Immigrants and the Making of America

SANDRA SEQUEIRA

London School of Economics and CEPR

NATHAN NUNN

Harvard University, NBER and BREAD

and

NANCY QIAN

Northwestern Kellogg MEDS, Fanhai International School of Finance Fudan University, NBER, CEPR, BREAD

We study the effects of European immigration to the U.S. during the Age of Mass Migration (1850–1920) on economic prosperity. Exploiting cross-county variation in immigration that arises from the interaction of fluctuations in aggregate immigrant flows and of the gradual expansion of the railway network, we find that counties with more historical immigration have higher income, less poverty, less unemployment, higher rates of urbanization, and greater educational attainment today. The long-run effects seem to capture the persistence of short-run benefits, including greater industrialization, increased agricultural productivity, and more innovation.

Key words: Immigration, Historical persistence, Economic development.

JEL Codes: B52; F22; N72; O10; O40.

1. INTRODUCTION

An important issue within current American political discourse is the effect that immigrants have on the communities in which they settle. While this topic has received significant attention, the focus has generally been on the short-term effects of immigrants.[1] We know much less about the long-run consequences of immigration. This is particularly important because the short-run and long-run effects could be very different, in both magnitude and sign.

We contribute to an improved understanding of the long-run effects of immigration by taking a historical perspective and studying the effects of immigration into the U.S. during the Age of Mass Migration (1850–1920). This wave of immigration is notable because it is the period of U.S.

---

1. Immigrants have been found to positively affect entrepreneurial activity (Kerr and Kerr, 2017), productivity (Peri, 2012), occupational specialization (Peri and Sparber, 2009), innovation (Hunt and Gauthier-Loiselle, 2010), and wages (Card, 2012).

The editor in charge of this paper was Nicola Gennaioli.

history with the highest levels of immigration and also because the new arrivals were quite different from previous immigrants. While prior immigrants were primarily from Western Europe, the new wave also included large numbers of immigrants from Southern, Northern, and Eastern Europe (Daniels, 2002, pp. 121–137; Hatton and Williamson, 2005, p. 51; Abramitzky and Boustan, 2017).

Empirically studying the long-run effects of immigration is challenging. A natural strategy is to examine the relationship between historical immigration and current economic outcomes across counties in the U.S. However, such an exercise has important shortcomings. Given the historical evidence, one is particularly concerned about negative selection. Immigrants may have only been able to settle in more marginal locations, where land and rents were cheaper and the potential for future growth was lower. Given the historical accounts of congestion and discrimination that kept immigrants from well-paying, attractive jobs and occupations (Handlin, 1957; McGouldrick and Tannen, 1977; Blau, 1980; Hannon, 1982), this form of selection, which would cause ordinary least squares (OLS) estimates of the long-run benefit of immigrants to be biased downwards, is likely to have been particularly important. By contrast, immigrants were also attracted to places with economic opportunity, which may have been locations with more long-run growth potential. This would cause OLS estimates to be biased upwards. Lastly, classical measurement error in the immigration data would cause the OLS estimates to be biased towards zero.

An important contribution of our analysis is the implementation of an empirical strategy that overcomes these identification problems. We use an instrumental variable (IV) strategy that exploits two facts about immigration during this period. The first is that after arriving in the U.S., immigrants tended to use the railway to travel inland to their eventual places of residence (Faulkner, 1960; Foerster, 1969). Therefore, a county's connection to the railway network affected the number of immigrants that settled in the county. The second fact is that the aggregate inflow of immigrants coming to the U.S. during this period fluctuated greatly from one decade to the next. If a county was connected to the railway network during periods of high aggregate immigration to the U.S., then the county tended to receive more immigrants. The benefit of combining the two sources of variation—the timing of railway construction and the timing of immigration booms—is that the interaction between the two produces variation that is unlikely to affect our contemporary outcomes of interest other than through historical immigration to the county.

Our analysis proceeds in three steps. First, to help understand the intuition behind our instrument, we begin with a "zero-stage" regression where we examine a panel of counties every census decade from 1850 to 1920, and estimate the determinants of the share of the population that was foreign born.[2] The specification includes county fixed effects and time-period fixed effects, as well as our interaction of interest, which is between the aggregate inflow of European immigrants into the U.S. (normalized by total U.S. population) during the prior ten years and an indicator variable that equals one if the county was connected to the railway network at the beginning of the ten-year period. This interaction captures the differential effect of connection to the railway network on immigrant settlement in decades with high aggregate immigrant inflows relative to decades with low aggregate immigrant inflows. This is the underlying variation of our instrument.

We find that the interaction term is a strong predictor of the settlement of foreign immigrants into a county. Counties experienced more immigrant settlement if they were connected to the railway network *and* the aggregate flow of immigrants into the country was high at the time. In addition, the coefficient of the uninteracted railway indicator is very close to zero, which suggests

---

2. As we explain in more detail below, while the zero-stage is not necessary to construct the instrument, it is useful to provide an intuition for the instrument and to assess its plausibility.

that connection to the railway would have no effect on immigrant settlement if there was no aggregate inflow of immigrants to the U.S. This is reassuring since it provides evidence that the estimates of the effect of railway access on immigrant settlement is unlikely to capture other mechanisms.

Secondly, we begin the long-run analysis by estimating the share of the population that was foreign born (for each county and decade), which is predicted using the interaction term only. Following the same intuition as in the zero-stage analysis, the only variation that we interpret as exogenous is the differential effect of being connected to the railway during an aggregate immigration boom versus being connected during an aggregate immigration lull. This yields a predicted immigrant share for each county and decade. We then calculate the average across decades from 1860 to 1920 to construct a measure of the average predicted immigrant share.

Lastly, we estimate the cross-county relationship between average historical immigrant share (from 1860–1920) and economic outcomes today using the predicted immigrant share as an instrument for the actual immigrant share.

One concern with our identification strategy is that the interaction of connection to the railway network and aggregate immigrant inflows might be correlated with how early a county became connected to the railway. To address this, we always control for a measure of how early the county became connected to the railway. Another potential concern with our estimation strategy is that decades with high aggregate immigration inflows may have been different in other ways. For example, if high levels of aggregate immigration coincided with high levels of industrial development or movements in the business cycle, then our estimates will be biased. Given such concerns, our zero-stage specification includes two additional interaction terms: the interaction of the railway connection indicator and an index of aggregate industrialization and the interaction of the railway connection indicator and the decadal change in real per-capita GDP. These control for differential effects of railway connection that depend on industrialization or changes in the business cycle. Following the same procedure as with our instrument, we create two measures of predicted immigration using each interaction term and control for them in all specifications.

Another potential concern is the possibility that the aggregate flow of immigrants could have been endogenous to railway expansion. If immigrant inflows tended to increase once the railway became connected to counties with a greater future growth potential, then our instrument would suffer from reverse causality and be invalid. As a robustness check, we construct a measure of the predicted flow of European emigrants to the U.S. that is determined solely by weather shocks in the origin countries. We find that predicted immigrant flows are strongly correlated with actual flows, and that using the predicted values yields estimates that are qualitatively identical to our baseline estimates.

Our main findings show that historical immigration resulted in significantly higher incomes, less poverty, less unemployment, more urbanization, and higher educational attainment today. The estimates, in addition to being statistically significant, are also economically meaningful. For example, they indicate that moving a county with no historical immigration to the 50th percentile of the sample (which is 0.049) results in a 13% increase in average per capita income today. We find no evidence that historical immigration affects social cohesion as measured by social capital, voter turnout, or crime rates. Consistent with historical accounts of congestion and discrimination leading to negative selection in immigrant settlement, we find that the two stage least squares (2SLS) estimates are often larger than the OLS estimates.

We then turn to an examination of mechanisms and examine whether the economic gains enjoyed by counties that received more immigrants appear to come at the expense of other nearby counties that received fewer immigrants. We do this by testing for the presence of spillover effects. If our main findings are due to the relocation of economic activity, we may find that immigration

to a location has negative effects in nearby regions. The estimates provide no evidence for such negative spillover effects.

Another way to shed light on mechanisms is to ask when the economic benefits of immigrants began to emerge. It is possible that in the short-run, immigrants were a burden on the economy and the benefits they brought were only felt in the medium- or long-runs. The estimates show that immigration resulted in benefits that were felt soon after their arrival. Immigration resulted in more and larger manufacturing establishments, greater agricultural productivity, and higher rates of innovation. These findings are consistent with a long-standing narrative in the historical literature suggesting that immigrants contributed to economic growth by providing an ample supply of unskilled labour, which was crucial for early industrialization, as well as a smaller, but also important, supply of skilled individuals, who brought with them knowledge, skills, and innovations that were particularly important for industrial development.[3]

The results of our article improve our understanding of the short- and long-run effects of immigration in the U.S. We find that in the long-run, immigration provides large economic benefits. At the same time, we see no systematic effects on long-run social outcomes. It is informative to compare these long-run effects with the estimated short-run effects from Tabellini (2018), who studies the effects of immigration from 1910 to 1930 across 180 U.S. cities. He finds, as we do, that immigration led to immediate economic benefits. He also finds that immigration had adverse political and social consequences due to an anti-immigrant backlash. Our findings suggest that, although the short-run economic benefits of historical immigration appear to have persisted until today, the short-run social costs appear to have died out. Thus, although there may have been social costs that coincided with the economic benefits of immigrants in the short-run, this does not appear to be the case in the long-run.

Our findings complement recent scholarship examining the selection of immigrants to the U.S. (Abramitzky *et al.*, 2012, 2013; Spitzer and Zimran, 2013) and their experiences after arrival (Abramitzky *et al.*, 2014), as well as the existing literature on the importance of effects of immigration that arise due to culture, genetics, or networks (Fischer, 1989; Ottaviano and Peri, 2006; Ager and Bruckner, 2013; Grosjean, 2014; Burchardi and Hassan, 2015; Bandiera *et al.*, 2016; Ager and Bruckner, 2017). They also complement existing studies that find long-term benefits of historical immigration outside of the U.S., for example, in Brazil (Rocha *et al.*, 2017), Argentina (Droller, 2017), and Prussia (Hornung, 2014).

Our article examines the effect of immigrants in general and not the different effects of immigrants from different countries, which has been the focus of some lines of research (Fischer, 1989; Burchardi and Hassan, 2015; Fulford *et al.*, 2017). In theory, our identification strategy could be used to instrument separately for immigrants from different countries.[4] However, in practice, the large number of origin countries (and thus endogenous variables and instruments) results in first-stage estimates that are weak and often counterintuitive.[5]

The article is structured as follows. We begin with a description of the historical setting, which is followed, in Sections 3 and 4, by overviews of our data and identification strategy. In Section 5, we report our baseline estimates and in Section 6, we turn to an examination of mechanisms,

---

3. On average, immigrants appear to have been less educated than native-born populations. We find that, consistent with this, immigration is associated with lower levels of education in the short-run (prior to 1920), but higher levels in the medium- and long-run (1950 and later).

4. Following the same logic as for all immigrants, one could construct instruments based on the interaction of the aggregate inflow of immigrants from a sending-country and a county's connection to the railway network at that time.

5. One has sixteen endogenous immigrant-share variables (and instruments), one for each sending-country for which we have data. We find that our first-stage estimates tend to be weak and country-specific immigration often loads on the instruments for other counties. This is most likely due to the collinearity that is present in the endogenous variables and the instruments.

estimating the short- and medium-run effects of immigration. In Section 7, we report the findings from a range of robustness and sensitivity checks before concluding in Section 8.

2. HISTORICAL BACKGROUND

2.1. *Immigration and the railway*

Historical immigration into the U.S. peaked during the Age of Mass Migration. The new immigrants were different from earlier ones. While previous waves were primarily from Western Europe, the new wave included large numbers of immigrants from Southern, Northern, and Eastern Europe who spoke different languages and had different religious practices (Daniels, 2002, pp. 121–137; Hatton and Williamson, 2005, p. 51). In 1850, at the onset of the Age of Mass Migration, over 90% of the foreign born living in the U.S. were from either Great Britain, Ireland, or Germany. By the end of the Age of Mass Migration, in 1920, this figure was only 45% (Abramitzky and Boustan, 2017, p. 7).

Throughout this period, immigration was facilitated by the railways. The best land was often granted to railway companies by the Federal government in an attempt to promote the development of uninhabited territories. The railway companies, including the Union Pacific, Santa Fe, Burlington, Northern Pacific, through a variety of mechanisms, intentionally promoted the settlement of these tracts of land contiguous to their railway lines (Luebke, 1977, p. 410). They did so by selling the land cheaply, subsidizing trans-Atlantic travel, and through information campaigns run through advertising offices in Europe. Upon arrival in the U.S., railroads were the primary means of transportation to the interior (Hedges, 1926). As argued by historian James Hedges (1926, p. 312), the settlement of the Western U.S. is a story of journeying "not with wagon and ox-teams but in the drab passenger coaches of early western railroads. It is the story of… [immigrants] who sought new homes where the railroads led them".

