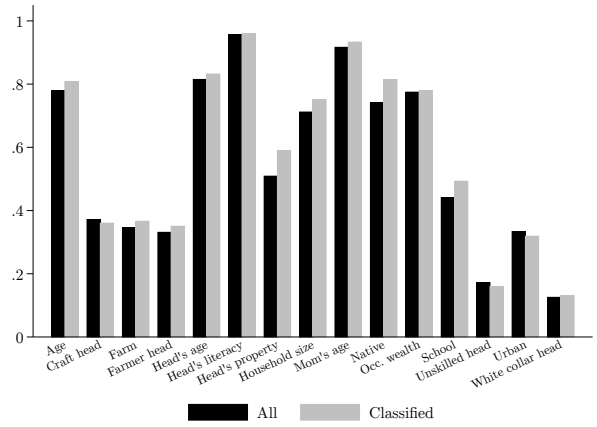


Figure C.1(b): Germans

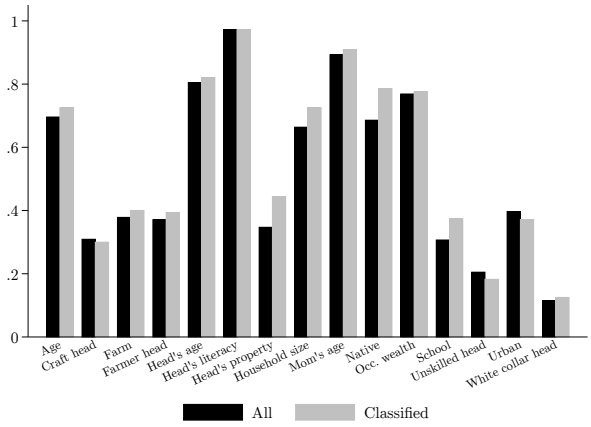
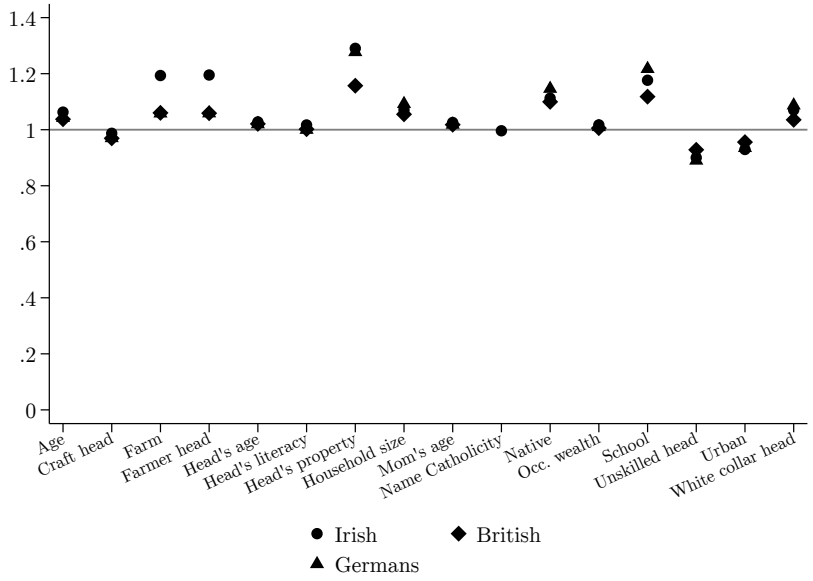


Figure C.1(c): Comparison across ethnicities



Note: Panels (a) and (b) are analogous to Figure 1 of the paper. Panel (c) presents ratios of the value of the variable on the *x*-axis for the classified group relative to the total population (i.e., the ratio of the light and dark bars in Figure 1 and panels (a) and (b) of this figure) for each ethnicity.

Figure C.2: Results with simulated assignment of unclassified arrival cohorts, unskilled occupation

Figure C.2(a): Famine probability 0.12

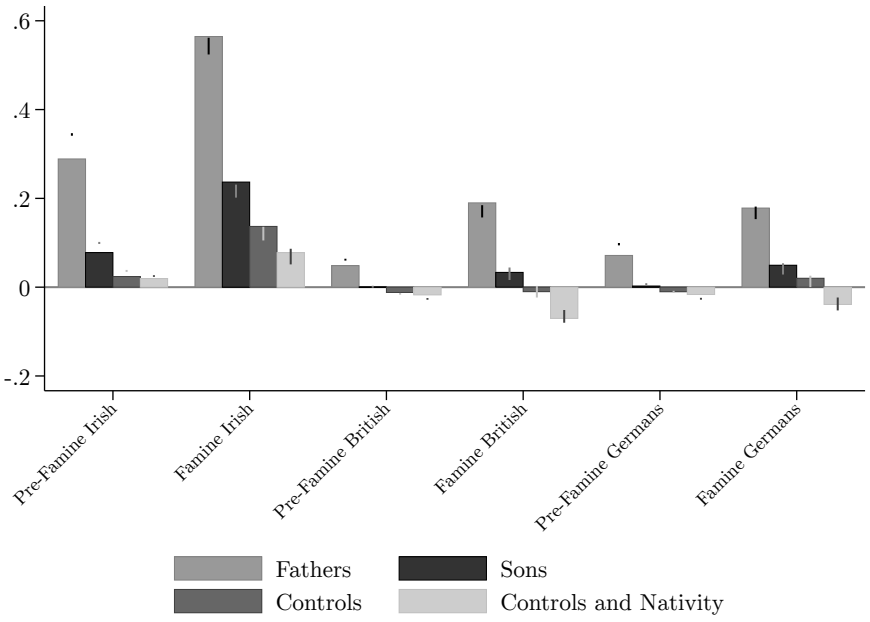
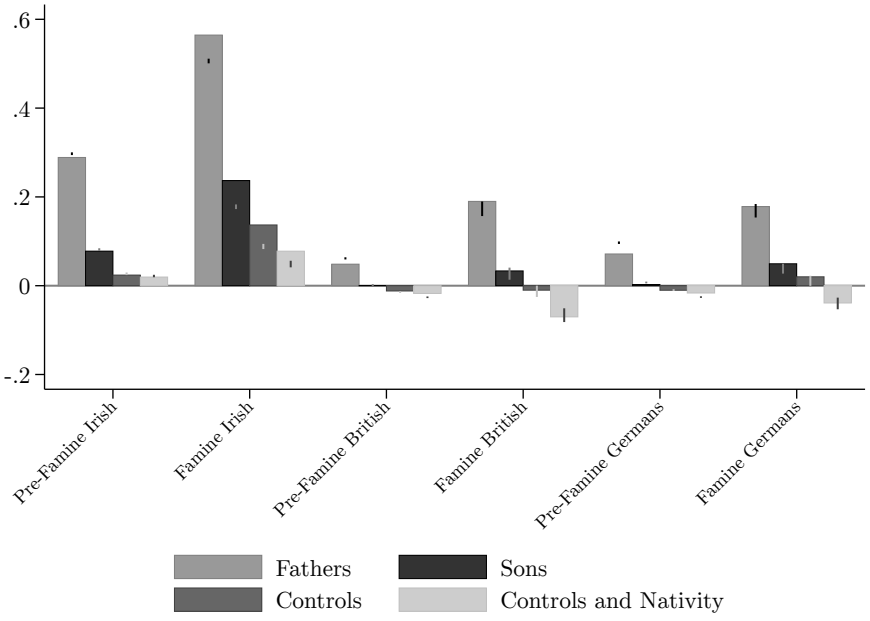


Figure C.2(b): Famine probability 0.90



Note: The bars present the point estimates from our baseline sample excluding the unclassified. The vertical lines present the range of point estimates reached from 500 random assignments of the unclassified group to the famine or pre-famine cohort for their ethnicity using the probability in the subfigure title for the Irish and the probabilities listed in text for the British and Germans.

Figure C.3: Results with simulated assignment of unclassified arrival cohorts, occupational wealth

Figure C.3(a): Famine probability 0.12

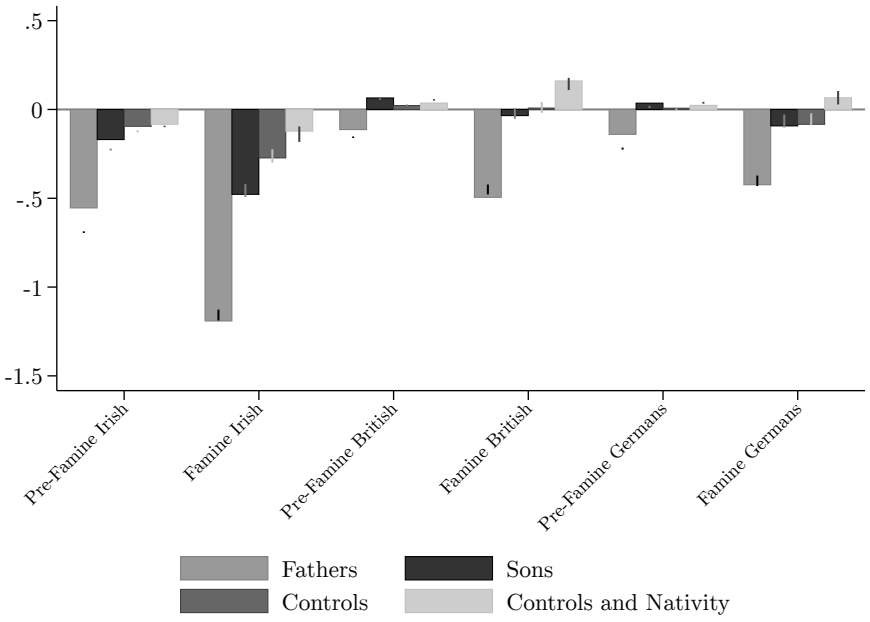
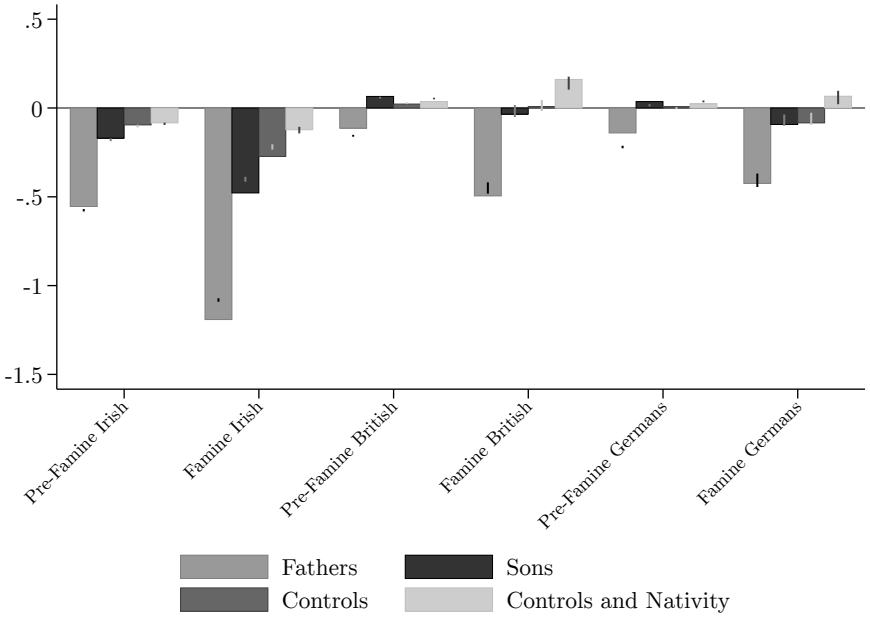


Figure C.3(b): Famine probability 0.90



Note: See Figure C.2.

Table C.1: Regressions of 1850 characteristics, including unclassified

Variables	(1) Literacy	(2) Numeracy	(3) Unskill	(4) Farmer	(5) Occ. Wealth	(6) Occ. Rank	(7) Property	(8) School
Pre-Famine Irish	-0.047^a	-0.177^a	0.289^a	-0.345^a	-0.554^a	-11.230^a	-0.624^a	0.043^a
	(0.004)	(0.005)	(0.005)	(0.005)	(0.012)	(0.307)	(0.058)	(0.007)
Famine Irish	-0.113^a	-0.249^a	0.565^a	-0.511^a	-1.192^a	-25.323^a	-1.136^a	-0.142^a
	(0.012)	(0.014)	(0.013)	(0.008)	(0.026)	(0.595)	(0.036)	(0.023)
Pre-Famine British	0.055^a	-0.040^a	0.048^a	-0.274^a	-0.113^a	-1.607^a	-0.302^a	0.067^a
	(0.002)	(0.005)	(0.004)	(0.005)	(0.010)	(0.303)	(0.049)	(0.007)
Famine British	0.044^a	-0.022^c	0.190^a	-0.400^a	-0.496^a	-11.958^a	-0.998^a	-0.090^a
	(0.007)	(0.012)	(0.012)	(0.011)	(0.024)	(0.633)	(0.027)	(0.021)
Pre-Famine Germans	0.066^a	-0.020^a	0.071^a	-0.237^a	-0.140^a	-1.617^a	-0.579^a	-0.023^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	-9.243^a	-1.059^a	-0.237^a
	(0.004)	(0.011)	(0.011)	(0.011)	(0.025)	(0.597)	(0.033)	(0.019)
Unclassified Irish	-0.100^a	-0.221^a	0.484^a	-0.508^a	-1.038^a	-22.166^a	-1.103^a	-0.158^a
	(0.007)	(0.008)	(0.008)	(0.005)	(0.016)	(0.374)	(0.020)	(0.015)
Unclassified British	0.051^a	-0.057^a	0.112^a	-0.360^a	-0.313^a	-7.001^a	-0.858^a	-0.104^a
	(0.004)	(0.009)	(0.008)	(0.008)	(0.018)	(0.487)	(0.049)	(0.015)
Unclassified Germans	0.072^a	-0.040^a	0.145^a	-0.298^a	-0.368^a	-7.458^a	-1.044^a	-0.245^a
	(0.002)	(0.006)	(0.006)	(0.006)	(0.014)	(0.337)	(0.022)	(0.011)
Observations	423,451	423,451	423,451	423,451	423,451	423,451	423,451	206,809
R-squared	0.009	0.012	0.085	0.076	0.063	0.040	0.009	0.029
F-Test	14.232^a	10.684^a	45.890^a	31.535^a	42.910^a	24.015^a	4.227^b	1.846

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 C.2: Regressions of 1880 characteristics, including unclassified

Variables	(1) Unskill	(2) Occ. Wealth	(3) Occ. Rank
Pre-Famine Irish	0.078^a	-0.170^a	-4.426^a
	(0.005)	(0.012)	(0.308)
Famine Irish	0.237^a	-0.480^a	-11.419^a
	(0.014)	(0.031)	(0.759)
Pre-Famine British	0.000	0.066^a	1.811^a
	(0.004)	(0.011)	(0.309)
Famine British	0.034^a	-0.036	-2.067^a
	(0.011)	(0.025)	(0.691)
Pre-Famine Germans	0.003	0.035^a	0.934^a
	(0.004)	(0.011)	(0.288)
Famine Germans	0.050^a	-0.095^a	-2.568^a
	(0.010)	(0.026)	(0.687)
Unclassified Irish	0.154^a	-0.370^a	-9.714^a
	(0.008)	(0.018)	(0.436)
Unclassified British	0.007	0.033^c	0.261
	(0.007)	(0.019)	(0.515)
Unclassified Germans	0.016^a	-0.023	-0.744^b
	(0.005)	(0.014)	(0.370)
Observations	423,451	411,392	411,392
R-squared	0.015	0.019	0.016
F-Test	25.964^a	13.208^a	5.837^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 C.3: Conditional differences between groups, including unclassified

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Pre-Famine Irish	0.025^a	-0.097^a	0.020^a	-0.086^a
	(0.005)	(0.012)	(0.005)	(0.012)
Famine Irish	0.140^a	-0.279^a	0.075^a	-0.123^a
	(0.014)	(0.031)	(0.015)	(0.034)
Pre-Famine British	-0.012^a	0.023^b	-0.018^a	0.037^a
	(0.004)	(0.011)	(0.004)	(0.011)
Famine British	-0.009	0.006	-0.075^a	0.165^a
	(0.011)	(0.025)	(0.012)	(0.029)
Pre-Famine Germans	-0.011^a	0.007	-0.017^a	0.023^b
	(0.004)	(0.011)	(0.004)	(0.011)
Famine Germans	0.021^b	-0.086^a	-0.044^a	0.070^b
	(0.010)	(0.027)	(0.012)	(0.030)
Unclassified Irish	0.068^a	-0.196^a	0.033^a	-0.110^a
	(0.008)	(0.018)	(0.008)	(0.018)
Unclassified British	-0.022^a	0.037^b	-0.061^a	0.130^a
	(0.007)	(0.018)	(0.008)	(0.020)
Unclassified Germans	-0.008	-0.014	-0.045^a	0.076^a
	(0.005)	(0.014)	(0.006)	(0.016)
Native			-0.074^a	0.179^a
			(0.006)	(0.015)
Observations	423,451	411,392	423,451	411,392
R-squared	0.048	0.091	0.049	0.092
F-Test	18.739^a	7.471^a	18.702^a	7.524^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.

Appendix D: Alternative Matching Methods (For Online Publication)

To verify that our results are not the spurious products of some peculiarity in our linkage method, we created two alternative linked samples and repeated the main analysis using these methods. This appendix describes the results obtained when using these alternative linkage methods.

We use two methods described and analyzed by Abramitzky et al. (2019a).[49] The first is the ABE linkage method based on NYSIIS standardization of names, and requiring uniqueness within a five-year band. The second is the ABE linkage method requiring exact matches by name, and requiring uniqueness within a five-year band. We do not use middle names in either of these matches. We refer to these two methods as "ABE-NYSIIS, 5-Year Band" and "ABE-Exact, 5 Year Band," respectively, in the following discussion.

Table D.1 presents the match rates by birth place for each method. Table D.2 presents the match rates by birth place for each famine-cohort group. For both matching methods, the size of the resulting linked sample is very similar to the linked sample obtained by our linking method (Tables A.1 and 2). This is particularly true of the Irish, in which the difference in the sample size is only 13 observations between our linkage method and the ABE-Exact method (Tables 2 and D.2).

Table D.3 compares the matches made by the various linkage methods, limiting the sample to individuals who are sons of heads of household in 1850 (the candidate group for inclusion in our main sample), and categorizing them by their fathers' birthplaces (the categorization used in the paper). Table D.3(a) compares the matches made by our method to those made by the two ABE methods. In the first row of each panel, columns (1), (2), (4), and (5) show that for 1850 observations in which both our linkage method and either of the ABE linkage methods made a link to 1880, these links almost always agree. For the Irish, our algorithm and the ABE-Exact algorithm agree in over 98 percent of cases (6,505 out of 6,636 matches). This rate of agreement is roughly comparable to those for the other nationalities and the NYSIIS-based linking method.

The first row in each panel of columns (3) and (6) of Table D.3(a) and the second row in each panel of columns (1), (2), (4), and (5) also show that there is a large number of 1850 observations for which either our linkage method made a match in 1880 and the ABE methods did not, or the ABE methods made a match but our method did not. This is unsurprising given the different linkage rules used by each method. For instance, our method allows matches to be made between observations whose age-implied birthyears are three or four years apart, whereas the ABE method does not. Similarly, by using orthographic distance, our method allows matches in cases where there is not an

[49] We executed these linkages using Stata code provided by Abramitzky et al. (2019a).

exact agreement of name, or an agreement of the NYSIIS standardization of the names. Conversely, the ABE algorithm makes a link in cases where the observations are unique within a five-year band but not in a nine-year band, whereas our algorithm requires uniqueness within a nine-year band. Table D.3(b) contextualizes these observations that are matched by only one algorithm by producing an analogous table for the two ABE algorithms. When matches are made by both of these algorithms, they always agree. But the rate at which only one algorithm makes a match, while smaller than the comparable rate in Table D.3(a), is not substantially smaller.

Figure D.1 studies the selection into linkage on the basis of observable characteristics for each of the two linkage methods. The over-representation of individuals from households of higher wealth and occupational status, as in our algorithm (Figure 2), is standard and similar across the three methods.

Tables D.4-D.9 and Figures D.2 and D.3 repeat the main results of the paper using the two alternate linked samples. The results are remarkably similar to those using our linkage method. The only exception is in Table D.8, which studies the correlates of upward mobility among the famine Irish. In this case, the advantage of the native-born children of the famine Irish is smaller than the analogous estimate in Table 7 and not statistically significant. Nonetheless, the sign of the coefficient is the same as that of Table 7.