2.2. *Why immigrants matter in both the short- and long-run*

There are several reasons why immigration during America's Age of Mass Migration may have mattered in both the short- and long-runs. The contributions of immigrants are nicely summarized by John F. Kennedy in his book, *A Nation of Immigrants*, where he writes: "Between 1880 and 1920 America became the industrial and agricultural giant of the world … This could not have been done without the hard labour, the technical skills and entrepreneurial ability of the 23.5 million people who came to America in this period" (Kennedy, 1964, p. 34). We discuss each of these potential contributions of immigration below.

Provision of unskilled labour: Immigrants may have spurred industrialization by providing a large supply of unskilled labour to newly established factories. The role of immigration may have been particularly important in manufacturing, which is often characterized by agglomeration effects arising from increasing returns to scale or network externalities. Immigrants from England, Ireland, and Germany also often had previous experience in the industrial cities in their homelands (Bergquist, 2007, pp. 264–265). Many have hypothesized that the rapid increase in industrialization in the U.S. was fuelled by such immigrant labour. For example, Foerster (1924, p. 331) writes that "the sixfold increase in the capital invested in manufactures between the outbreak of the Civil War and the year 1890, a period in which the population in the country doubled, was largely made possible by the inpouring immigrants".

Evidence that immigration resulted in cheaper labour costs—that is, low wages—has been put forth by Goldin (1994). Examining variation across American cities from 1890 to 1903, she finds that greater immigration was associated with lower wage growth: a one-percentage-point

increase in the foreign-born population is associated with a decrease in wages of about 1.0–1.5%. Interestingly, these effects are found both for unskilled labourers and artisans.

Provision of important skills for industry: Although most immigrants worked in unskilled occupations, a disproportionate share appears to have engaged in more specialized and skilled occupations. Malone (1935) reports that among the noteworthy and exceptional individuals summarized in the Volume fifteen *Dictionary of American Biography*, 12.5% of those born after 1790 were foreign born, which is higher than the national proportion of foreigners (10.1% in our sample). Abramitzky *et al.* (2014) examine the occupational distribution of immigrants and natives in 1900 and find that immigrants, relative to natives, were equally likely to be in unskilled occupations, less likely to be in farming, and more likely to hold semi-skilled or skilled blue collar occupations.

Immigrants from particular countries were especially represented in a range of skilled occupations, such as carpentry, cabinetmaking, blacksmithing, brewing, distilling, clockmaking, etc. For example, in 1870, 37% of German-born workers were employed in skilled occupations (Daniels, 2002, p. 150). Genoese Italians, with their tradition in the commercial trades, opened saloons and restaurants and also went into confectionary and fresh fruit businesses, while Jewish immigrants specialized in retail and other professional occupations (Bergquist, 2007, p. 195).

Provision of agricultural know-how: Immigrants, who represented a small but important proportion of farm operators (15.3% in 1900 and 10.5% in 1920), are widely believed to have had knowledge of superior farming practices, which contributed to productivity improvements within agriculture (Cance, 1925, pp. 102–113). The most notable group of immigrant farmers were the Germans, who were also the largest immigrant group within the farming sector, accounting for 25% of all foreign-born farm operators in 1920 (Cance, 1925, p. 113). German immigrants have been credited with adopting, perfecting, and popularizing new crops and better livestock. They were the first to breed the Congesta horse, to introduce alfalfa seed, and to adopt increased rotation and diversification of crops (Kollmorgen, 1942, pp. 53–54; Saloutos, 1976, p. 66). Immigrants were also particularly known for their innovations (Jordan, 1966). Although empirical evidence is limited, the few studies of the role of immigrants in agricultural innovation show that they were significantly over-represented among agricultural innovators (Gripshover and Bell, 2012).

Provision of knowledge and innovation: Immigrants contributed directly to the U.S. economy through important technological innovations. One example of such an innovation is the suspension bridge, which was pioneered by John A. Roebling, a German-born and trained civil engineer, who built numerous suspension bridges, including the Niagara Fall Suspension Bridge and the Brooklyn Bridge (Faust, 1916, p. 10). Other notable immigrant engineers include Charles Conrad Schneider (born in Saxony), who constructed the famous cantilever bridge across the Niagara River in 1883; the Austrian Gustav Lindenthal, who built the Hell Gate Bridge; and John F. O'Rourke, an Irish engineer, who built seven of the tunnels under the East and Hudson Rivers, and six of the tunnels of the New York subway systems (Wittke, 1939, pp. 389–390).[6]

Many important inventions were developed by immigrants. A notable example is Alexander Graham Bell, who was born in Scotland in 1847 and moved to Boston in 1871. In 1876, he developed an acoustic telegraph that could transmit voices and sounds telegraphically and, within a year, established the Bell Telephone company. Other notable inventors include David Thomas (Welsh), who invented the hot blast furnace; John Ericsson (Swedish), who invented the ironclad

---

6. The historical importance of engineers for economic development is highlighted by the recent findings from Maloney and Caceido (2017) which shows that engineers were critically important for long-run economic growth.

ship and the screw propeller; Conrad Hubert (Russian), who invented the flashlight; and Ottmar Mergenthaler (German), who invented the linotype machine (Kennedy, 1964, pp. 33–34).

Immigrants also made important contributions to the educational system of the U.S. For example, the State University system was modelled after the Prussian system (Faust, 1916, pp. 10–11). The concept of kindergarten, which has been shown to have had important economic effects, was brought to the U.S. by German immigrant Friederich Fröbel (Paz, 2015; Ager *et al.*, 2016).

3. DATA

Our zero-stage estimation uses a panel of counties and census decades from 1860 to 1920.[7] The key variables of the analysis are a measure of whether a county was connected to the railway network in each decade and a measure of the inflow of immigrants into the U.S. during each year.

Data on a county's historical connectivity to the railway network were constructed using thirty-eight historical maps, fifteen with national coverage, and twenty-three with regional coverage only.[8] Using these, we digitized the estimated railway network for each decade from 1830 to 1920.[9] To construct the digitized railway network, we first obtained an accurate and geo-referenced shape file of the current railway network from the U.S. Department of Transportation.[10] We then laid the modern shapefile over a digitized version of a paper map of the most recent historical time period of interest: 1920. We then proceeded to remove all railway lines that exist today but did not exist in 1920. We repeated this for each earlier time period in sequence—that is, 1910, 1900, etc—at each point removing railway lines that did not exist in the previous decade. This procedure ensures the greatest precision in digitizing the exact location of the railway lines. Because of mapping imprecisions from the original historical maps, simply tracing the lines from each paper map would have generated inaccurate maps of historical railway networks. In instances where railway lines existed at some point in the past, but are not in the modern shapefile, the historical railway lines were drawn using the geo-referenced paper maps.[11] As a measure of whether a county was connected to the railway network, we use an indicator variable that equals one if a county's boundary is intersected by at least one railway line.[12]

The second key measure used in our analysis is the aggregate inflow of immigrants into the U.S. We measure this using data from Willcox (1929, pp. 377–393) on the total number of immigrants from Europe who were admitted to the U.S. each year from 1820 to 1920.[13] The figures are from passengers lists provided by the masters of arriving vessels. Although the data from the lists are

---

7. Although 1860 is the first year of our panel, we measure the presence of the railway one decade prior. Therefore, 1850 is the earliest period of railway data that we use in our analysis. It is the decade in which the census started to consistently record whether an individual was foreign-born. The census data were obtained through the Natural Historical Geographic Information System (NHGIS), which is available at www.nhgis.org (see Minnesota Population Center, 2011), and the Inter-university Consortium for Political and Social Research (ICPSR), which is available at www.icpsr.umich.edu (see Haines and Inter-university Consortium for Political and Social Research, 2010).

8. See Online Appendix Table A1 for a summary of the maps that were used.

9. One source of imprecision arises because the years for which the railway maps are available does not match the Census year exactly. In some, cases there is a 1- to 3-years discrepancy.

10. The shapefile used is the 2009 version of the National Transportation Atlas Railroads (NTAR), which is at a 1:100,000 scale.

11. Figures A1–A11 of the Online Appendix show, for time periods from 1850 to 1920, the digitized and geo-referenced railway network overlaid on the original paper maps from which the data were obtained.

12. The proportion of connected counties steadily increased over time from just under 20% in 1850 to over 90% in 1920 (see Online Appendix Figure A12).

13. We use Willcox (1929) rather than the already-digitized data available from Migration Policy Institute (2016) because Willcox (1929) reports immigrants by sending country, while Migration Policy Institute (2016) does not. This information is necessary to construct predicted immigration inflows that are due to sending country weather shocks.
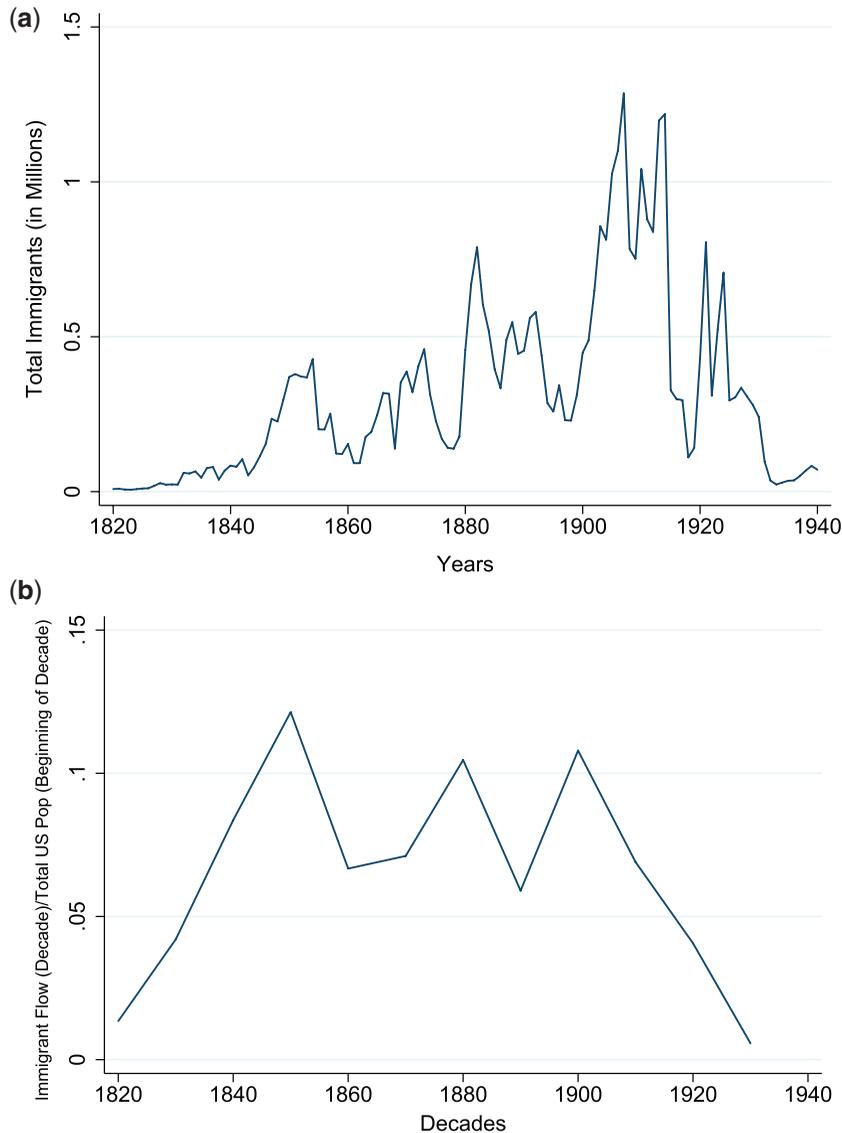
FIGURE 1

Immigration into the U.S. during the Age of Mass Migration. (a) Annual flow of immigrants to the U.S., 1820–1940.
Source: Migration Policy Institute. (b) Total flow of immigrants to the U.S. by decade normalized by total U.S.
population at the beginning of the decade, 1820–1939. *Source*: Willcox (1929), Tables 1–3 on pages 377–393, for the
immigration data and the U.S. Census for the population data.

believed to be reliable (they were reported quarterly to the Secretary of State as mandated by the
first immigration Act of 1819), there is some imprecision and variation over time in who was
defined as an immigrant.[14] For most of the time period, alien immigrants were defined as foreign
passengers who arrived in the country with the intention of staying. This measure double counts

14. See Willcox (1929, pp. 374–376) for a full discussion of measurement issues.

any immigrants who had already arrived in the U.S. but left and were again returning. Another source of imprecision is that prior to 1908, the figures omitted immigrants that arrived by land through Canada and Mexico.