Given the strong similarity of results reached regardless of the linkage method, we conclude that our results are not spurious products of peculiarities of our linkage method.

Figure D.1: Selection into linkage, alternate matches

Figure D.1(a): ABE-NYSIIS 5 Year Band

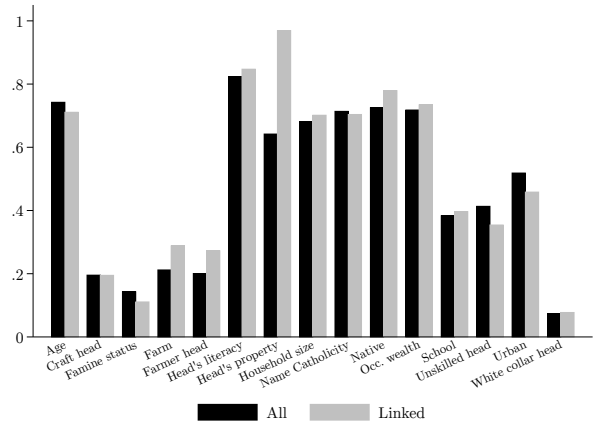
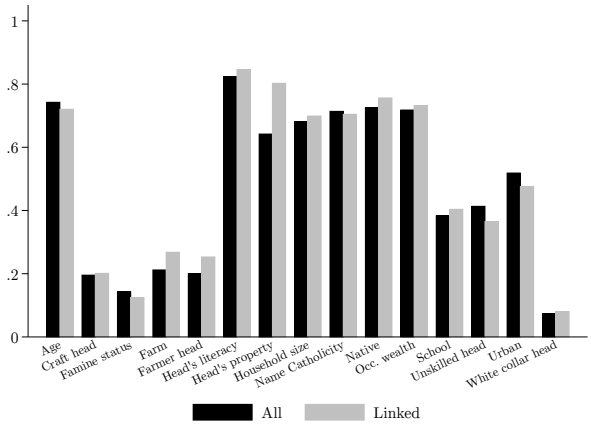


Figure D.1(b): ABE-Exact 5 Year Band



Note: Sample covers sons of 1850 Irish male heads of household. The "Linked" group contains 1850 sons who were successfully linked to 1880 by the method listed above the graph. The "All" group includes these individuals as well as those whose could not be linked. Values for household size, age, and occupational wealth are divided by 10 needed to generate a mean below 1.

Figure D.2: Comparison of adults' and childrens' outcomes, alternate matches

Figure D.2(a): Unskilled Occupation, NYSIIS

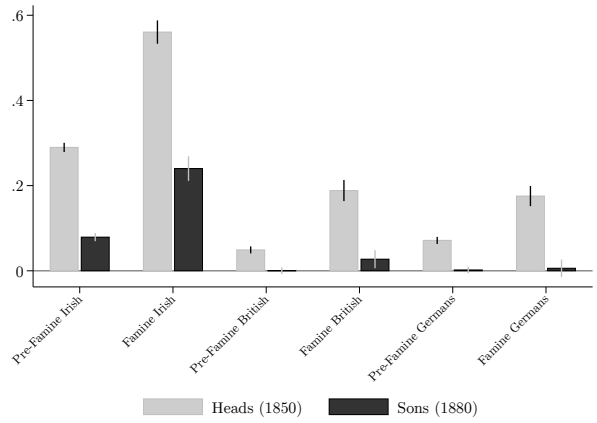


Figure D.2(b): Occupational Wealth Score, NYSIIS

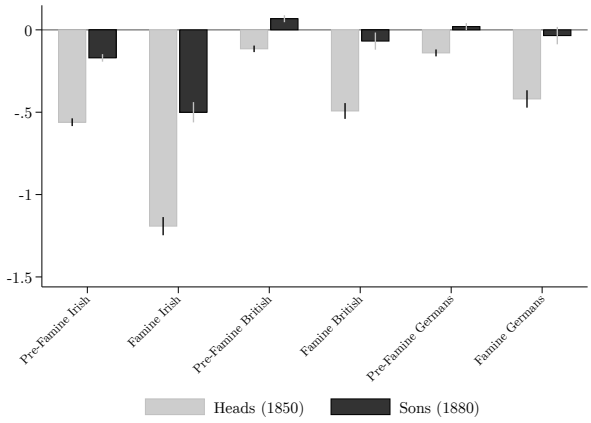


Figure D.2(c): Unskilled Occupation, Exact

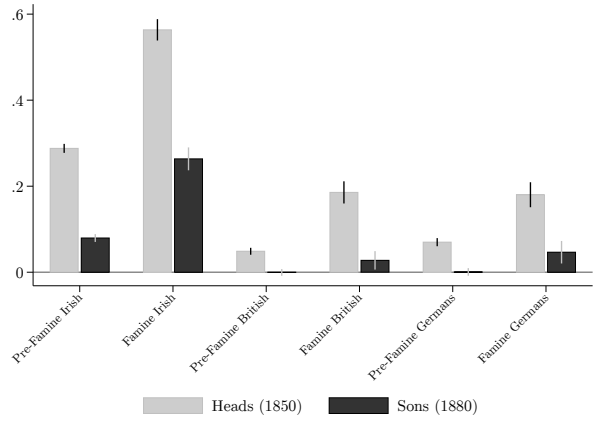
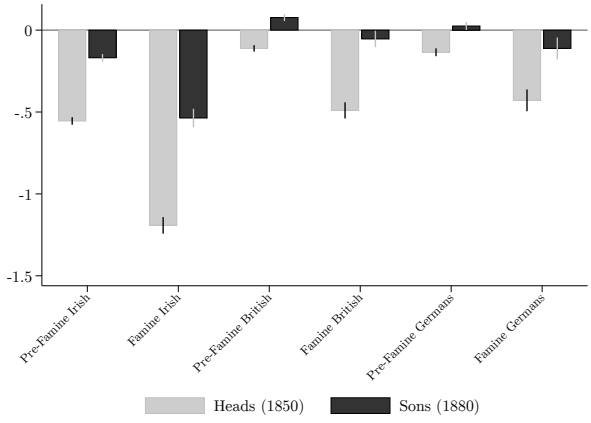


Figure D.2(d): Occupational Wealth Score, Exact



Note: Linkage methods are ABE-NYSIIS 5 Year Band for panels (a) and (b), and ABE-Exact 5 Year Band for panels (c) and (d). Bars in panel (a) and (c) for 1850 heads of household are coefficients from column (3) of Tables D.5(a) and D.5(b), respectively. Bars in panels (b) and (d) for 1880 sons are coefficients from column (1) of Tables D.7(a) and D.7(b), respectively. Bars in panels (b) and (d) for 1850 heads of household are coefficients from column (5) of Tables D.5(a) and D.5(b), respectively. Bars in panel (b) and (d) for 1880 sons are coefficients from column (2) of Tables D.7(a) and D.7(b), respectively.

Figure D.3: Unconditional and conditional coefficients for 1880 characteristics, alternate matches

Figure D.3(a): Unskilled Occupation, NYSIIS

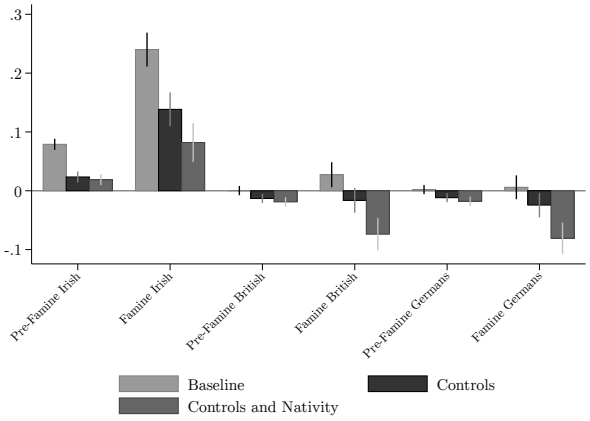


Figure D.3(b): Occupational Wealth Score, NYSIIS

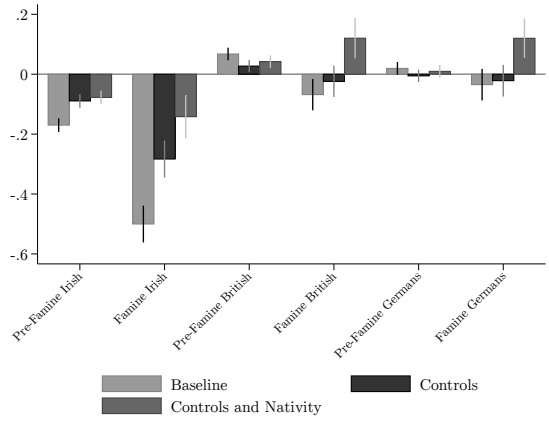


Figure D.3(c): Unskilled Occupation, Exact

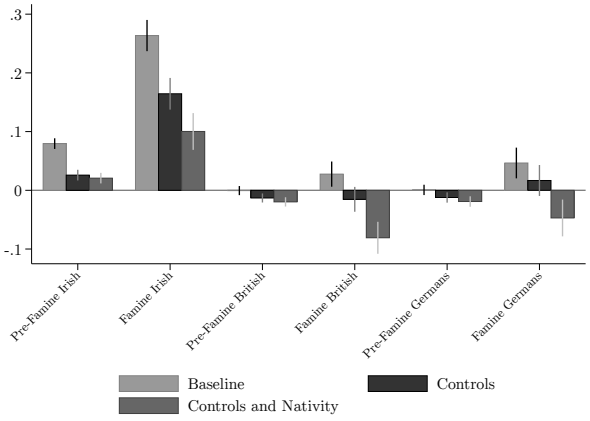
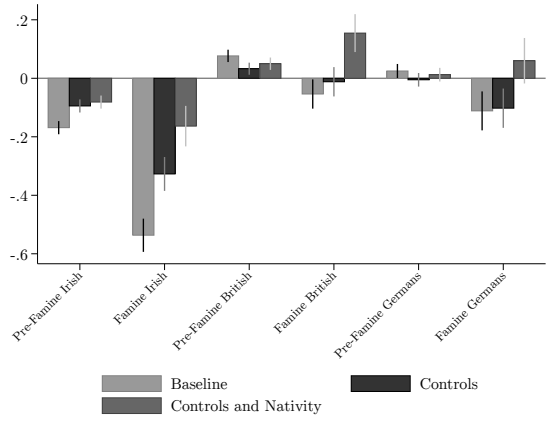


Figure D.3(d): Occupational Wealth Score, Exact



Note: Linkage methods are ABE-NYSIIS 5 Year Band for panels (a) and (b), and ABE-Exact 5 Year Band for panels (c) and (d). 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 D.7. 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 D.9. 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 D.9.

Table D.1: Linkage rates by place of birth, alternate matches

	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
Linked (NYSIIS)	18,143	18,566	24,165	1,004,266
	(0.035)	(0.083)	(0.070)	(0.112)
Linked (Exact)	20,655	18,187	16,656	934,470
	(0.040)	(0.081)	(0.048)	(0.104)

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 rows labeled Linked are those for whom a unique match could be made in the 1880 census by the ABE method using the listed name-matching and the five-year uniqueness band with no middle names. Numbers in parentheses indicate the fraction of observations for which a successful link was made.

Table D.2: Linkage rates for the seven categories, alternate matches

Famine Status	(1) Start	(2) NYSIIS	(3) Exact
Panel A: Irish			
Pre-Famine	139,968	10,328	10,326
		(0.074)	(0.074)
Famine	23,116	1,285	1,472
		(0.056)	(0.064)
Panel B: British			
Pre-Famine	84,036	9,809	9,915
		(0.117)	(0.118)
Famine	12,127	1,510	1,458
		(0.125)	(0.120)
Panel C: Germans			
Pre-Famine	93,063	10,030	7,714
		(0.108)	(0.083)
Famine	17,229	1,576	1,025
		(0.091)	(0.059)
Panel D: Americans			
All	3,725,019	467,413	442,875
		(0.125)	(0.119)

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) who were successfully linked to the 1880 census using the ABE-NYSIIS 5 year band uniqueness method without middle names. Column (3) enumerates those individuals from column (1) who were successfully linked to the 1880 census using the ABE-Exact 5 year band uniqueness method without middle names. Numbers in parentheses indicate the fraction of observations from column (1) that were linked.

Table D.3: Comparison of matches made by the various linkage methods

Table D.3(a): Comparing our matching method to ABE

	ABE-NYSIIS, 5 yr Band			ABE-Exact, 5 yr Band		
	Matched		Unmatched	Matched		Unmatched
Our Linkage	Agree (1)	Disagree (2)	(3)	Agree (4)	Disagree (5)	(6)
Panel A: Irish						
Matched	7,701	338	9,342	6,505	131	10,745
Unmatched	8,775		217,640	10,647		215,768
Panel B: British						
Matched	8,334	209	6,067	7,754	115	6,741
Unmatched	6,705		106,956	7,391		106,270
Panel C: Germans						
Matched	7,818	409	10,321	5,877	95	12,576
Unmatched	9,718		148,422	7,204		150,936
Panel D: Americans						
Matched	270,034	3,509	150,957	239,512	2,365	182,623
Unmatched	193,870		3,099,464	200,998		3,092,336

Notes: Sample limited to individuals who are male sons of heads of household and under age 18 in 1850. Nationalities in the panel headers refer to the birthplace of the father. This Table compares the matches created by our matching method to those created by the ABE-NYSIIS, 5 yr band and ABE-Exact, 5 yr band methods, neither using middle names.

Table D.3(b): Comparing the two ABE methods

	ABE-NYSIIS, 5 yr Band		
	Matched		Unmatched
ABE-Exact, 5 yr Band	Agree (1)	Disagree (2)	(3)
Panel A: Irish			
Matched	8,915	0	7,899
Unmatched	8,368		218,614
Panel B: British			
Matched	9,992	0	5,256
Unmatched	5,268		107,755
Panel C: Germans			
Matched	6,878	0	11,067
Unmatched	6,298		152,445
Panel D: Americans			
Matched	321,841	0	145,572
Unmatched	121,034		3,129,387

Notes: Sample limited to individuals who are male sons of heads of household and under age 18 in 1850. Nationalities in the panel headers refer to the birthplace of the father. This Table compares the matches created by the ABE-NYSIIS, 5 yr band and ABE-Exact, 5 yr band methods, neither using middle names.

Table D.4: Summary statistics for linked children, 1850, alternate matches

Table D.4(a): ABE-NYSIIS 5 Year Band

	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	0.856	0.788	0.960	0.947	0.972	0.975	0.904
	(0.351)	(0.409)	(0.195)	(0.223)	(0.165)	(0.158)	(0.295)
Numeracy	0.563	0.493	0.697	0.740	0.718	0.698	0.741
	(0.496)	(0.500)	(0.460)	(0.439)	(0.450)	(0.459)	(0.438)
Head's labor market characteristics							
ln(Occ. Wealth)	7.388	6.714	7.840	7.402	7.809	7.499	7.937
	(1.101)	(0.948)	(0.897)	(0.916)	(0.981)	(1.025)	(0.807)
Occ. Rank	29.834	15.008	39.514	27.905	39.445	31.362	40.462
	(27.790)	(21.637)	(26.590)	(24.219)	(26.718)	(24.921)	(22.233)
White collar	0.094	0.048	0.137	0.077	0.133	0.081	0.079
	(0.292)	(0.214)	(0.343)	(0.267)	(0.339)	(0.273)	(0.270)
Craft	0.227	0.185	0.353	0.422	0.300	0.303	0.185
	(0.419)	(0.389)	(0.478)	(0.494)	(0.458)	(0.460)	(0.388)
Farmer	0.295	0.101	0.370	0.204	0.402	0.339	0.639
	(0.456)	(0.301)	(0.483)	(0.403)	(0.490)	(0.474)	(0.480)
Unskilled	0.384	0.666	0.141	0.296	0.165	0.277	0.097
	(0.486)	(0.472)	(0.348)	(0.457)	(0.371)	(0.448)	(0.296)
Real property (1,000)	1.024	0.107	1.297	0.232	0.922	0.296	1.556
	(8.126)	(1.064)	(4.295)	(1.065)	(2.731)	(0.806)	(7.829)
Son's characteristics							
School (ages 5-15)	0.718	0.528	0.739	0.568	0.646	0.417	0.675
	(0.450)	(0.500)	(0.439)	(0.496)	(0.478)	(0.493)	(0.468)
Native	0.921	0.124	0.905	0.109	0.905	0.126	
	(0.270)	(0.329)	(0.293)	(0.311)	(0.294)	(0.332)	
Irish name Catholicity	0.686	0.749					
	(0.309)	(0.275)					
Famine in utero		0.402					
		(0.491)					
Place of residence characteristics							
Urban	0.451	0.547	0.313	0.377	0.360	0.423	0.106
	(0.498)	(0.498)	(0.464)	(0.485)	(0.480)	(0.494)	(0.308)
Fraction Irish in 100	0.248	0.289	0.117	0.162	0.079	0.094	0.034
	(0.221)	(0.223)	(0.140)	(0.169)	(0.111)	(0.120)	(0.067)
Observations	9,961	1,209	9,547	1,466	9,765	1,502	415,569

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 Catholicity, Famine in utero, and Fraction Irish in 100.

Table D.4(b): ABE-Exact 5 Year Band

	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	0.855	0.791	0.960	0.946	0.971	0.975	0.903
	(0.353)	(0.407)	(0.196)	(0.226)	(0.168)	(0.157)	(0.296)
Numeracy	0.566	0.488	0.707	0.744	0.708	0.699	0.741
	(0.496)	(0.500)	(0.455)	(0.436)	(0.455)	(0.459)	(0.438)
Head's labor market characteristics							
ln(Occ. Wealth)	7.392	6.712	7.842	7.403	7.812	7.489	7.935
	(1.101)	(0.952)	(0.896)	(0.918)	(0.980)	(1.031)	(0.808)
Occ. Rank	30.256	15.085	39.908	28.119	39.916	31.471	40.834
	(28.115)	(21.811)	(26.793)	(24.368)	(26.900)	(25.331)	(22.500)
White collar	0.096	0.051	0.137	0.076	0.134	0.078	0.079
	(0.294)	(0.221)	(0.343)	(0.266)	(0.341)	(0.269)	(0.270)
Craft	0.227	0.180	0.351	0.424	0.299	0.302	0.185
	(0.419)	(0.385)	(0.477)	(0.494)	(0.458)	(0.459)	(0.388)
Farmer	0.295	0.100	0.371	0.206	0.403	0.338	0.638
	(0.456)	(0.300)	(0.483)	(0.405)	(0.490)	(0.473)	(0.480)
Unskilled	0.382	0.669	0.141	0.294	0.164	0.282	0.097
	(0.486)	(0.471)	(0.348)	(0.456)	(0.370)	(0.450)	(0.296)
Real property (1,000)	0.963	0.117	1.301	0.236	0.936	0.300	1.606
	(4.171)	(0.877)	(4.278)	(1.126)	(2.749)	(0.987)	(8.713)
Son's characteristics							
School (ages 5-15)	0.721	0.524	0.738	0.563	0.644	0.429	0.675
	(0.448)	(0.500)	(0.440)	(0.496)	(0.479)	(0.496)	(0.468)
Native	0.925	0.123	0.907	0.107	0.906	0.130	
	(0.263)	(0.328)	(0.291)	(0.309)	(0.293)	(0.337)	
Irish name Catholicity	0.689	0.757					
	(0.308)	(0.273)					
Famine in utero		0.398					
		(0.490)					
Place of residence characteristics							
Urban	0.453	0.547	0.313	0.375	0.360	0.426	0.106
	(0.498)	(0.498)	(0.464)	(0.484)	(0.480)	(0.495)	(0.308)
Fraction Irish in 100	0.251	0.296	0.116	0.163	0.079	0.097	0.035
	(0.227)	(0.228)	(0.138)	(0.172)	(0.110)	(0.126)	(0.068)
Observations	9,966	1,402	9,647	1,411	7,515	972	394,926

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 Catholicity, Famine in utero, and Fraction Irish in 100.