Annual aggregate immigration inflows from 1820 to 1940 are shown in Figure 1a (Migration Policy Institute, 2016). It is clear from the figure that aggregate immigrant flows into the U.S. fluctuated significantly from year to year. Figure 1b shows that, even after aggregating flows to the decade level (which is the unit of our analysis) and normalizing by the total U.S. population at the beginning of the decade, one still observes significant variation over time.[15] This volatility, combined with the expansion of the railway network, is the variation that lies at the heart of our analysis.

4. EMPIRICAL STRATEGY

4.1. *Estimating equations*

Our empirical strategy exploits two facts about immigration during the period 1850–1920. First, the total inflow of immigrants fluctuated greatly across decades (as seen in Figures 1a and 1b). Secondly, the arriving immigrants tended to use the railway to travel inland to their eventual place of residence (Faulkner, 1960; Foerster, 1969). Therefore, during the period of railway construction, the timing of a county's connection to the railway network in relation to the aggregate inflow of immigrants affected the number of immigrants that settled in the county.

To verify and better understand this source of variation, our analysis begins by estimating the following zero-stage equation:

$$\text{Immigrant Share}_{i,t} = \alpha_t + \alpha_i + \gamma \, \text{Immigrant Share}_{i,t-1} + \delta I_{i,t-1}^{\text{RR Access}}$$
$$+ \beta \, \text{Immigrant Flow}_{t-1} \times I_{i,t-1}^{\text{RR Access}} + \theta \, \text{Industrialization}_{t-1} \times I_{i,t-1}^{\text{RR Access}}$$
$$+ \phi \, \text{GDP Growth}_{t-1} \times I_{i,t-1}^{\text{RR Access}} + \mathbf{X}_{i,t-1} \Gamma + \varepsilon_{i,t}, \tag{1}$$

where i indexes counties and t indexes census years (1860, 1870, 1880, 1890, 1900, 1910, 1920); α_t denotes decade fixed effects and α_i county fixed effects. The outcome of interest, Immigrant Share$_{it}$, is the share of the population in county i that is foreign born during census year t. Immigrant Share$_{i,t-1}$ denotes a one-decade lag of the dependent variable, which captures the mechanical relationship between the previous decade's population of immigrants and this decade's population of immigrants.[16] Immigrant Flow$_{t-1}$ is the total number of European immigrants arriving in the U.S. during decade t, normalized by the total U.S. population at the beginning of that decade. For example, if $t = 1860$, then Immigrant Flow$_{t-1}$ measures all immigrants arriving from 1850 to 1859 normalized by total population in 1850. $I_{i,t-1}^{\text{RR Access}}$ is an indicator variable that equals one if county i is connected to the railway network in decade $t-1$. For example, if $t = 1860$, then $I_{i,t-1}^{\text{RR Access}}$ is an indicator variable for connection in 1850.

Core to our identification strategy is the interaction between the aggregate flow of immigrants into the U.S. and whether a county was connected to the railway network: Immigrant Flow$_{t-1} \times$

---

15. The figure reports total immigrant flows during a decade and normalized by the total U.S. population at the beginning of the decade. Flows reported in decade t refer to flows during that year and the nine years that follow. For example, 1820 in the figure refers to aggregate flows from 1820 to 1829, which are normalized by total population in 1820. We use this convention unless stated otherwise. The population data are taken from the U.S. Census.

16. Due to the presence of a Nickell bias, there is concern that the estimate of γ may be biased, which could bias our estimate of β. As we discuss below, the estimates of equation (1) are nearly identical without the inclusion of a lagged dependent variable.

$I_{i,t-1}^{\text{RR Access}}$. The interaction captures the differential effect that connection to the railway had on immigrant settlement during periods of high aggregate immigration relative to periods of low aggregate immigration. Thus, we expect the estimate of β in equation (1) to be positive.

The two variables that comprise the interaction terms are also included in equation (1). The coefficient δ for the variable $I_{i,t-1}^{\text{RR Access}}$ reflects the estimated effect of access to the railway on immigrant settlement during a decade when there are no immigrants coming into the U.S. Thus, a test of the logic of our IV strategy is whether the estimate of δ is close to zero. The variable Immigrant Flow$_{t-1}$ is absorbed by the time period fixed effects, and, thus, does not appear in the equation.

Motivated by the possibility that the timing of connection to the railway may have a direct impact on long-term development through increased specialization and industrialization, we also allow the effect of railway connection to differ depending on the level of aggregate industrial development at the time: Industrialization$_{t-1} \times I_{i,t-1}^{\text{RR Access}}$.[17] Industrialization$_{t-1}$ is the annual average of the level of industrialization during the ten years prior to census year t.[18] The interaction term captures any differential effects that connection to the railway network has depending on the level of aggregate industrial development at the time.

Similarly, it is possible that the arrival of immigrants coincided with variation in the business cycle. Thus, we also allow the effect of railway connection to vary differentially depending on decadal growth in national GDP: GDP Growth$_{t-1} \times I_{i,t-1}^{\text{RR Access}}$.[19] This interaction term captures any differential effects that connection to the railway network may have had depending on the business cycle. Equation (1) also includes a vector of additional control variables, $\mathbf{X}_{i,t-1}$, that are intended to capture the potential influence that cities and more populous counties had in attracting immigrants: log population density, a one-decade lag of an urbanization indicator, and an interaction of the lagged urbanization indicator with the lagged aggregate immigrant flow variable.

To help understand the intuition behind our instrument, we undertake the following exercise. After estimating equation (1), we first calculate the immigrant share in each county and decade that is predicted by the interaction between the aggregate inflow of immigrants and connection to the railway network: $\widehat{\text{Immigrant Share}}_{i,t} = \widehat{\beta}\,\text{Immigrant Flow}_{t-1} \times I_{i,t-1}^{\text{RR Access}}$, where $\widehat{\beta}$ is the estimate of β from equation (1). We then average this county decade-specific predicted immigrant share over the seven census years from 1860 to 1920:

$$\text{Avg}\,\widehat{\text{Immigrant Share}}_i = \frac{1}{T}\sum_{t=1}^{T}\widehat{\beta}\,\text{Immigrant Flow}_{t-1} \times I_{i,t-1}^{\text{RR Access}},$$

where T is the total number of time periods.[20] Since some counties were still in the process of being formed during this period, our panel is unbalanced with counties entering over time.[21] When constructing Avg $\widehat{\text{Immigrant Share}}_i$, we use the average immigrant share for all census years from

17. The logged industrialization index closely approximates a linear time trend. Thus, the estimates are very similar if one uses the interaction between a linear time trend and the railroad access indicator rather than the industrialization index and the railway access interaction.

18. The level of industrialization is measured using the log of the annual industrial production index taken from Davis (2004). The data are shown in Online Appendix Figure A13.

19. The measure of GDP per capita is from Maddison (2001).

20. For counties that are present in all periods of our analysis, $T = 7$. For those that enter our sample after the initial period, the value of T is less than 7 and equals the number of time periods for which they are observed in our panel.

21. In 1860, there are 1,532 counties in our sample, there are 1,922 counties in 1870; 2,137 in 1880; 2,416 in 1890; 2,692 in 1900; 2,752 in 1910; and 2,935 in 1920.

1860 to 1920 for which the county is in existence. We implement our IV procedure using 2SLS, with Avg Immigrant $\widehat{\text{Share}}_i$ as an instrument for the actual average immigrant share from 1860 to 1920.

There are two important points to keep in mind about our instrument. The first is that, in the end, the estimate $\widehat{\beta}$ from equation (1) is effectively inconsequential since it simply scales $\frac{1}{T}\sum_{t=1}^{T}$ Immigrant Flow$_{t-1} \times I_{i,t-1}^{\text{RR Access}}$ by a constant. The second point is that, for counties that are present during the full sample period, $T = 7$, and Avg Immigrant $\widehat{\text{Share}}_i$ takes on seven potential values. The range of variation is not a problem per se; for example, an indicator variable, which takes on two values only, can serve as a valid instrument. The only concern is whether there is sufficiently-rich variation to estimate a strong first stage in the 2SLS estimates. As we will see, this is the case.

Our 2SLS equations are given by equations (2) and (3), where equation (2) is the first stage and equation (3) is the second stage.

$$\text{Avg Immigrant Share}_{i,s} = \zeta_s + \mu \text{Avg Immigrant } \widehat{\text{Share}}_{i,s} + \omega \text{RR Duration}_{i,s} + \mathbf{X}_{i,s}\Omega + \epsilon_{i,s} \quad (2)$$

$$Y_{i,s} = \xi_s + \psi \text{Avg Immigrant Share}_{i,s} + \pi \text{RR Duration}_{i,s} + \mathbf{X}_{i,s}\Pi + \nu_{i,s} \quad (3)$$

where i indexes counties and s states. $Y_{i,s}$ is a contemporary outcome of interest; for example current per capita income, poverty, unemployment, education, etc. These variables are generally measured in 2000. Avg Immigrant share$_{i,s}$ is the average immigrant share in county i in census years from 1860 to 1920; and Avg Immigrant $\widehat{\text{Share}}_{i,s}$ is the predicted average immigrant share described above.

In equations (2) and (3), ζ_s and ξ_s denote state fixed effects, which capture differences between counties due to, for example, geography or historical experience. RR Duration$_{i,s}$ is the log number of years, as of 2000, that a county has been connected to the railway network. The variable is included to address the possibility that our instrument may be correlated with early connection to the railway network, which could have an independent long-run effect on our outcomes of interest.

The vector $\mathbf{X}_{i,s}$ includes the remaining covariates. This includes a cubic polynomial in the latitude and longitude of each county's centroid, which controls flexibly for potential relationships between our instrument and a county's spatial orientation. The vector also includes a pair of regressors that are meant to account for the fact that the timing of connection to the railway could have affected long-term economic growth through other mechanisms. As the U.S. industrialized, counties that became connected to the railway network during certain periods (*e.g.* early industrialization or rapid economic growth) may have disproportionately benefitted from being connected to the railway. Thus, we also include two control variables that use the same logic and procedure as for the immigration instrument and that account for these differential historical effects. These controls capture any potential differential effects of the timing of the connection to the railway that is due to different levels of industrialization and different effects that are due to the business cycle (*e.g.*, economic growth): $\frac{1}{T}\sum_{t=1}^{T}\widehat{\theta}$ Industrialization$_{t-1} \times I_{i,t-1}^{\text{RR Access}}$ and $\frac{1}{T}\sum_{t=1}^{T}\widehat{\phi}$ GDP Growth$_{t-1} \times I_{i,t-1}^{\text{RR Access}}$, where T is the number of census years from 1860 to 1920 for which county i is in the sample, where $\hat{\theta}$ and $\hat{\phi}$ are the coefficient estimates from equation (1).[22]

---

22. As with the case for $\hat{\beta}$, the values of $\hat{\theta}$ and $\hat{\phi}$ are inconsequential since they only scale $\frac{1}{T}\sum_{t=1}^{T}$ Industrialization$_{t-1} \times I_{i,t-1}^{\text{RR Access}}$ and $\frac{1}{T}\sum_{t=1}^{T}\widehat{\phi}$ GDP Growth$_{t-1} \times I_{i,t-1}^{\text{RR Access}}$ by a constant.

4.2. *Identification*

Our IV strategy exploits the differential effect that a county's connection to the railway network has in decades with high aggregate immigration relative to decades with low aggregate immigration. During the period of analysis, once a county became connected to the railway network, it generally stayed connected. Therefore, whether a county was connected during periods with relatively high aggregate immigration is primarily determined by whether a county became connected to the railway network just prior to a decade with high aggregate immigration rather than just prior to a decade with low aggregate immigration. To illustrate this variation, Figure 2 presents examples of pairs of counties that are within the same state (recall that we control for state fixed effects), but became connected to the railway at different times. Within each pair, one county became connected just prior to a high-immigration decade (*i.e.* a boom) and the other became connected just prior to a low-immigration decade (*i.e.* a lull). Counties connected just prior to a boom decade (1850s, 1880s, 1900s, 1910s) are shaded dark and counties connected just prior to a lull decade (1860s, 1870s, and 1890s) are shaded light.[23] The figure also reports the subsequent average immigrant share for the census years from 1860 to 1920. The examples demonstrate how the exact timing of a county's connection to the railway network can have significant effects on the extent of subsequent immigration into a county.

An important question regarding the validity of our empirical strategy is the comparability of counties that were connected just prior to immigration booms and lulls. In Table 1, we compare baseline economic, demographic, and geographic characteristics that may be correlated with the timing of railroad construction, the settlement of immigrants, and our long-run outcomes of interest. We find that the two sets of counties appear similar in terms of initial characteristics. Panel A examines differences in the initial share of foreign-born in 1820 or 1830, panel B examines initial economic characteristics, and panel C examines geographic characteristics, namely whether a county is located in the Midwest/West, or in the South.

We also undertake an extended and more generalized balance test by repeating the same exercise for the decades following our initial time period—that is, 1850–90. For each decade, we consider the sample of counties that had not yet been connected to the railway, and then check for the balance of characteristics measured at that time between counties that subsequently became connected in a lull decade relative to those that subsequently were connected in a boom decade. Estimates of this more exhaustive balance test are reported in Online Appendix Table A2.[24] The two sets of counties appear balanced across observable characteristics. We find that for the vast majority of characteristics there is little to no significant difference between the two groups. Exceptions are for how early the railway was connected, and the geographic characteristics of counties. These differences underscore the importance of our inclusion of the date of connection to the railway network, state fixed effects, and a cubic polynomial for latitude and longitude as controls in our baseline 2SLS equations.

Another concern is that the aggregate inflow of immigrants may have been endogenous to railway expansion. For example, the inflow of immigrants may have increased when the railway became connected to counties with greater future growth potential. We address this concern by constructing a measure of aggregate immigrant inflows that is solely due to sending-country weather shocks. As we report in Section 7, this alternative procedure generates estimates that are very similar to our baseline estimates.

---

23. We code all counties that enter our panel being connected to the railway as obtaining connection in the 1850s. Thus, connection in the 1850s should be understood as being connected in the 1850s or earlier.

24. Balance statistics are reported for all variables for which data are available. For some variables and periods a comparison is not available due to missing data.

(a)

(b)
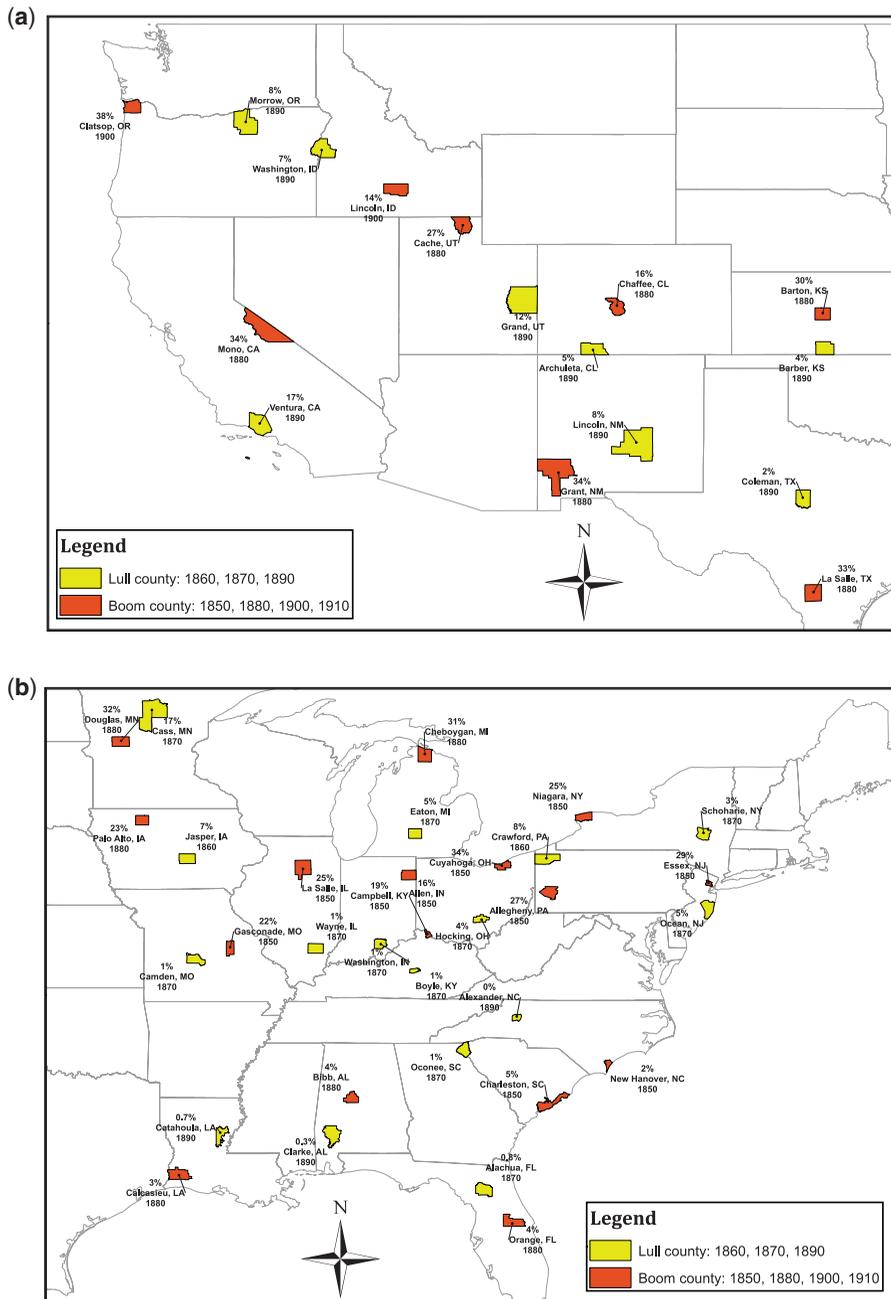
FIGURE 2

Illustration of the variation behind the identification strategy. Pairs of counties within the same state are shown. One county was connected just prior to an immigration boom and the other county was connected just prior to an immigration lull. Reported next to each county is the average immigration share from 1860–1920, the county name, and the first full decade in which the county was connected to the railway. (a) Map of the Western U.S. (b) Map of the Eastern U.S.

TABLE 1
Examining differences in baseline characteristics between lull- and boom-connection counties

Sample	(1)	(2)	(3)	(4)	(5)	(6)	(7)
	Boom-connection counties			Lull-connection counties			Equality of means
	Obs	Mean	Std Err	Obs	Mean	Std Err	*p*-value
Panel A: demographic composition							
Foreign share of the population, 1820	392	0.005	(0.0006)	312	0.004	(0.0005)	0.622
Foreign share of the population, 1830	524	0.004	(0.0005)	408	0.004	(0.0007)	0.482
Panel B: economic characteristics							
Urban share, 1840	626	0.975	(0.470)	496	0.695	(0.172)	0.575
Population density, 1830	670	0.147	(0.0480)	531	0.131	(0.022)	0.754
Share of the population in commerce, 1840	653	0.004	(0.0002)	509	0.005	(0.0003)	0.374
Share of the population in agriculture, 1840	653	0.259	(0.005)	509	0.260	(0.005)	0.916
Share of the population in mining, 1840	654	0.0009	(0.0002)	511	0.001	(0.0002)	0.582
Value of agricultural output per capita, 1840	663	46.332	(1.092)	527	44.25	(1.412)	0.244
Value of agricultural crops per capita, 1840	663	42.300	(1.076)	527	40.35	(1.404)	0.272
Post Offices per 1,000 inhabitants, 1840	672	0.698	(0.022)	536	0.652	(0.050)	0.403
Newspapers per 1,000 inhabitants 1840	242	0.175	(0.020)	120	0.112	(0.024)	0.048
Water connection indicator, 1840	670	0.467	(0.019)	531	0.514	(0.022)	0.106
Panel C: geographic characteristics							
Latitude	1,305	38.115	(0.125)	1,502	38.47	(0.124)	0.045
Longitude	1,305	−90.029	(0.306)	1,502	−92.164	(0.281)	0.000
Share of counties in the midwest and west	1,305	0.474	(0.014)	1,503	0.476	(0.013)	0.942
Share of counties in the south	1,305	0.436	(0.014)	1,503	0.460	(0.013)	0.195

Notes: "Boom-connection counties" are counties that we observe as connected to the railway for the first time in either 1850, 1880, or 1900. "Lull-connection counties" are counties that we observe as being connected to the railway for the first time in 1860, 1870, 1890, or 1910. Column 7 reports the *p*-value from a test of equality of means with unequal variances.

5. ESTIMATES

5.1. *Zero-stage analysis: verification of the sensibility of the instrument*

Estimates of the zero-stage equation (1) are reported in column 1 of Table 2. The reported standard errors are Conley standard errors that use a five-degree window.[25] The estimated coefficient for our interaction of interest—the railroad-access indicator multiplied by normalized aggregate immigrant inflows—is positive and statistically significant. This indicates that connection to the railway network did have a significant effect on immigrant settlement in a county.[26] The estimated coefficient for the uninteracted railway-connection indicator is close to zero, which means that connection to the railway has no effect on immigrant settlement when aggregate immigration flows are zero.

To further examine the variation underlying the interaction term, we estimate a more flexible version of equation (1), where we interact the indicator for whether a county had access to the railway network in the previous decade with decade fixed effects, instead of with the aggregate inflow of immigrants to the U.S. This allows the importance of being connected to the railway to vary flexibly over time. We then examine the relationship between the coefficients of each interaction term and the normalized aggregate inflow of immigrants during the previous decade.

25. The standard errors are very similar if we cluster at the county level.

26. This finding is consistent with relatively low levels of within-county mobility of immigrants after initial settlement. Examining the microdata from the 1940 Census, one finds that only 2% of immigrants report having moved states during the past five years, and only 3% report having moved across counties within the same state during the same time. These figures are about half the magnitude of those for native-born, which are 4% across states and 6% across counties within the same state.

TABLE 2
Zero-stage OLS panel estimates

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
	Immigrant share of total county population					
	All Counties	Excluding Northeast	Northeast	Excluding South	South	Midwest and West
Interaction of interest:						
Lag Rail Access	0.172***	0.183***	0.046	0.239***	0.051*	0.292***
x Lag Immigrant Inflow/ Total US Pop	[0.045]	[0.051]	[0.107]	[0.076]	[0.027]	[0.085]
Other variables:						
Lag Rail Access	−0.002	0.010	0.087***	−0.056	−0.007	−0.038
	[0.015]	[0.015]	[0.015]	[0.036]	[0.009]	[0.040]
Lag Rail Access	−0.005	−0.009	−0.042***	0.014	0.002	0.007
x Lag Log Industrialization Index	[0.006]	[0.006]	[0.006]	[0.014]	[0.004]	[0.016]
Lag Rail Access	−0.001	−0.008	0.022	−0.023	−0.010	−0.028
x Lag GDP Per Cap Decadal Growth	[0.010]	[0.010]	[0.027]	[0.027]	[0.008]	[0.027]
Lag Immigrant Share	Yes	Yes	Yes	Yes	Yes	Yes
Lag Urban Indicator	Yes	Yes	Yes	Yes	Yes	Yes
Lag Urban Indicator	Yes	Yes	Yes	Yes	Yes	Yes
x Lag Immigrant Inflow/ Total US Pop						
Log County Population Density	Yes	Yes	Yes	Yes	Yes	Yes
County Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Decade Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	16,386	14,903	1,483	8,612	7,774	7,129
R-squared (within)	0.408	0.408	0.652	0.463	0.414	0.495
Mean Dependent Variable	0.085	0.080	0.138	0.144	0.021	0.145
SD Dependent Variable	0.109	0.109	0.088	0.112	0.057	0.116

Notes: OLS estimates are reported. An observation is a county in a time period (1860, 1870, 1880, 1890, 1900, 1910 or 1920). The dependent variable "Immigrant Share of Total County Population" is the proportion of a county's population that is foreign born in period t. "Lag Rail Access" is an indicator variable that equals one if a county has a railway in period $t−1$. Coefficients are reported, with Conley standard errors in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Figure 3 shows a strong positive relationship between the two variables (the correlation coefficient is 0.66). The decades for which connection to the railway network in the previous decade had the largest effects on county-level immigrant settlement (1850, 1880, and 1900) are also the decades for which we observe the largest aggregate immigrant inflows.

While our baseline sample includes all counties, one could argue that the logic of our identification strategy applies less well (or not at all) to the Northeast of the U.S., where there are many urban centres located on the coast, travel distances are relatively short, and the railway network was already developed prior to the first period in our analysis. Thus, as a robustness check, we re-estimate equation (1), but omit from the sample counties located in the Northeast as defined by the Census.[27] The findings, which are reported in column 2 of Table 2, show that omitting these counties produces results that are nearly identical to our baseline estimates.

The characteristics of the Northeastern counties also provide an opportunity to test whether omitted factors are biasing our estimates. In particular, looking within the Northeast only, we should not observe the same effects as we do for the rest of the country. As reported in column 3 of Table 2, this is exactly what we find. The estimated coefficient for the interaction term is small in magnitude and not statistically different from zero.