Table D.5: Regressions of 1850 characteristics, alternate matches

Table D.5(a): ABE-NYSIIS 5 Year Band

Variables	(1) Literacy	(2) Numeracy	(3) Unskill	(4) Farmer	(5) Occ. Wealth	(6) Occ. Rank	(7) Property	(8) School
Pre-Famine Irish	-0.050^a	-0.176^a	0.290^a	-0.345^a	-0.561^a	-10.933^a	-0.586^a	0.047^a
	(0.004)	(0.005)	(0.006)	(0.005)	(0.012)	(0.302)	(0.079)	(0.007)
Famine Irish	-0.115^a	-0.251^a	0.561^a	-0.512^a	-1.192^a	-24.714^a	-1.194^a	-0.133^a
	(0.012)	(0.015)	(0.014)	(0.009)	(0.028)	(0.634)	(0.031)	(0.025)
Pre-Famine British	0.054^a	-0.041^a	0.049^a	-0.275^a	-0.115^a	-1.408^a	-0.362^a	0.063^a
	(0.002)	(0.005)	(0.004)	(0.005)	(0.010)	(0.294)	(0.047)	(0.007)
Famine British	0.046^a	-0.006	0.188^a	-0.402^a	-0.493^a	-11.549^a	-0.991^a	-0.098^a
	(0.006)	(0.012)	(0.013)	(0.011)	(0.024)	(0.655)	(0.032)	(0.022)
Pre-Famine Germans	0.065^a	-0.021^a	0.071^a	-0.236^a	-0.140^a	-1.334^a	-0.682^a	-0.023^a
	(0.002)	(0.005)	(0.004)	(0.005)	(0.011)	(0.281)	(0.036)	(0.007)
Famine Germans	0.070^a	-0.045^a	0.175^a	-0.280^a	-0.420^a	-8.702^a	-1.093^a	-0.245^a
	(0.004)	(0.012)	(0.012)	(0.012)	(0.027)	(0.637)	(0.031)	(0.020)
Observations	449,022	449,022	449,019	449,019	449,022	449,022	449,022	223,804
R-squared	0.007	0.008	0.055	0.054	0.040	0.024	0.007	0.025
F-Test	12.677^a	15.246^a	38.107^a	27.447^a	37.245^a	21.291^a	7.099^a	1.992

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 D.5(b): ABE-Exact 5 Year Band

Variables	(1) Literacy	(2) Numeracy	(3) Unskill	(4) Farmer	(5) Occ. Wealth	(6) Occ. Rank	(7) Property	(8) School
Pre-Famine Irish	-0.051^a	-0.174^a	0.288^a	-0.344^a	-0.555^a	-10.884^a	-0.701^a	0.051^a
	(0.004)	(0.005)	(0.005)	(0.004)	(0.012)	(0.295)	(0.058)	(0.007)
Famine Irish	-0.111^a	-0.257^a	0.564^a	-0.514^a	-1.192^a	-25.012^a	-1.232^a	-0.137^a
	(0.012)	(0.014)	(0.013)	(0.008)	(0.026)	(0.591)	(0.038)	(0.022)
Pre-Famine British	0.054^a	-0.033^a	0.049^a	-0.273^a	-0.111^a	-1.391^a	-0.414^a	0.064^a
	(0.002)	(0.005)	(0.004)	(0.005)	(0.010)	(0.292)	(0.050)	(0.007)
Famine British	0.045^a	-0.002	0.186^a	-0.401^a	-0.490^a	-11.682^a	-1.023^a	-0.103^a
	(0.006)	(0.012)	(0.013)	(0.011)	(0.025)	(0.662)	(0.042)	(0.023)
Pre-Famine Germans	0.064^a	-0.031^a	0.070^a	-0.236^a	-0.135^a	-1.244^a	-0.728^a	-0.025^a
	(0.002)	(0.005)	(0.005)	(0.006)	(0.012)	(0.317)	(0.049)	(0.008)
Famine Germans	0.071^a	-0.044^a	0.180^a	-0.281^a	-0.429^a	-8.950^a	-1.138^a	-0.237^a
	(0.005)	(0.015)	(0.015)	(0.015)	(0.034)	(0.823)	(0.046)	(0.026)
Observations	425,845	425,845	425,839	425,839	425,845	425,845	425,845	213,178
R-squared	0.007	0.008	0.055	0.054	0.040	0.024	0.006	0.026
F-Test	12.613^a	16.798^a	39.333^a	22.188^a	35.232^a	18.583^a	3.953^b	0.778

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 D.6: Summary statistics for linked children, 1880, alternate matches

Table D.6(a): ABE-NYSIIS 5 Year Band

	Irish		British		Germans		Americans
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Pre-Famine	Famine	Pre-Famine	Famine	Pre-Famine	Famine	All
Labor market characteristics							
ln(Occ. Wealth)	7.660	7.289	7.906	7.721	7.841	7.757	7.827
	(1.041)	(1.043)	(0.968)	(0.969)	(1.018)	(0.989)	(0.978)
Occ. Rank	39.914	30.809	46.038	40.640	44.517	41.664	43.905
	(27.713)	(25.951)	(27.014)	(26.274)	(27.103)	(26.757)	(25.055)
White collar	0.203	0.155	0.226	0.183	0.204	0.205	0.158
	(0.402)	(0.362)	(0.418)	(0.387)	(0.403)	(0.404)	(0.365)
Craft	0.300	0.295	0.260	0.353	0.275	0.337	0.167
	(0.458)	(0.456)	(0.439)	(0.478)	(0.447)	(0.473)	(0.373)
Farmer	0.231	0.126	0.332	0.243	0.332	0.262	0.491
	(0.422)	(0.332)	(0.471)	(0.429)	(0.471)	(0.440)	(0.500)
Unskilled	0.235	0.408	0.154	0.195	0.161	0.173	0.157
	(0.424)	(0.492)	(0.361)	(0.396)	(0.367)	(0.378)	(0.363)
Place of residence							
Moved state	0.376	0.679	0.378	0.626	0.324	0.667	0.340
	(0.484)	(0.467)	(0.485)	(0.484)	(0.468)	(0.471)	(0.474)
Observations	9,649	1,186	9,274	1,430	9,490	1,470	404,024

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 D.6(b): ABE-Exact 5 Year Band

	Irish		British		Germans		Americans
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Variable	Pre-Famine	Famine	Pre-Famine	Famine	Pre-Famine	Famine	All
Labor market characteristics							
ln(Occ. Wealth)	7.661	7.254	7.914	7.736	7.847	7.681	7.827
	(1.050)	(1.055)	(0.979)	(0.932)	(0.999)	(1.037)	(0.978)
Occ. Rank	39.967	30.151	46.449	40.683	44.448	40.363	43.892
	(27.820)	(25.827)	(27.375)	(26.099)	(26.955)	(27.176)	(25.034)
White collar	0.206	0.137	0.232	0.183	0.204	0.191	0.158
	(0.405)	(0.344)	(0.422)	(0.387)	(0.403)	(0.394)	(0.365)
Craft	0.294	0.277	0.261	0.356	0.275	0.317	0.167
	(0.456)	(0.448)	(0.439)	(0.479)	(0.446)	(0.465)	(0.373)
Farmer	0.232	0.134	0.328	0.240	0.335	0.266	0.492
	(0.422)	(0.341)	(0.469)	(0.427)	(0.472)	(0.442)	(0.500)
Unskilled	0.235	0.430	0.153	0.195	0.159	0.213	0.156
	(0.424)	(0.495)	(0.360)	(0.396)	(0.365)	(0.410)	(0.363)
Place of residence							
Moved state	0.377	0.726	0.376	0.629	0.316	0.632	0.338
	(0.485)	(0.446)	(0.484)	(0.483)	(0.465)	(0.482)	(0.473)
Observations	9,647	1,370	9,387	1,375	7,307	960	383,939

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 D.7: Regressions of 1880 characteristics, alternate matches

Table D.7(a): ABE-NYSIIS 5 Year Band

Variables	(1) Unskill	(2) Occ. Wealth	(3) Occ. Rank
Pre-Famine Irish	0.079^a	-0.170^a	-4.073^a
	(0.005)	(0.012)	(0.311)
Famine Irish	0.240^a	-0.500^a	-12.270^a
	(0.015)	(0.031)	(0.780)
Pre-Famine British	0.000	0.068^a	1.887^a
	(0.004)	(0.011)	(0.299)
Famine British	0.027^b	-0.068^b	-2.444^a
	(0.011)	(0.027)	(0.712)
Pre-Famine Germans	0.002	0.019^c	0.717^b
	(0.004)	(0.011)	(0.289)
Famine Germans	0.006	-0.035	-1.477^b
	(0.010)	(0.027)	(0.717)
Observations	449,015	436,530	436,530
R-squared	0.011	0.014	0.010
F-Test	36.362^a	20.007^a	14.127^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 D.7(b): ABE-Exact 5 Year Band

Variables	(1) Unskill	(2) Occ. Wealth	(3) Occ. Rank
Pre-Famine Irish	0.080^a	-0.169^a	-3.986^a
	(0.005)	(0.011)	(0.302)
Famine Irish	0.264^a	-0.537^a	-12.950^a
	(0.014)	(0.029)	(0.706)
Pre-Famine British	-0.001	0.077^a	2.326^a
	(0.004)	(0.011)	(0.299)
Famine British	0.028^b	-0.054^b	-2.388^a
	(0.011)	(0.026)	(0.717)
Pre-Famine Germans	0.001	0.025^b	0.666^b
	(0.004)	(0.012)	(0.326)
Famine Germans	0.047^a	-0.111^a	-2.762^a
	(0.013)	(0.034)	(0.891)
Observations	425,837	413,993	413,993
R-squared	0.012	0.014	0.011
F-Test	39.421^a	19.206^a	12.577^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 D.8: Correlates of upward mobility among the famine Irish, alternate matches

Table D.8(a): ABE-NYSIIS 5 Year Band

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Head Literacy (1850)	−0.019	0.146	−0.017	0.145
	(0.043)	(0.089)	(0.044)	(0.089)
Native	−0.025	0.177	−0.014	0.173
	(0.070)	(0.144)	(0.070)	(0.145)
Famine in utero	0.077	−0.106	0.067	−0.103
	(0.056)	(0.118)	(0.057)	(0.118)
Irish Name Catholicity	0.107^{c}	−0.389^{a}	0.099^{c}	−0.386^{a}
	(0.056)	(0.117)	(0.056)	(0.118)
Fraction Irish in 100	−0.010	−0.124	−0.011	−0.124
	(0.078)	(0.159)	(0.078)	(0.159)
Moved State			0.062^{c}	−0.020
			(0.037)	(0.076)
Observations	917	898	917	898
R-squared	0.036	0.044	0.040	0.044

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.