A related question is the applicability of the model to the U.S. South, which received few immigrants during this time. In column 4, we report estimates for a sample that omits counties

27. These include: Connecticut, Maine, Massachusetts, New Hampshire, New Jersey, New York, Pennsylvania, Rhode Island, and Vermont.
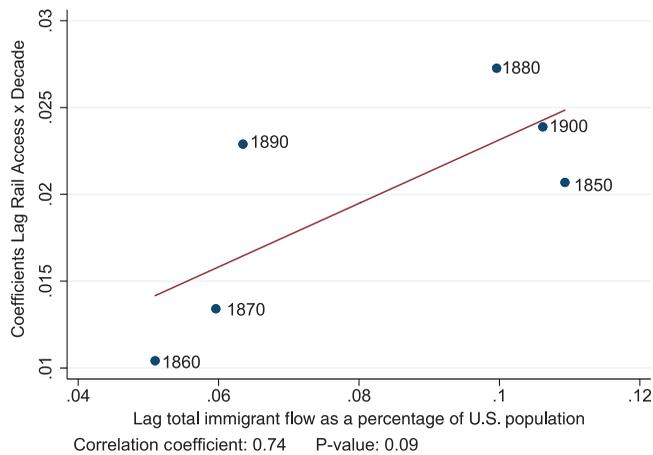
FIGURE 3

Bivariate relationship between the estimated effect of a county's connection to the railway on the subsequent share of foreign-born in a county and total immigration into the U.S.

in the South. We find that our estimates are similar. The magnitude of the point estimate of the interaction of interest increases slightly and remains statistically significant. Despite the fact that the South experienced less immigration than the rest of the country, as shown in column 5, we still find a positive and significant effect within the South, although one that is much smaller in magnitude. For completeness, in column 6, we also report estimates omitting counties from both the Northeast and the South—that is, the Midwest and West only. The estimates are similar to, although noticeably larger in magnitude than, the baseline estimates.

The last set of tests that we perform check the sensitivity of the zero-stage estimates. These are reported in Online Appendix Table A3, which reports estimates that omit all covariates (column 1),[28] omit the lagged dependent variable from the baseline specification (column 2) that includes region by decade fixed effects (column 3), and omit influential observations (column 4).[29] In all cases, we find that the estimates remain robust. The final check tests the robustness of the estimates to the omission of each time period of our sample. The estimates, which are reported in Online Appendix Table A5, show that our estimates remain similar with the exclusion of any of the decades in our sample.[30]

5.2. *The long-term effects of immigration on economic outcomes*

5.2.1. First-stage estimates. We now turn to our 2SLS estimation, starting with the first-stage estimates, which are are reported in Table 3. The baseline estimate using the full sample is reported in column 1. We find that the predicted immigrant share instrument is strongly correlated with actual immigrant share, with a Kleibergen-Paap F-statistic of approximately 21. Taken literally, the magnitude of the point estimate suggests that a 1 percentage point increase

28. Formal tests of coefficient stability are reported in panel A of Online Appendix Table A4.

29. Influential observations are defined as those with a Cook's distance greater than $4/N$, where N is the number of observations in the regression.

30. The results also indicate that the relationship of interest does not appear to be systematically stronger in earlier or later decades. We confirm this fact more formally by allowing a differential effect for the 1900–20 period (reported in column 5 of Online Appendix Table A3).

TABLE 3
First-stage relationship between the predicted and actual average immigrant shares

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
	Average Immigrant Share, 1860-1920					
	All Counties	Excluding northeast	northeast	Excluding south	South	Midwest and west
Predicted Average Immigrant Share, 1860–1920	4.559*** [1.311]	4.184*** [1.270]	−3.585 [9.271]	9.730*** [1.807]	1.303** [0.615]	8.903*** [1.891]
Controls :						
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,935	2,720	215	1,593	1,342	1,378
Kleibergen Paap *F*-statistic	21.22	23.98	0.07	36.07	4.29	29.87
Standardized 'beta' coeff. for Pred. Avg. Immig. Share	0.167	0.153	−0.095	0.311	0.099	0.285
Mean of Dependent Variable	0.098	0.095	0.137	0.160	0.024	0.163
SD of Dependent Variable	0.111	0.112	0.082	0.107	0.057	0.110

Notes: An observation is a county. Coefficient estimates are reported, with Conley standard errors reported in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

in the predicted immigrant share is associated with a 4.5 percentage point increase in the actual average immigrant share. While this coefficient appears large, it is important to recognize that the absolute magnitude of the coefficient is highly sensitive to the estimated value of β from the zero-stage equation (1). Since β scales our instrument, it does not materially affect our first-stage coefficient (or the 2SLS estimates) but it does affect its value. If, due to measurement error, the estimate of β from equation (1) is biased towards zero, then this will mechanically decrease the values of the calculated predicted average immigrant share and mechanically increase the magnitude of the instrument's coefficient in the first-stage. Given this, a preferred measure of the magnitude of the first-stage coefficient is the standardized 'beta' coefficient, which is modest with a value of 0.167.[31]

As with the zero-stage, we find that our estimates are robust to omitting the Northeast (column 2) or the South (column 4). When we examine the Northeast only, we find that the railway-based predicted immigrants share is not correlated with actual immigrant share (column 3), which is unsurprising and consistent with the weak zero-stage estimates for this region. Also consistent with the zero-stage estimates, we find that our instrument has predictive power for the U.S. South, although the relationship is weaker than for the full sample (column 5). For completeness, we also report estimates for the Midwest and West (column 6).

Online Appendix Figures A14 and A15 report binscatter plots that show the first-stage relationship between predicted immigrant share and the actual immigrant share, both with and without the baseline covariates. They both show that the relationship between predicted and actual immigration is not driven by a small set of observations.

5.2.2. 2SLS estimates.

The 2SLS estimates examining economic outcomes are reported in Table 4. Panel A reports OLS estimates for comparison, panel B reports reduced form estimates,

---

31. The statistic is the predicted effect on the dependent variable, measured in standard deviations, for a one-standard-deviation increase in the independent variable. Summary statistics of the key variables in our analyses are reported in Online Appendix Table A6.

TABLE 4

OLS and 2SLS estimates of the effects of historical immigration on economic prosperity today

Dependent variable	(1) Log average per capita income, 2000	(2) Share of Pop. below poverty line, 2000	(3) Unemployment rate, 2000	(4) Urbanization rate, 2000	(5) Average years of schooling, 2000
	A. OLS estimates				
Average Immigrant Share, 1860–1920	0.243* [0.130]	0.015 [0.028]	0.020 [0.015]	0.949*** [0.184]	0.020 [0.307]
	B. Reduced form				
Predicted Average Imigrant Share, 1860–1920	11.942*** [3.629]	−2.229*** [0.777]	−1.876*** [0.500]	22.382*** [6.820]	41.925*** [10.562]
	C. 2SLS estimates				
Average Immigrant Share, 1860–1920	2.619*** [1.022]	−0.489** [0.209]	−0.411*** [0.151]	4.909*** [2.008]	9.195*** [3.392]
	D. First stage estimates				
	Dependent variable: average immigrant share, 1860–1920				
Predicted Average Immigrant Share, 1860–1920	4.559*** [1.311]	4.559*** [1.311]	4.559*** [1.311]	4.559*** [1.311]	4.559*** [1.311]
Kleibergen Paap *F*-statistic	21.222	21.222	21.222	21.222	21.222
Controls (in all Panels):					
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	2,935	2,935	2,935	2,935	2,935
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	10.022	0.136	0.047	0.401	11.445
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.203	0.054	0.025	0.305	0.558

Notes: An observation is a county. Panel A reports OLS estimates, Panel B reports reduced-form estimates, Panel C reports 2SLS estimates, and Panel D reports first-stage estimates. Coefficient estimates are reported, with Conley standard errors in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

panel C reports the 2SLS estimates, and panel D reports the first-stage estimates.[32] According to the 2SLS estimates in panel C, counties with a greater share of immigrants from 1860 to 1920 have significantly higher average per capita income in 2000 (column 1).[33] The magnitude of the coefficient suggests that moving a county's average historical immigrant share from zero to the 50th percentile of the sample—a change of 0.049 or 4.9%—results in an increase in average income of $2.62 \times 0.049 = 0.128$ or 13%.[34]

32. Partial correlation plots of the reduced-form are reported in Online Appendix Figure A17.

33. Throughout the table, we report Conley standard errors adjusted for spatial correlation using a five-degree window. The results are very similar when we use smaller or larger windows, e.g. one degree or ten degrees.

34. In reporting magnitudes, we focus on the median rather than the mean because the distribution of average immigrant share is noticeably right skewed, with a large number of counties with very low levels of average immigrant share, and a small number of counties with high levels (see Online Appendix Figure A16). The mean of average immigrant share is 0.098 and the standard deviation is 0.111. The median is 0.049, the 25th percentile is 0.007, and the 75th percentile is 0.163.

We also examine a number of alternative measures of the economic prosperity: the proportion of the population living below the poverty line (column 2) and the unemployment rate (column 3). We estimate a negative effect of historical immigrant share on both poverty and unemployment. According to the estimates, moving a county with no historical immigration to the 50th percentile of the distribution (0.049) is associated with a decrease in the proportion of people living under the poverty line by 2 percentage points and a decrease in the unemployment rate by 2 percentage points. These findings are consistent with the long-run increase in income found in column 1. In columns 4 and 5, we consider two additional measures of economic development: the urbanization rate and average years of schooling. We estimate a large positive effect on both urbanization and education. An increase in average immigrant share from zero to the 50th percentile (0.049) is associated with a 24 percentage point increase in the urbanization rate and 0.45 additional years of schooling.

Overall, the estimates show that historical immigration had large positive effects on long-run economic growth and prosperity.

5.3. *Differences between OLS and 2SLS estimates*

Comparing the estimates of panels A and C in Table 4, it is clear that the OLS correlation between historical immigrant share and current per capita income is smaller than the 2SLS estimate. One explanation for this differences is that the 2SLS estimates are causal while the OLS estimates are not, with the difference between the two arising due to the negative selection by immigrants, which results in OLS estimates that understate the long-run economic benefits of immigrants. Negative selection would occur if immigrants tended to move to places that counterfactually would have had lower long-run economic growth. This is consistent with historical evidence suggesting that immigrants tended to settle in less desirable lower-income neighbourhoods and counties, for example, the immigrant tenements in New York City (Muller, 1993, pp. 74–75, 104–109). These were locations without the desirable amenities that are important for attracting labour, which can then create agglomeration benefits and lead to long-run growth.[35] Counterfactually, these places would not have been likely candidates for long-run economic prosperity. This is also consistent with evidence showing that immigrants were systematically excluded from attractive well-paying jobs, either due to direct discrimination, state legislation, language requirements, or union rules (Handlin, 1957; McGouldrick and Tannen, 1977; Blau, 1980; Hannon, 1982). Legislation in the mid-1890s in both New York and Pennsylvania excluded all foreign aliens from jobs in state and local municipal public works. Pennsylvania had residence and language requirements for all those who were foreign born, while Idaho legislation prevented companies from hiring aliens who had declared their intention to stay permanently in the U.S. (Higham, 2011, pp. 68–74, 158–165).

Another potential explanation for the difference in estimates is that the 2SLS estimates are biased due to a violation of the exclusion restriction. It is possible that the aggregate immigrant inflow into the U.S., which we take as given, was actually influenced by whether the railway network had recently become connected to counties with long-run growth potential. However, as we show in Section 77.1, when we use variation in immigrant inflow that is predicted by sending-country weather shocks—and therefore exogenous to factors within the U.S.—we obtain similar estimates to those reported in Table 4. If anything, we find that these estimates are actually slightly larger in magnitude (not smaller), which is evidence against the endogeneity of aggregate immigrant inflows causing bias in our 2SLS estimates.

A third potential reason for the difference between the two estimates is that 2SLS estimates a local average treatment effect (LATE), which is the average effect amongst compliers, which in

---

35. See Desmet *et al*. (2018) for theory and evidence of such a mechanism.

our setting are counties whose immigrant population was strongly affected by the presence of a railway. It is possible that because immigrants who settled in a location due to the railway were different than those who did not, the LATE is different than the average treatment effect (ATE). To gain some sense of the importance of selection as a potential explanation, we re-estimate the regressions of Table 4 by region. As shown in Online Appendix Tables A7 and A8, the estimates are noticeably different across regions. In comparing the three regions (Northeast, South, and Midwest and West), we expect selection to be strongest in the parts of the country that were already populated and economically developed; namely, the Northeast and South. By contrast, for the Midwest and West there is less economic variation that would induce selection. Thus, we expect the OLS estimates to be less biased for the Midwest and West than for the Northeast and South. We find that the OLS estimates for income are not statistically different from zero in both the Northeast and South, but positive and significant in the Midwest. This is consistent with the presence of negative selection that is stronger in the Northeast and South than in the Midwest and West.

5.4. *The long-term effects of immigration on social outcomes*

We now turn to an examination of the social effects of immigrants. It is possible that although immigration had positive economic benefits, these coincided with long-run social costs, such as an erosion of social cohesion, civic mindedness, or an increase in crime. This is particularly likely given the evidence put forth by Tabellini (2018) showing that this period of immigration resulted in short-run social and political frictions.

The first social measure that we consider is a composite index of social capital created by Rupasingha and Goetz (2008) that applies principal component analysis to a range of variables such as the total number of associations and not-for-profit organizations per 10,000 people, as well as census mail response rates and voter turnout. For ease of interpretation, we normalize the original variable, which ranged from -3.9 to 17.5, to lie between zero and one. The 2SLS estimates are reported in column 1 of Table 5.[36] We find a small and statistically insignificant effect of historical immigration on social capital today. According to the estimate, an increase in historical immigration from zero to the 50th percentile (0.049) is associated with an increase in social capital of 0.0023, which is very small relative to the mean of 0.18.

Next, we turn to two alternative measures of social cohesion: political participation and crime. Column 2 of Table 5 reports 2SLS estimates of the long-term effects of immigration on political participation, measured by voter turnout in the 2000 presidential election. We find a positive, but insignificant and modest, effect of historical immigration on voter turnout.[37] Columns 3–5 report estimates of the effects of immigration on the crime rate (crimes in 2000 per capita) for: any crime, crimes against persons, and property crimes.[38] We estimate positive, but statistically insignificant effects of historical immigration on each type of crime.[39]

Overall, based on the estimates in Table 5, we find no evidence of historical immigration having statistically significant effects on social capital, political participation, or crime.

---

36. Partial correlation plots of the reduced-form are reported in Online Appendix Figure A18.

37. According to the estimated magnitude, an increase in historical immigration from zero to the 50th percentile (0.049) is associated with an increase in voter turnout of 1 percentage point, which is small when compared to the mean turnout rate of 54%.

38. The measures are from the *County and City Data Book*, which is produced by the U.S. Census Bureau.

39. According to the point estimate from column 3, an increase in historical immigration from zero to the 50th percentile (0.049) is associated with an increase of 0.0009 crimes per year per capita.

TABLE 5

OLS and 2SLS estimates of the effects of historical immigration on social outcomes

Dependent variable	(1) Social Capital, 2000	(2) Voting Turnout, 2000	(3) Total Crime Rate, 2000	(4) Crimes Against Persons, 2000	(5) Crimes Against Property, 2000
	A. OLS estimates				
Average Immigrant Share, 1860–1920	−0.048 [0.030]	−0.071 [0.046]	0.008*** [0.002]	0.002*** [0.001]	0.004*** [0.001]
	B. Reduced form				
Predicted Average Immigrant Share, 1860–1920	0.210 [0.958]	1.244 [1.662]	0.086 [0.070]	0.020 [0.013]	0.054 [0.053]
	C. 2SLS estimates				
Average Immigrant Share, 1860–1920	0.046 [0.209]	0.271 [0.347]	0.019 [0.017]	0.004 [0.003]	0.012 [0.012]
	D. First-stage estimates				
	Dependent variable: average immigrant share, 1860–1920				
Predicted Average Immigrant Share, 1860–1920	4.588*** [1.329]	4.596*** [1.330]	4.559*** [1.311]	4.559*** [1.311]	4.559*** [1.311]
Kleibergen Paap *F*-statistic	21.206	21.712	21.222	21.222	21.222
Controls (in all Panels):					
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	2,934	2,925	2,935	2,935	2,935
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.182	0.535	0.006	0.001	0.004
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.061	0.090	0.004	0.001	0.003

Notes: An observation is a county. Panel A reports OLS estimates, Panel B reports reduced-form estimates, Panel C reports 2SLS estimates, and Panel D reports first-stage estimates. Coefficient estimates are reported, with Conley standard errors reported in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

6. MECHANISMS

We now turn to an examination of the mechanisms for the finding that historical immigration is associated with better economic outcomes and similar social outcomes today.

6.1. *Evidence for the reallocation of economic activity*

In linking our findings to aggregate economic growth, it is important to recognize that our cross-county estimates can capture two different effects. One is the creation of economic activity in the county and the other is the reallocation of economic activity into the county (from other counties). From the perspective of those living in a county today, it may not be important where the long-run benefits are coming from. However, for a better understanding of the exact reasons for the benefits, it is important. Thus, to assess the extent to which our estimates are due to reallocation effects, we test whether being close to a county with more historical immigration resulted in *less* long-term economic development today. If reallocation is important, we expect that being located near a county with historical immigration might cause economic activity to move from a given county to the nearby county.

We do this by first constructing a measure of average immigrant shares of all neighbouring counties, where we weight each neighbouring county in proportion to the length of the shared border, which we denote Nearby Share$_{i,s}$. We then estimate the following set of equations using 2SLS. The two first-stage equations are:

$$\text{Avg Immigrant Share}_{i,s} = \alpha_s + \alpha_1 \widehat{\text{Avg Immigrant Share}}_{i,s} + \alpha_2 \widehat{\text{Nearby Share}}_{i,s} + \mathbf{X}_{i,s}\Omega + \varepsilon_{i,s}$$
(4)

$$\text{Nearby Share}_{i,s} = \gamma_s + \gamma_1 \widehat{\text{Avg Immigrant Share}}_{i,s} + \gamma_2 \widehat{\text{Nearby Share}}_{i,s} + \mathbf{X}_{i,s}\Pi + \mu_{i,s}.$$
(5)

And, the second stage equation is:

$$Y_{i,s} = \alpha_s + \beta_1 \text{Avg Immigrant Share}_{i,s} + \beta_2 \text{Nearby Share}_{i,s} + \mathbf{X}_{i,s}\Gamma + \nu_{i,s}.$$
(6)

In all three equations, i indexes counties and s states. As before, Avg Immigrant Share$_{i,s}$ is the average share of a county's population from 1860 to 1920 who are foreign born. The new term, Nearby Share$_{i,s}$, is the average share of the population of neighbouring counties (during the same time period) who were foreign born.

The estimates are reported in Online Appendix Table A9. Panel A reports the OLS estimates of equation (6), panel B reports the reduced-form estimates, panel C reports 2SLS estimates of equation (6) and panel D reports estimates from the two first-stage equations—that is, equations (4) and (5).[40] The spillover coefficients are imprecisely estimated and not statistically different from zero. They provide no indication for the presence of negative spatial spillovers. The signs of the coefficients suggest that the spillovers may even be positive. That is, being close to a county with more historical immigration may be economically beneficial today. Most importantly, we also find that the point estimates of the own-county effects remain positive, and for unemployment, urbanization, and schooling, they remain statistically significant.

One potential concern with these results is that adjacent counties may be too close to each other to generate negative spillover effects, especially since these counties may have become part of the same city, commuting zone, or economic region over time. Motivated by this concern, we also estimate the spillover effects of immigration to all other counties in the same state. The measure of Nearby Share$_{i,s}$ used in equations (4)–(6) is the average historical immigrant share of all other counties within the same state. The estimates are reported in Online Appendix Table A10. We continue to find no evidence for negative spillovers, and if anything, weak evidence for positive spillovers.

6.2. *Evidence from short-run estimates*

In an attempt to better understand the mechanisms underlying our long-run estimates, we now study the short-run effects of immigration on industry, agriculture, human capital, and innovation.

Industry: One explanation for the long-run economic benefits of immigration is that, during the early stages of industrial development, immigration provided a large supply of labour that was necessary for the take-off of industry and sustained modern economic growth (Goldin, 1994; Hatton and Williamson, 1998; Hirschman and Mogford, 2009). Several historians have documented that immigrants were disproportionately represented in the industrial workforce

---

40. Because we have multiple endogenous variables, we report Angrist–Pischke first-stage *F*-statistics in the table.

([Engerman and Sokoloff, 2000](#); [Alexander, 2007](#)). For example, in 1880, despite only accounting for approximately 10% of the total U.S. population, immigrants accounted for 57% of the manufacturing workforce ([Hirschman and Mogford, 2009](#)).

Given this, we test whether the data are consistent with immigrants helping to spur early industrialization by estimating the effects of immigration on manufacturing output during the Age of Mass Migration and immediately afterwards. The estimates are reported in Table 6. In the odd numbered columns, we report outcomes measured during our period of interest, 1860–1920.[41] In the even numbered columns, we report outcomes measured in 1930, the decade immediately following the Age of Mass Migration. In columns 1 and 2, we examine the natural log of real manufacturing output per capita. We find that the presence of immigrants caused a large and significant increase in manufacturing output both during the Age of Mass Migration (1860–1920) and immediately afterwards (1930). According to the magnitude of the estimated effects, moving a county with no historical immigration to the 50th percentile (an increase of 0.049) led to a 44% increase in average manufacturing output per capita from 1860 to 1920 and an 80% increase in 1930.

In columns 3–6, we further probe specific channels by examining the effect of immigrants on establishment size, measured using average output per establishment (columns 3 and 4), as well as the effect of immigrants on the number of establishments per 10,000 inhabitants (columns 5 and 6).[42] We find that during the Age of Mass Migration (1860–1920), the primary effect of immigrants was to increase the number of manufacturing establishments and not their size, while in 1930, after the end of the Age of Mass Migration, the primary effect of immigration is to increase the size of establishments. These estimates, which show that immigrants had an immediate positive effect on industrialization, are consistent with historical accounts of immigrants bringing raw labour and manufacturing know-how, both of which were crucial for the growth of manufacturing during this time ([Hirschman and Mogford, 2009](#)).

Agriculture: We next turn to estimates of the short-run effect of immigrants on the agricultural sector, measured using total farm value.[43] Estimates are reported in Table 7, where columns 1 and 2 use farm value per farm (in 1860–1920 and 1930) as the dependent variable, while columns 3 and 4 use farm value per acre. For both sets of estimates, we see modest positive effects of immigration on farm values in 1860–1920, with these effects becoming large and statistically significant in 1930. According to the estimates, moving a county with no historical immigration to the 50th percentile (0.049) is associated with a 22% or 37% increase in 1930 farm values, depending on the measure used. Thus, immigration appears to have had large positive effects in the agricultural sector, with the largest benefits arising just after the end of the Age of Mass Migration.

Human capital: We next turn to the possibility that immigrants may have helped to create a greater stock of technology and human capital. We examine this potential channel by estimating the short-run effects of immigration on educational outcomes. Specifically, we consider the average share of children enrolled in school during the Age of Mass Migration (1870–1920) and immediately

---

41. We note that the decade 1910 is missing from our manufacturing census.

42. We measure establishment size using output per establishment. We use output rather than value-added because value added data are only available for one year of our sample period, 1920. Using this alternative measure, we obtain estimates that are very similar to the estimates of columns 3 and 4.

43. All data are from the Agricultural Census. Acres of land are only reported as being within the following categories: less than 3 acres, 3–9 acres, …, 1000+ acres. We approximate total acreage by using the midpoint of each category, and 1000 for the 1000-or-more-acre category.

TABLE 6
OLS and 2SLS estimates of the effects of historical immigration on manufacturing output

Dependent variable	(1) (2) Log Average Manufacturing Output per Capita		(3) (4) Log Average Manufacturing Output per Establishment		(5) (6) Log Number of Establishments per 10,000 Inhabitants	
	1860-1920	1930	1860-1920	1930	1860-1920	1930
A. OLS estimates						
Average Immigrant Share, 1860–1920	3.476*** [0.631]	4.216*** [0.796]	3.301*** [0.537]	3.343*** [0.648]	0.319** [0.249]	0.783*** [0.248]
B. Reduced form						
Predicted Average Immigrant Share 1860–1920	40.765 [33.988]	74.736*** [26.368]	20.778 [29.227]	71.924*** [23.653]	32.710*** [6.462]	2.079 [6.765]
C. 2SLS estimates						
Average Immigrant Share, 1860–1920	9.014 [8.460]	16.197*** [7.343]	4.594 [6.838]	15.588*** [6.868]	7.253*** [2.389]	0.453 [1.467]
D. First-stage estimates						
Dependent variable: average immigrant share, 1860–1920						
Predicted Average Immigrant Share, 1860–1920	4.523*** [1.381]	4.614*** [1.466]	4.523*** [1.381]	4.614*** [0.927]	4.510*** [1.381]	4.590*** [1.463]
Kleibergen Paap *F*-statistic	16.584	17.729	16.584	17.729	16.376	17.456
Controls (in all Panels):						
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,805	2,463	2,805	2,463	2,804	2,462
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	6.561	7.206	12.578	14.030	3.352	2.487
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	1.197	1.419	0.906	1.156	0.536	0.509

Notes: An observation is a county. The decade of 1910 is missing from the Manufacturing Census. Panel A reports OLS estimates, Panel B reports reduced-form estimates, Panel C reports 2SLS estimates, and Panel D reports first-stage estimates. Coefficient estimates are reported, with Conley standard errors reported in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

afterwards (1930).[44] As reported in columns 1 and 2 of Table 8, we find that counties with a higher share of immigrants actually had lower enrollment rates. This finding is also confirmed by the estimates, reported in columns 3 and 4, which show that immigration is associated with more illiteracy.[45]

Our finding that immigration resulted in less education in the short-run is consistent with the fact that immigrants tended to be less educated than native-born populations, particularly towards the end of the Age of Mass Migration. According to Census figures, in 1850, 9% of immigrants were illiterate compared to 4% of natives. In 1870, these figures are close to equal

44. The education data are from the U.S. Census. Because the first year for which the measures are available is 1870, we examine average education from 1870 to 1920.