Table D.8(b): ABE-Exact 5 Year Band

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Head Literacy (1850)	−0.055	0.146^{b}	−0.051	0.143^{c}
	(0.037)	(0.073)	(0.037)	(0.073)
Native	−0.116^{c}	0.297^{b}	−0.107^{c}	0.288^{b}
	(0.061)	(0.134)	(0.062)	(0.134)
Famine in utero	0.049	−0.090	0.038	−0.079
	(0.049)	(0.104)	(0.049)	(0.105)
Irish Name Catholicity	0.167^{a}	−0.439^{a}	0.158^{a}	−0.430^{a}
	(0.047)	(0.103)	(0.047)	(0.104)
Fraction Irish in 100	0.026	−0.076	0.032	−0.084
	(0.064)	(0.131)	(0.064)	(0.131)
Moved State			0.067^{b}	−0.063
			(0.033)	(0.066)
Observations	1,249	1,220	1,249	1,220
R-squared	0.036	0.047	0.039	0.048

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.

Table D.9: Coefficients underlying Figure D.3, alternate matches

Table D.9(a): ABE-NYSIIS 5 Year Band

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Pre-Famine Irish	0.023[a]	−0.089[a]	0.019[a]	−0.077[a]
	(0.005)	(0.012)	(0.005)	(0.012)
Famine Irish	0.138[a]	−0.283[a]	0.082[a]	−0.142[a]
	(0.015)	(0.032)	(0.017)	(0.037)
Pre-Famine British	−0.013[a]	0.027[a]	−0.019[a]	0.042[a]
	(0.004)	(0.011)	(0.004)	(0.011)
Famine British	−0.016	−0.024	−0.074[a]	0.120[a]
	(0.011)	(0.027)	(0.014)	(0.034)
Pre-Famine Germans	−0.012[a]	−0.006	−0.018[a]	0.010
	(0.004)	(0.011)	(0.004)	(0.011)
Famine Germans	−0.024[b]	−0.022	−0.081[a]	0.120[a]
	(0.011)	(0.027)	(0.014)	(0.033)
Native			−0.065[a]	0.163[a]
			(0.010)	(0.023)
Observations	449,012	436,527	449,012	436,527
R-squared	0.048	0.091	0.049	0.091
F-Test	25.341[a]	8.921[a]	25.099[a]	8.705[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.

Table D.9(b): ABE-Exact 5 Year Band

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Pre-Famine Irish	0.026[a]	−0.095[a]	0.021[a]	−0.081[a]
	(0.005)	(0.011)	(0.005)	(0.011)
Famine Irish	0.164[a]	−0.327[a]	0.100[a]	−0.164[a]
	(0.014)	(0.030)	(0.016)	(0.035)
Pre-Famine British	−0.013[a]	0.033[a]	−0.020[a]	0.050[a]
	(0.004)	(0.011)	(0.004)	(0.011)
Famine British	−0.015	−0.012	−0.081[a]	0.154[a]
	(0.011)	(0.026)	(0.014)	(0.033)
Pre-Famine Germans	−0.012[a]	−0.005	−0.019[a]	0.012
	(0.004)	(0.012)	(0.005)	(0.012)
Famine Germans	0.017	−0.102[a]	−0.047[a]	0.060
	(0.013)	(0.034)	(0.016)	(0.040)
Native			−0.074[a]	0.187[a]
			(0.010)	(0.023)
Observations	425,831	413,987	425,831	413,987
R-squared	0.050	0.091	0.050	0.091
F-Test	30.531[a]	10.551[a]	30.406[a]	10.401[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.

Appendix E: Results with Exact Matches (For Online Publication)

To verify that our results are not the products of false matches in linkage, we repeat the main results using a subset of our main dataset that is limited only to matches whose quality suggests that they are less likely to be false positives. Specifically, these matches, in addition to meeting the criteria described in Appendix A, must also satisfy the following conditions:

1. The difference in age-implied birth years across the two censuses is not more than one year.
2. The first and last names across the two records are identical, except for the potential removal of double letters or the use of common name abbreviations (e.g., "Wm." is considered an exact match for "William" and "Terrence" is considered an exact match for "Terence").

Imposing this restriction limits the sample to 274 sons of famine-era Irish, 3,387 sons of pre-famine Irish, 574 sons of famine-era British, 4,243 sons of pre-famine British, 245 sons of famine-era Germans, 3,083 sons of pre-famine Germans, and 195,550 sons of the native-born.

Figures E.1 and E.2, and Tables E.1-E.3 present results analogous to the main results for the sample limited to these exact matches. Figure E.1 and Table E.1 show the position of the famine-era Irish relative to, earlier Irish immigrants, the other major immigrant groups, and natives in 1850. In the exact-match sample, the famine-era Irish heads of household exhibit a 52.8 percentage point greater likelihood than native heads of household to hold an unskilled occupation, as compared to 56.5 percentage points in the complete sample. As in the main sample, the *F*-test supports the view that the difference between the outcomes of the pre-famine Irish and the famine-era Irish is greater than similar differences for the other ethnicities. The differences in the other characteristics are also similar to those obtained using the full sample.

Similarly, Figure E.1 and Table E.2 show the position of the children of the famine-era Irish relative to the children of natives, the children of earlier Irish immigrants, and the children of other immigrant groups as adults in 1880. In the exact-match sample, the famine-era Irish children were 16.4 percentage points more likely to hold an unskilled occupation than the children of natives, as compared to a difference of 23.7 percentage points in the full sample. This result of substantial but incomplete convergence of the famine-era Irish relative to natives, which is also present in the full sample, is also present in columns (2) and (3) of Table E.2, where the other occupational status measures are used.

The main difference that emerges between the results with the exact matches and the results with the full sample is that the degree to which the gap between natives and the famine-era Irish closed over generations appears larger when using only the exact matches.[50] Figure E.2 and Table E.3 further support this notion. Column (1) of Table E.3 shows that the children of the famine-era Irish in the exact-match sample were only 6.4 percentage points more likely to hold an unskilled occupation than were the children of natives. Column (3) shows that this gap falls to 4.6 percentage points and is not statistically significant when further controlling for nativity. The statistical insignificance of this difference is likely the product of the smaller sample available when restricting to exact matches. But it is also notable and not surprising that the assimilation appears greater in the restricted sample, which, because of the requirement that information be consistently reported (a marker for human capital), is likely to be disproportionately composed of those with unobservable characteristics suited for labor market success.

Given the qualitative similarity of the results when limiting the sample to only exact matches to the main results, we conclude that our results are unlikely to be the products of false matches in linkage between the 1850 and 1880 censuses.

[50] That is, the point estimates for the difference between the famine Irish and natives are smaller in 1880, whereas they are similar in 1850. Naturally, the smaller sample sizes also render these estimates less precise.

Figure E.1: Comparison of adults' and children's outcomes, exact matches

Figure E.1(a): Unskilled Occupation

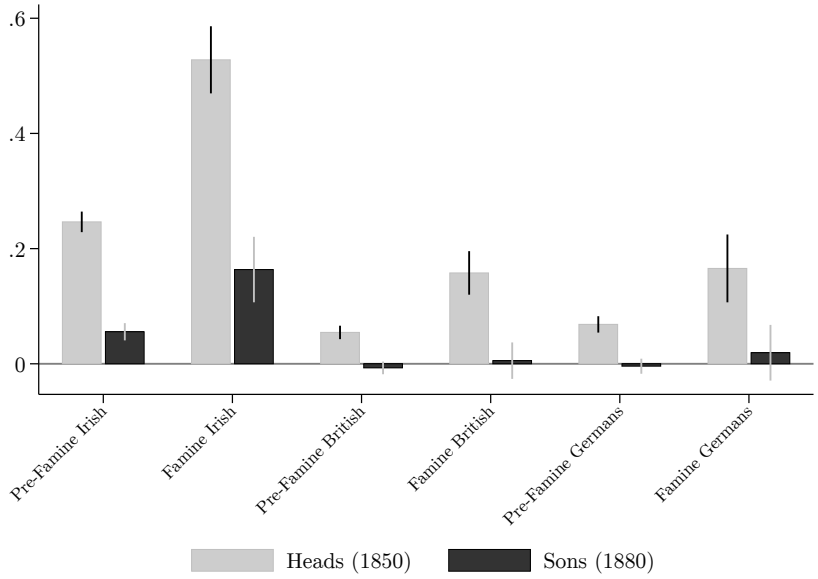
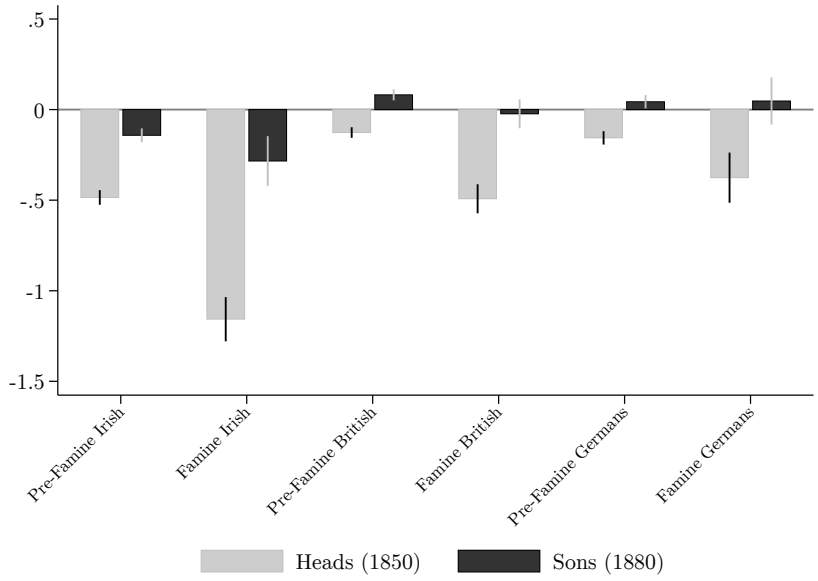


Figure E.1(b): Occupational Wealth Score



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

Figure E.2: Unconditional and conditional coefficients for 1880 characteristics, exact matches

Figure E.2(a): Unskilled Occupation

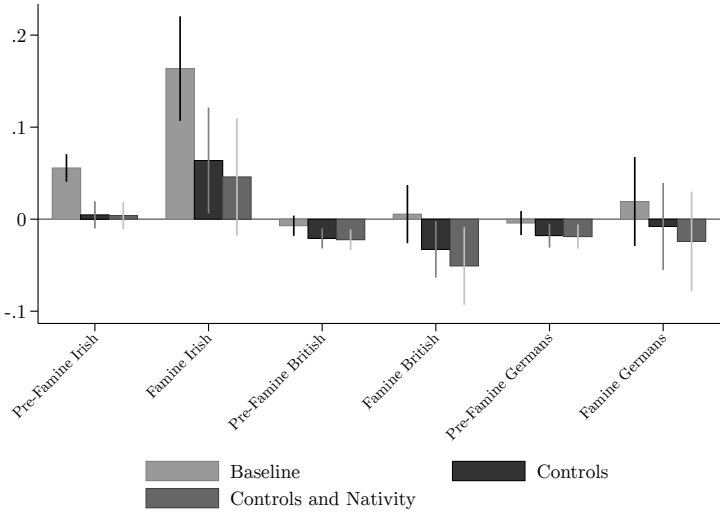
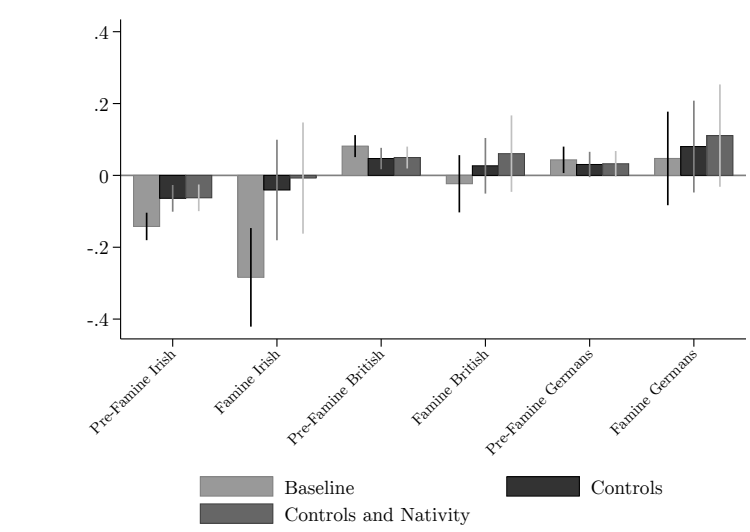


Figure E.2(b): Occupational Wealth Score



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 E.2. 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 E.4. 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 E.4.

Table E.1: Regressions of 1850 characteristics, exact matches

Variables	(1) Literacy	(2) Numeracy	(3) Unskill	(4) Farmer	(5) Occ. Wealth	(6) Occ. Rank	(7) Property	(8) School
Pre-Famine Irish	-0.045^a	-0.156^a	0.246^a	-0.300^a	-0.485^a	-9.329^a	-0.681^a	0.019
	(0.006)	(0.009)	(0.009)	(0.008)	(0.021)	(0.531)	(0.089)	(0.011)
Famine Irish	-0.117^a	-0.214^a	0.528^a	-0.509^a	-1.157^a	-23.845^a	-1.298^a	-0.159^a
	(0.027)	(0.031)	(0.030)	(0.018)	(0.062)	(1.412)	(0.055)	(0.050)
Pre-Famine British	0.037^a	-0.034^a	0.054^a	-0.252^a	-0.127^a	-1.799^a	-0.443^a	0.051^a
	(0.003)	(0.007)	(0.006)	(0.008)	(0.015)	(0.438)	(0.076)	(0.009)
Famine British	0.045^a	0.005	0.158^a	-0.388^a	-0.492^a	-11.313^a	-1.043^a	-0.106^a
	(0.009)	(0.018)	(0.019)	(0.018)	(0.041)	(1.080)	(0.051)	(0.035)
Pre-Famine Germans	0.052^a	-0.022^a	0.068^a	-0.245^a	-0.156^a	-1.671^a	-0.678^a	-0.047^a
	(0.003)	(0.008)	(0.007)	(0.009)	(0.019)	(0.500)	(0.082)	(0.012)
Famine Germans	0.067^a	-0.035	0.166^a	-0.280^a	-0.376^a	-7.324^a	-1.079^a	-0.188^a
	(0.007)	(0.030)	(0.030)	(0.030)	(0.071)	(1.641)	(0.113)	(0.051)
Observations	197,080	197,080	197,080	197,080	197,080	197,080	197,080	98,536
R-squared	0.005	0.004	0.033	0.039	0.027	0.016	0.007	0.023
F-Test	4.597^b	3.492^b	12.963^a	11.418^a	11.919^a	7.796^a	0.997	0.125

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 E.2: Regressions of 1880 characteristics, exact matches

Variables	(1) Unskill	(2) Occ. Wealth	(3) Occ. Rank
Pre-Famine Irish	0.056^a	-0.142^a	-3.478^a
	(0.008)	(0.019)	(0.527)
Famine Irish	0.164^a	-0.284^a	-6.281^a
	(0.029)	(0.070)	(1.874)
Pre-Famine British	-0.007	0.082^a	2.197^a
	(0.006)	(0.016)	(0.438)
Famine British	0.005	-0.023	-1.353
	(0.016)	(0.041)	(1.139)
Pre-Famine Germans	-0.004	0.043^b	1.326^a
	(0.007)	(0.019)	(0.514)
Famine Germans	0.019	0.047	1.530
	(0.025)	(0.066)	(1.898)
Observations	197,080	191,644	191,644
R-squared	0.008	0.011	0.008
F-Test	3.921^b	1.238	1.329

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 E.3: Coefficients underlying Figure D.2

Variables	(1) Unskill	(2) Occ. Wealth	(3) Unskill	(4) Occ. Wealth
Pre-Famine Irish	0.005	-0.064^{a}	0.004	-0.062^{a}
	(0.008)	(0.019)	(0.008)	(0.019)
Famine Irish	0.064^{b}	-0.041	0.046	-0.007
	(0.029)	(0.071)	(0.032)	(0.079)
Pre-Famine British	-0.021^{a}	0.047^{a}	-0.022^{a}	0.050^{a}
	(0.006)	(0.015)	(0.006)	(0.016)
Famine British	-0.033^{b}	0.027	-0.051^{b}	0.061
	(0.016)	(0.039)	(0.022)	(0.054)
Pre-Famine Germans	-0.018^{a}	0.030^{c}	-0.019^{a}	0.032^{c}
	(0.007)	(0.018)	(0.007)	(0.018)
Famine Germans	-0.008	0.080	-0.024	0.111
	(0.024)	(0.065)	(0.028)	(0.073)
Native			-0.021	0.039
			(0.017)	(0.042)
Observations	197,080	191,644	197,080	191,644
R-squared	0.045	0.097	0.045	0.097
F-Test	2.143	0.431	2.123	0.408

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.

Appendix F: Alternate Occupational Wealth Score Definitions (For Online Publication)

This appendix tests the sensitivity of the occupational wealth results to extreme outliers of the wealth distribution in the 1870 one-percent sample. Our benchmark measure, which we use in our analysis in the main text, assigns each occupation a score based on average wealth holdings of the occupation (or a broader class to which it belongs if there are too few individuals in a particular occupation) among men ages 30-65 in the 1870 one-percent sample. To determine whether our results are sensitive to outliers of the wealth distribution, we construct four other occupational wealth scores. For the first three, we compute average occupational wealth after topcoding all wealth observations at either the 99th, 95th, or 90th percentile of the wealth distribution in the 1870 one-percent sample. For the fourth, we use median wealth by occupation rather than the mean, without topcoding. We study all measures in logs rather than levels. Because there are some occupations (including large classes, such as laborer) where median wealth is zero, we use the log of median occupational wealth plus one instead of the log of median occupational wealth alone.

In Table F.1, we show that all five of the measures that we consider are highly correlated with one another for both the household heads of our linked sons in 1850 and the linked sons in 1880. The smallest coefficient of correlation is 0.896 in 1850 between our base occupational wealth score using means without topcoding, and the occupational wealth score using medians, also without topcoding. These correlations are even higher when we use ranks of these scores, as shown in Table F.2. Thus, it is unlikely that outliers of the wealth distribution have qualitatively affected our results.

To more directly determine the impact on our results of changing the way in which we define the occupational wealth score, Tables F.3-F.8 repeat the main results of the paper. Column (1) of each table uses the benchmark occupational wealth measure, thus reproducing the results of the main text of the paper. Columns (2)-(5) of each table then repeat the results with the alternate occupational wealth scores. The results with the topcoded mean scores are quantitatively and qualitatively similar to those with the benchmark measure. For the median-based scores, the results are qualitatively similar to those with the benchmark measure, showing considerable intergenerational assimilation by the famine Irish with upgrading at rates similar to those of observationally similar natives. The results with the occupational rank based on the median are also quantitatively similar to the benchmark occupational rank results.

Given these results, we are confident that our results are not driven by extreme outliers of wealth used to construct the occupational wealth scores.

Table F.1: Correlations between the various occupational wealth measures

	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Panel A: 1850 Household Heads					
Base	1.000				
99th	0.994	1.000			
95th	0.978	0.994	1.000		
90th	0.969	0.989	0.999	1.000	
Median	0.896	0.917	0.926	0.930	1.000
Panel B: 1880 Sons					
Base	1.000				
99th	0.993	1.000			
95th	0.979	0.995	1.000		
90th	0.969	0.989	0.999	1.000	
Median	0.907	0.928	0.939	0.942	1.000

Notes: Each cell presents the pairwise correlation coefficient between the logs of the various occupational wealth score measures, weighted to correct for selection into linkage. Base refers to the scoring described in the text, which uses the mean wealth. 99th, 95th, and 90th refer to means topcoding wealth at the 99th, 95th, or 90th percentiles, respectively. Median uses median wealth instead of mean wealth; the logarithm in this case is of median wealth plus one.

Table F.2: Correlations between the various occupational rank measures

	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Panel A: 1850 Household Heads					
Base	1.000				
99th	0.987	1.000			
95th	0.973	0.986	1.000		
90th	0.969	0.983	0.998	1.000	
Median	0.945	0.960	0.976	0.980	1.000
Panel B: 1880 Sons					
Base	1.000				
99th	0.988	1.000			
95th	0.979	0.991	1.000		
90th	0.975	0.987	0.997	1.000	
Median	0.943	0.957	0.969	0.971	1.000

Notes: Each cell presents the pairwise correlation coefficient between the ranks of the logs of the various occupational wealth score measures, weighted to correct for selection into linkage. Base refers to the scoring described in the text, which uses the mean wealth. 99th, 95th, and 90th refer to means topcoding wealth at the 99th, 95th, or 90th percentiles, respectively. Median uses median wealth instead of mean wealth; the logarithm in this case is of median wealth plus one.