45. Unlike our previous estimates, the difference between the OLS and IV estimates might be viewed as evidence for positive selection. At the same time, they are also consistent with negative selection once one recognizes that economic activity can be associated with a greater opportunity cost of schooling and less education in the short-run (Atkin, 2016).

TABLE 7

OLS and 2SLS estimates of the effect of historical immigration on farming

Dependent variable	(1)	(2)	(3)	(4)
	Log Average Total Farm Value (per Farm)		Log Average Total Farm Value (per Acre)	
	1860-1920	1930	1860-1920	1930
	A. OLS estimates			
Average Immigrant Share, 1860–1920	0.571 [0.417]	1.321*** [0.340]	1.866*** [0.699]	2.224*** [0.721]
	B. Reduced form			
Predicted Average Immigrant Share, 1860–1920	−1.771 [18.545]	32.991*** [13.341]	−12.372 [22.477]	14.961 [19.545]
	C. 2SLS estimates			
Average Immigrant Share, 1860–1920	−0.393 [4.116]	7.455*** [3.485]	−2.743 [5.005]	3.367 [4.519]
	D. First-stage estimates			
	Dependent variable: average immigrant share, 1860–1920			
Predicted Average Immigrant Share, 1860–1920	4.510*** [1.381]	4.425*** [1.360]	4.510*** [1.381]	4.443*** [1.359]
Kleibergen Paap *F*-statistic	16.376	15.543	16.376	16.065
Controls (in all Panels):				
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes
Observations	2,804	2,800	2,804	2,799
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	10.420	11.513	5.907	6.558
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.913	0.829	0.701	0.793

Notes: An observation is a county. Panel A reports OLS estimates, Panel B reports reduced-form estimates, Panel C reports 2SLS estimates, and Panel D reports first-stage estimates. Coefficient estimates are reported, with Conley standard errors reported in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

at 15% and 14%, respectively. After this period, the rates began to diverge noticeably: in 1900, 13% of immigrants were illiterate compared to 3% of natives; in 1910, these figures were 12% and 2%; and, in 1920, they were 12% and 1%, respectively.

Comparing the short-run effects of immigration on education in columns 1–4 of Table 8 to the long-run education effects reported in column 5 of Table 4, it is clear that there has been a reversal. In the short-run, immigrants reduced average education, while in the long-run they increased it. There are several possible explanations for this. One explanation is that the long-run effects arise due to the long-run effect of immigrants on income, which is associated with more education. A second explanation is the mechanism found in the recent study by Foged and Peri (2016). The presence of immigrants, and their supply of unskilled labour, could have led native workers to pursue less manual-intensive occupations and to obtain more schooling in the long-run. A final potential explanation is that although immigrants were (on average) less skilled than the native population, they may have had values and aspirational beliefs that facilitated the rapid accumulation of education among their children and/or future generations of children in their communities. This is consistent with the fact that although immigrants were less educated than native populations, their children tended to be more educated.[46]

46. For example, in the 1910 *Report of the Immigration Commission*, which studied 12,011 male iron and steel workers from the Midwest, although the proportion of foreign-born men that could read and write was lower than for

TABLE 8
OLS and 2SLS estimates of the effects of historical immigration on human capital and innovation

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)
	Educational attainment				Innovation	
	Share Enrolled in School		Share Illiterate		Log Patents per 10,000 Inhabitants:	
					All Nationalities	European
	1870-1920	1930	1870-1920	1930	1860-1920	1860-1920
A. OLS estimates						
Average Immigrant Share, 1860–1920	−0.119***	−0.059***	0.103***	0.044***	1.379***	2.992***
	[0.015]	[0.016]	[0.035]	[0.013]	[0.474]	[0.555]
B. Reduced form						
Predicted Average Immigrant Share,	−3.350***	−1.314***	7.558***	1.415***	139.378***	38.047***
1860–1920	[1.009]	[0.451]	[3.190]	[0.781]	[21.909]	[9.336]
C. 2SLS estimates						
Average Immigrant Share, 1860–1920	−0.735***	−0.288***	1.658**	0.310**	30.366***	8.289***
	[0.308]	[0.113]	[0.862]	[0.123]	[9.277]	[1.998]
D. First-stage estimates						
Dependent variable: average immigrant share, 1860–1920						
Predicted Average Immigrant Share,	4.559***	4.559***	4.559***	4.559***	4.590***	4.590***
1860–1920	[1.311]	[0.849]	[0.849]	[0.849]	[1.332]	[1.332]
Kleibergen Paap *F*-statistic	21.222	21.222	21.222	21.222	21.151	21.151
Controls (in all Panels):						
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,935	2,935	2,935	2,935	2,929	2,929
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.190	0.217	0.104	0.041	3.561	0.312
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.035	0.026	0.104	0.042	1.263	0.589

Notes: An observation is a county. Panel A reports OLS estimates, Panel B reports reduced-form estimates, Panel C reports 2SLS estimates, and Panel D reports first-stage estimates. Coefficient estimates are reported, with Conley standard errors reported in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

Innovative activity: Another mechanism through which immigrants could have affected early economic development is through innovative activities and knowledge creation (Fairlie and Lofstrom, 2015). Although most immigrants were unskilled, an important subset of immigrants were highly skilled and important innovators (Wegge, 2002; Long and Ferrie, 2013; Abramitzky and Boustan, 2017). There are many examples of immigrants, who were involved in early industrialization in Europe, bringing more advanced European technologies to the U.S. (Rosenberg, 1972). It has also been argued that the increased availability of unskilled labour due to immigration facilitated the introduction of technological and managerial innovations, such as assembly lines and the rise of the managerial firm (Denison, 1974; Chandler, 1977; Hounshell, 1984; Wright, 1990; Abramovitz and David, 2000; Hirschman and Mogford, 2009). Others have argued that the

native-born men (81.6% versus 98.9%), native-born men with a foreign-born father had *higher* literacy than native-born men with a native-born (and white) father (99.8% versus 98.2%) (Dillingham, 1911, p. 27).

increase in the labour force enabled economies of scale in production, leading to increased profits that spurred innovation (Carter and Sutch, 1999).

To test whether immigration affected innovative activity, we examine patenting rates from 1860–1920, using utility patent data that were obtained from the *United States Patent and Trademark Office*.[47] Estimates are reported in column 5 of Table 8. We find a positive and significant effect of immigration on innovation during this time. An increase in historical immigration from zero to the 50th percentile (0.049) results in a 148% increase in the number of patents per capita.

To assess the extent to which this increase in innovation is due to immigrants innovating themselves or due to them facilitating innovation by native-born Americans, we attempt to identify the country of birth of the innovators in the patent applications. The main challenge is that the citizenship of patent applicants was not consistently reported prior to 1880. Consequently, we are only able to identify the citizenship of the patent applicant in 50% of our sample of 1,297,086 applications. Moreover, following the Naturalization Act of 1798, immigrants could become U.S. citizens after only fourteen years of residence in the country. Therefore, it is possible that several patent applicants are registered as U.S. citizens, despite them being foreign born. Another concern is that there were significant challenges and costs associated with obtaining a patent, which might have placed recently landed foreigners with a limited understanding of English at a disadvantage.[48]

With these caveats in mind, we estimate the effect of immigration on the rate of patenting by inventors that report themselves as being foreign born. The estimates are reported in column 6 of Table 8. We find a positive and statistically significant effect of immigration on foreign patents. However, the magnitude is smaller than for total patents. According to the estimates, an increase in historical immigration from zero to the 50th percentile (0.049) results in an increase in foreign patenting by 41%.[49] This suggests that the direct effect of immigrants on foreign patents was lower than the indirect effect of immigrants on innovation by native-born inventors. Such an indirect effect of immigrants on native inventiveness is consistent with the findings of Moser *et al.* (2014). Although the authors examine a slightly later period than our analysis (post-1920), they show that innovations by German-Jewish immigrants had a significant effect on the rate of innovation of U.S.-born inventors.

A closer analysis of the types of patents that tended to be registered by European-born inventors suggests that, while they were fairly small in number, they may have been disproportionately influential. The importance of their contribution can be inferred by relative citation rates. Of the patents in our sample, 16% are cited by patents in the *NBER Patent Citation Database*, which contains patents from 1975 to 1999. Among the cited patents, 12% are historical patents held by individuals who are European-born, a figure that is significantly higher than the share of all patents that are registered by European-born inventors (which is 3%).

47. Prior to 1927, the introductory paragraph of a patent stated citizenship and residence. Since this is not reported after then, we do not have patent measures for 1930.

48. While the Patent Act of 1793 might have benefited foreigners by removing the requirement of a thorough oral examination as part of the process of granting patents, the cost of a patent was \$35 in 1861, which corresponds to about \$891 in 2010 U.S. dollars. Note, however, that the 1869 Report of the Commissioner of Patents compared the \$35 fee for a U.S. patent to the significantly higher charges in European countries such as Britain, France and Russia (\$450); Belgium (\$420); and Austria (\$350).

49. These findings are robust to an alternative specification that uses a poisson estimator, while controlling for the natural log of the average population of a country from 1860 to 1920 (see Online Appendix Table A11).

6.3. *Examining effects over both the short- and long-run*

We now attempt to connect the short- and long-run effects of immigration by examining the full time series of effects immediately after the Age of Mass Migration. To do so, we examine urbanization, which is positively associated with income and is available at regular time intervals during our time span of interest. We estimate equation (3) with urbanization measured in each decade from 1920 to 2000 as the outcome of interest. The estimates, which are reported in Online Appendix Table A12, show that by 1920 one already observes a large positive effect of immigration on urbanization. This effect remains stable until about 2000, when it increases slightly. Thus, the economic benefits of immigrants were felt early and persisted over time. This is consistent with immigration affecting early industrialization, which, due to agglomeration, increasing returns, network externalities, or other lock-in effects, caused a persistent increase in urbanization.

Unfortunately, unlike urbanization, other measures are not available during the full time span. For education and per capita income, we can examine how the effects evolve over time, but only in the post-WWII era. These estimates, which we report in Online Appendix Tables A13 and A14, show that we observe the same basic trends for education and income as we do for urbanization. In the medium- and long-runs, we see that the effects of immigrants persist over time.[50]

6.4. *Historical migration and subsequent immigration*

We next consider the possibility that part of the mechanism for the estimated effects is due to the impact that historical immigration had on the movement of native-born populations within the U.S. We test for this by estimating the effects of historical immigration on a number of proxies for the migration of native-born of populations into a county. Estimates are reported in Online Appendix Table A15 for the following outcome variables: growth in a county's native-born population from 1860 to 1920 (column 1), growth in a county's native-born population who were born in another state from 1850 to 1920 (column 2), average fraction of the native-born population who were born out of state from 1850 to 1920 (column 3); and the average fraction of the native-born population of men twelve or older who were born out of state from 1850 to 1920 (column 4).[51] We find limited evidence that immigration led to greater internal migration. Although the estimated effects of historical immigration on each of the four proxies for internal migration are positive, only one is significant and only marginally so.

We next test the possibility subsequent immigration following the Age of Mass Migration explains part of the long-term effects we observe. We check for this by estimating the effect of historical immigration on the size of the foreign-born population in each decade since 1920. The estimates, which are reported in Online Appendix Table A16, show that immediately following the Age of Mass Migration, historical immigration is (mechanically) associated with a greater foreign-born population share. However, the relationship quickly fades over time, suggesting that it is unlikely that contemporary immigration is an important channel for our findings.[52] By 1950, the magnitude of the estimated effect is reduced by more than half and by 1980, it is not statistically different from zero.

50. These findings are consistent with the recent findings from Bleakley and Lin (2012), who find evidence of lock-in effects in the context of historical U.S. portage sites.

51. The measures from columns 2–4 are from the complete count U.S. Census Microdata. Since 1860 is unavailable, in column 2, 1850 is used as the initial year. The average of columns 3 and 4 are average over the following Census decades: 1850, 1880, 1890, 1900, 1910, and 1920.

52. As we report in Online Appendix Tables A17 and A18, our baseline estimates remain very similar when we control for the share of the population that is foreign born in 2000.

7. ROBUSTNESS AND SENSITIVITY CHECKS

We now check the robustness of our estimates. Table 11 summarizes the findings for per capita income that are discussed in this section. Due to space limitations, the findings for all other outcomes are reported in Online Appendix tables.

7.1. *Using aggregate immigration that is predicted by origin country weather shocks*

A potential concern with our identification strategy is that the aggregate inflow of immigrants into the U.S. may have been greater when the railway became connected to counties with more future growth potential. To address this concern, we test the robustness of our findings to the use of a measure of predicted aggregate immigration that is driven only by origin-country weather shocks. The strategy is motivated by existing evidence that European weather shocks were important determinants of emigration during this period (Solomou and Wu, 1999; Karadja and Prawitz, 2016).