Table F.3: Regressions of 1850 occupational wealth scores

Variables	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Pre-Famine Irish	-0.555^a	-0.556^a	-0.533^a	-0.508^a	-2.099^a
	(0.012)	(0.011)	(0.011)	(0.010)	(0.038)
Famine Irish	-1.191^a	-1.168^a	-1.103^a	-1.046^a	-4.208^a
	(0.026)	(0.025)	(0.023)	(0.022)	(0.086)
Pre-Famine British	-0.113^a	-0.139^a	-0.147^a	-0.145^a	-0.630^a
	(0.010)	(0.010)	(0.009)	(0.009)	(0.031)
Famine British	-0.495^a	-0.533^a	-0.523^a	-0.502^a	-1.988^a
	(0.024)	(0.024)	(0.023)	(0.022)	(0.086)
Pre-Famine Germans	-0.141^a	-0.149^a	-0.148^a	-0.145^a	-0.635^a
	(0.011)	(0.010)	(0.010)	(0.009)	(0.031)
Famine Germans	-0.424^a	-0.417^a	-0.393^a	-0.375^a	-1.488^a
	(0.025)	(0.025)	(0.023)	(0.022)	(0.080)
Observations	410,278	410,278	410,278	410,278	410,278
R-squared	0.039	0.043	0.044	0.045	0.061
F-Test	42.866^a	42.037^a	42.189^a	41.980^a	48.997^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. Base refers to the standard occupational wealth scoring. 99th, 95th, and 90th refer to occupational wealth scoring top coding wealth at the 99th, 95th, and 90th percentiles, respectively. Median refers to median occupational wealth scoring.

Table F.4: Regressions of 1850 occupational wealth ranks

Variables	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Pre-Famine Irish	-10.798^a	-10.927^a	-10.769^a	-11.025^a	-12.060^a
	(0.306)	(0.306)	(0.304)	(0.306)	(0.310)
Famine Irish	-24.568^a	-24.829^a	-24.690^a	-25.051^a	-26.082^a
	(0.588)	(0.587)	(0.572)	(0.576)	(0.571)
Pre-Famine British	-1.390^a	-1.623^a	-1.958^a	-2.227^a	-3.512^a
	(0.305)	(0.302)	(0.295)	(0.298)	(0.303)
Famine British	-11.639^a	-11.931^a	-12.303^a	-12.764^a	-14.644^a
	(0.631)	(0.637)	(0.614)	(0.622)	(0.649)
Pre-Famine Germans	-1.351^a	-1.193^a	-1.219^a	-1.392^a	-2.842^a
	(0.279)	(0.280)	(0.278)	(0.280)	(0.278)
Famine Germans	-8.899^a	-8.598^a	-8.593^a	-8.860^a	-9.704^a
	(0.592)	(0.606)	(0.601)	(0.610)	(0.618)
Observations	410,278	410,278	410,278	410,278	410,278
R-squared	0.024	0.024	0.025	0.025	0.029
F-Test	22.584^a	24.146^a	25.312^a	24.832^a	29.518^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. Base refers to the standard occupational wealth scoring. 99th, 95th, and 90th refer to occupational wealth scoring top coding wealth at the 99th, 95th, and 90th percentiles, respectively. Median refers to median occupational wealth scoring.

Table F.5: Regressions of 1880 occupational wealth scores

Variables	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Pre-Famine Irish	-0.170^a	-0.188^a	-0.191^a	-0.187^a	-0.750^a
	(0.012)	(0.011)	(0.010)	(0.009)	(0.033)
Famine Irish	-0.478^a	-0.491^a	-0.478^a	-0.460^a	-1.823^a
	(0.031)	(0.029)	(0.027)	(0.025)	(0.093)
Pre-Famine British	0.065^a	0.043^a	0.026^a	0.019^b	-0.079^a
	(0.011)	(0.011)	(0.010)	(0.009)	(0.030)
Famine British	-0.035	-0.073^a	-0.085^a	-0.088^a	-0.491^a
	(0.025)	(0.024)	(0.023)	(0.021)	(0.077)
Pre-Famine Germans	0.036^a	0.022^b	0.011	0.006	-0.116^a
	(0.011)	(0.010)	(0.009)	(0.009)	(0.030)
Famine Germans	-0.093^a	-0.114^a	-0.123^a	-0.124^a	-0.552^a
	(0.026)	(0.024)	(0.022)	(0.021)	(0.073)
Observations	398,591	398,591	398,591	398,591	398,591
R-squared	0.017	0.018	0.019	0.020	0.021
F-Test	13.252^a	12.200^a	12.393^a	12.680^a	16.281^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. Base refers to the standard occupational wealth scoring. 99th, 95th, and 90th refer to occupational wealth scoring top coding wealth at the 99th, 95th, and 90th percentiles, respectively. Median refers to median occupational wealth scoring.

Table F.6: Regressions of 1880 occupational wealth ranks

Variables	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Pre-Famine Irish	-4.239^a	-4.325^a	-4.574^a	-4.788^a	-6.378^a
	(0.307)	(0.306)	(0.303)	(0.302)	(0.321)
Famine Irish	-11.075^a	-11.330^a	-11.609^a	-11.822^a	-14.290^a
	(0.756)	(0.760)	(0.753)	(0.752)	(0.806)
Pre-Famine British	1.891^a	1.908^a	1.714^a	1.433^a	0.386
	(0.309)	(0.308)	(0.305)	(0.304)	(0.323)
Famine British	-1.893^a	-2.095^a	-2.375^a	-2.711^a	-4.518^a
	(0.689)	(0.696)	(0.689)	(0.687)	(0.754)
Pre-Famine Germans	1.022^a	1.347^a	1.297^a	1.055^a	0.081
	(0.287)	(0.288)	(0.289)	(0.287)	(0.303)
Famine Germans	-2.356^a	-2.054^a	-2.260^a	-2.469^a	-3.750^a
	(0.685)	(0.695)	(0.691)	(0.690)	(0.732)
Observations	398,591	398,591	398,591	398,591	398,591
R-squared	0.014	0.014	0.015	0.015	0.017
F-Test	5.729^a	5.906^a	5.652^a	5.682^a	6.380^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. Base refers to the standard occupational wealth scoring. 99th, 95th, and 90th refer to occupational wealth scoring top coding wealth at the 99th, 95th, and 90th percentiles, respectively. Median refers to median occupational wealth scoring.

Table F.7: Regressions of 1880 occupational wealth scores with controls

Variables	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Pre-Famine Irish	-0.095^a	-0.101^a	-0.101^a	-0.097^a	-0.324^a
	(0.012)	(0.011)	(0.010)	(0.009)	(0.033)
Famine Irish	-0.273^a	-0.275^a	-0.267^a	-0.257^a	-1.040^a
	(0.031)	(0.030)	(0.027)	(0.026)	(0.094)
Pre-Famine British	0.022^b	0.015	0.009	0.006	0.033
	(0.011)	(0.010)	(0.009)	(0.009)	(0.030)
Famine British	0.008	-0.014	-0.022	-0.024	-0.129^c
	(0.025)	(0.024)	(0.022)	(0.021)	(0.076)
Pre-Famine Germans	0.007	0.001	-0.003	-0.005	0.001
	(0.011)	(0.010)	(0.009)	(0.009)	(0.029)
Famine Germans	-0.084^a	-0.098^a	-0.105^a	-0.104^a	-0.334^a
	(0.027)	(0.025)	(0.023)	(0.021)	(0.074)
Observations	398,591	398,591	398,591	398,591	398,591
R-squared	0.091	0.094	0.100	0.102	0.066
F-Test	7.389^a	6.547^a	6.743^a	6.945^a	9.554^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. Base refers to the standard occupational wealth scoring. 99th, 95th, and 90th refer to occupational wealth scoring top coding wealth at the 99th, 95th, and 90th percentiles, respectively. Median refers to median occupational wealth scoring.

Table F.8: Regressions of 1880 occupational wealth scores with controls and nativity

Variables	(1) Base	(2) 99th	(3) 95th	(4) 90th	(5) Median
Pre-Famine Irish	-0.084^a	-0.089^a	-0.090^a	-0.087^a	-0.286^a
	(0.012)	(0.011)	(0.010)	(0.009)	(0.033)
Famine Irish	-0.122^a	-0.122^a	-0.120^a	-0.117^a	-0.546^a
	(0.036)	(0.034)	(0.031)	(0.030)	(0.108)
Pre-Famine British	0.036^a	0.029^a	0.023^b	0.020^b	0.079^a
	(0.011)	(0.010)	(0.010)	(0.009)	(0.031)
Famine British	0.161^a	0.142^a	0.128^a	0.118^a	0.372^a
	(0.032)	(0.031)	(0.028)	(0.027)	(0.097)
Pre-Famine Germans	0.023^b	0.017^c	0.012	0.010	0.053^c
	(0.011)	(0.010)	(0.009)	(0.009)	(0.030)
Famine Germans	0.066^b	0.055^c	0.042	0.036	0.159^c
	(0.033)	(0.031)	(0.028)	(0.026)	(0.093)
Native	0.173^a	0.176^a	0.169^a	0.161^a	0.567^a
	(0.022)	(0.021)	(0.019)	(0.018)	(0.065)
Observations	398,591	398,591	398,591	398,591	398,591
R-squared	0.092	0.095	0.101	0.102	0.067
F-Test	7.388^a	6.547^a	6.762^a	6.970^a	9.447^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. Base refers to the standard occupational wealth scoring. 99th, 95th, and 90th refer to occupational wealth scoring top coding wealth at the 99th, 95th, and 90th percentiles, respectively. Median refers to median occupational wealth scoring.

Appendix G: Surname-Based Catholicity Imputation Algorithm (For Online Publication)

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 "Catholicity" 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.

References (For Online Publication)

Abramitzky, Ran, Leah Platt Boustan, Katherine Eriksson, James J. Feigenbaum, and Santiago Pérez (2019a). "Automated Linking of Historical Data." NBER Working Paper 25825.

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.