To construct measures of origin-country weather shocks, we use historical temperature data from Luterbacher *et al.* (2004) and historical precipitation data from Pauling *et al.* (2006). The data are measured four times annually (once during each season) and at a 0.5-degree (approx. 55 kilometer) spatial resolution. Because the emigration data are at the country-level, we create country-level averages of temperature and precipitation by taking an average over all grid-cells in a country that were under cultivation at the time.[53] Our sample includes sixteen European countries for which we have immigration, temperature, and crop data.[54]

We estimate the relationship between weather shocks and the outflows of emigrants using the following equation:

$$\ln \text{Immigrant Flow}_{c,t+1} = \sum_{s \in S} \sum_{k \in K} \beta_{c,s,k} I_{c,t}^{Temp,s,k} + \sum_{s \in S} \sum_{k \in K} \gamma_{c,s,k} I_{c,t}^{Precip,s,k} + \varepsilon_{c,t} \qquad (7)$$

where c denotes countries and t years (1850–1929). The dependent variable, $\ln \text{Immigrant Flow}_{c,t+1}$, is the natural log of the flow of immigrants from origin country c to the U.S. in year $t+1$. $I_{c,t}^{Temp,s,k}$ is an indicator variable that equals 1 if the average temperature in season $s \in \{$Spring, Summer, Winter, Autumn$\}$ falls within temperature range k, where k indexes a set K of six temperature categories: 3+ standard deviations below the mean, 2–3 standard deviations below the mean, 1–2 standard deviations below the mean, 1–2 standard deviations above the mean, 2–3 standard deviations above the mean, and 3+ standard deviations above the mean; the omitted category is for temperatures that are within 1 standard deviation of the mean (*i.e.* the absence of a shock). Since there are six temperature categories and four seasons, there are $6 \times 4 = 24$ temperature indicators in total. The equation also includes 24 precipitation indicator variables that are structured in exactly the same manner.

The coefficients for the weather variables, $\beta_{c,s,k}$ and $\gamma_{c,s,k}$, are allowed to vary across countries. In practice, we estimate equation (7) separately for each country in our sample and use the predicted values of the $\beta_{c,s,k}$'s and the $\gamma_{c,s,k}$'s to calculate the predicted log immigrant flow from each country in each year, $\ln \widehat{\text{Immigrant Flow}}_{c,t}$. We find that the predicted immigrant flows are

---

53. Information on historical land under cultivation is taken from Ramankutty and Foley (1999), who provide annual estimates at a 5-arc-minute resolution.

54. These are: Belgium, Denmark, France, Germany, Greece, Hungary, Ireland, Italy, Netherlands, Norway, Poland, Portugal, Russia, Spain, Sweden, and Switzerland. The countries account for 75% of European immigration into the U.S. from 1860 to 1920, as measured by Willcox (1929).

TABLE 9

OLS and 2SLS estimates of the effects of historical immigration on economic prosperity today, using immigrant inflows predicted by sending-country weather shocks

Dependent variable	(1) Log Average per Capita Income, 2000	(2) Share of Pop. Below Poverty Line, 2000	(3) Unemployment Rate, 2000	(4) Urbanization Rate, 2000	(5) Average Years of Schooling, 2000
A. OLS estimates					
Average Immigrant Share, 1860–1920	0.244* [0.130]	0.015 [0.028]	0.020 [0.015]	0.948*** [0.185]	0.022 [0.307]
B. Reduced form					
Predicted Average Immigrant Share, 1860–1920	28.934*** [9.683]	−4.571 [2.874]	−4.402*** [1.453]	65.007*** [19.282]	115.574*** [27.643]
C. 2SLS estimates					
Average Immigrant Share, 1860–1920	2.792*** [1.249]	−0.441 [0.298]	−0.425*** [0.195]	6.273*** [2.651]	11.152*** [4.397]
D. First-stage estimates					
Dependent variable: average immigrant share, 1860–1920					
Predicted Average Immigrant Share, 1860–1920	10.364*** [3.058]	10.364*** [3.058]	10.364*** [3.058]	10.364*** [3.058]	10.364*** [3.058]
Kleibergen Paap *F*-statistic	18.006	18.006	18.006	18.006	18.006
Controls (in all Panels):					
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	2,935	2,935	2,935	2,935	2,935
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	10.022	0.136	0.047	0.401	11.445
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.203	0.054	0.025	0.305	0.558

Notes: An observation is a county. Panel A reports OLS estimates, Panel B reports reduced-form estimates, Panel C reports 2SLS estimates, and Panel D reports the first-stage estimates. Coefficient estimates are reported, with Conley standard errors in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

strongly correlated with actual immigrant flows.[55] We then aggregate the predicted flows across countries to obtain an estimate of the total flow of emigrants from all sixteen countries to the U.S. in a given decade: $\text{Agg}\,\widehat{\text{Immigrant Flow}}_t = \sum_c \exp(\ln \widehat{\text{Immigrant Flow}}_{c,t})$.

We then re-estimate equations (2) and (3), replacing actual immigrant inflows with predicted immigrant inflows. The estimates are reported in Tables 9 and 10.[56] The findings are very similar to the baseline estimates.

55. The correlation coefficients between the actual and predicted flows measures range from 0.54 (for Switzerland) to 0.91 (for Hungary). The relationship between the two measures for each of our sixteen countries is shown in Online Appendix Figure A19.

56. The zero-stage estimates of equation (1) using predicted immigrant flows are reported in Online Appendix Table A19.

TABLE 10

OLS and 2SLS estimates of the effects of historical immigration on social outcomes, using immigrant inflows predicted by sending-country weather shocks

Dependent variable	(1) Social Capital, 2000	(2) Voting Turnout, 2000	(3) Total Crime Rate, 2000	(4) Crimes Against Persons, 2000	(5) Crimes Against Property, 2000
	A. OLS estimates				
Average Immigrant Share,	−0.048	−0.071	0.008***	0.002***	0.004***
1860–1920	[0.030]	[0.046]	[0.002]	[0.0005]	[0.001]
	B. Reduced form				
Predicted Average Immigrant Share,	1.506	2.733	0.348	0.063	0.241
1860–1920	[2.838]	[5.032]	[0.215]	[0.039]	[0.162]
	C. 2SLS estimates				
Average Immigrant Share,	0.144	0.254	0.034	0.006	0.023
1860–1920	[0.278]	[0.460]	[0.024]	[0.005]	[0.018]
	D. First-stage estimates				
	Dependent variable: average immigrant share, 1860–1920				
Predicted Average Immigrant Share,	10.424***	10.776***	10.364***	10.364***	10.364***
1860–1920	[3.097]	[3.161]	[3.058]	[3.058]	[3.058]
Kleibergen Paap *F*-statistic	17.806	18.143	18.006	18.006	18.006
Controls (in all Panels):					
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	Yes	Yes
Observations	2,934	2,925	2,935	2,935	2,935
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.182	0.535	0.006	0.001	0.004
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.061	0.090	0.004	0.001	0.003

Notes: An observation is a county. Panel A reports OLS estimates, Panel B reports reduced-form estimates, Panel C reports 2SLS estimates, and Panel D reports first-stage estimates. Coefficient estimates are reported, with Conley standard errors reported in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

7.2. *Alternative formulations of the instrument*

In our baseline specification, we control for the length of time a county has been connected to the railroad network to account for any potential relationship between our instrument and how early or late a county became connected to the railway network. We now check the robustness of our results to the use of alternative strategies to account for the potential relationship between the instrument and how early a county was connected to the railway network. Recall that our baseline instrument is: $\text{Avg Immigrant Share}_i = \frac{1}{T} \sum_{t=1}^{T} \widehat{\beta} \text{Immigrant Flow}_{t-1} \times I_{it-1}^{\text{RR Access}}$. During periods without railway access, $I_{i,t-1}^{\text{RR Access}} = 0$, which mechanically reduces the value of $\text{Avg Immigrant Share}_i$. We construct an alternative predicted immigrant share that is the mean of the predicted immigrant share, but only in the periods from 1860 to 1920 for which the county was connected to the railway network. Specifically, the alternative measure is: $\text{Avg Immigrant Share}_i = \frac{1}{N_i^{\text{RR}}} \sum_{t \in T_i^{\text{RR}}} \widehat{\beta} \text{Immigrant Flow}_{t-1} \times I_{i,t-1}^{\text{RR Access}}$, where N_i^{RR} is the number of time periods for which $I_{i,t-1}^{\text{RR Access}} = 1$ in county i, and T_i^{RR} is the set of census years for which $I_{it-1}^{\text{RR Access}} = 1$

TABLE 11

OLS and 2SLS estimates of the impacts of historical immigration on current income: Robustness checks

Dependent variable	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
		Variants of Instrument			Additional Covariates			Restricted Samples	
	Baseline Specification	Average in Connected Decades	Immigrant Inflow in First Decade	No Covariates	Frontier Experience	Civil War	Internat Trade	Removing Outliers	Counties w Constant Borders
	Log Average per Capita Income, 2000								
A. OLS estimates									
Average Immigrant Share, 1860–1920	0.243* [0.130]	0.246* [0.133]	0.269** [0.130]	0.395*** [0.134]	0.220 [0.128]	0.244* [0.130]	0.235*** [0.129]	0.333*** [0.087]	0.607*** [0.193]
B. Reduced form									
Predicted Average Immigrant Share, 1860–1920	11.942*** [3.629]	287.404*** [57.292]	49.319*** [12.435]	12.719*** [2.293]	9.793*** [3.413]	12.183*** [3.618]	10.527*** [2.912]	14.684*** [3.773]	15.778*** [5.794]
C. 2SLS estimates									
Average Immigrant Share, 1860-1920	2.619*** [1.022]	8.532* [4.497]	4.540*** [1.773]	2.865*** [0.779]	2.238*** [0.964]	2.693** [1.035]	3.564*** [1.581]	3.831*** [1.460]	3.119** [1.615]
D. First-stage estimates									
Dependent variable: average immigrant share, 1860–1920									
Predicted Average Immigrant Share, 1860–1920	4.559*** [1.311]	33.686*** [17.511]	10.862*** [3.275]	4.439*** [1.214]	4.375*** [1.301]	4.524*** [1.315]	2.954*** [1.132]	3.833*** [1.324]	5.059*** [2.129]
Kleibergen Paap *F*-statistic	21.222	18.803	21.005	75.973	19.883	21.256	21.225	17.027	10.675
Controls (in all Panels):									
Never Connected to the Railroad [0-1]	No	Yes	Yes	No	No	No	No	No	No
Number of Years with Frontier Experience	No	No	No	No	Yes	No	No	No	No
County Connected During Civil War (1860)	No	No	No	No	No	Yes	No	No	No
Trade-Based Predicted Immigrant Share	No	No	No	No	No	No	Yes	No	No
Industrialization-Based Predicted Immig. Share	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Business Cycle-Based Predicted Immig. Share	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Date of RR Connection (Log Years as of 2000)	Yes	No	No	No	Yes	Yes	Yes	Yes	Yes
Cubic Polynomial for Latitude and Longitude	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
State Fixed Effects	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes
Observations	2,935	2,935	2,935	2,935	2,934	2,935	2,935	2,761	1,489
Mean of Dep. Var. (OLS, Reduced Form, and 2SLS)	10.022	10.022	10.022	10.022	10.022	10.022	10.022	10.013	10.021
SD of Dep. Var. (OLS, Reduced Form, and 2SLS)	0.203	0.203	0.203	0.203	0.203	0.203	0.203	0.170	0.205

Notes: An observation is a county. Panel A reports OLS estimates, Panel B reports reduced form estimates, Panel C reports 2SLS estimates, and Panel D reports the first-stage estimates from the 2SLS. Coefficient estimates are reported, with Conley standard errors reported in square brackets. ***, **, and * indicate significance at the 1%, 5% and 10% levels.

for county i. Because periods without a connection to the railway network are not included in the average, not being connected to the railway, $I_{i,t-1}^{\text{RR Access}} = 0$, no longer mechanically reduces $\widehat{\text{Avg Immigrant Share}}_i$. Column 2 of Table 11 reports the income estimates with the alternative instrument. The estimates for all outcomes are reported in Online Appendix Tables A22 and A23.[57] The estimated effects of historical immigration on economic and social outcomes are qualitatively similar to the estimates using the baseline instrument.

The second alternative instrument that we use follows the same logic as our baseline instrument, but the instrument directly measures the extent to which a county became connected during a boom or a lull decade. Thus, we use the normalized aggregate immigrant inflow measure during the decade following a county's first connection to the railway network as an instrument. One complication with this strategy is that the instrument is not defined for counties that were not connected to the railway network during the Age of Mass Migration. For these counties (127 in total) we assign the instrument a value of zero and also control for an indicator variable that equals one for counties that were never connected. Estimates using this alternative instrument are reported in column 3 of Table 11 (for income), as well as Online Appendix Tables A20 and A21 (for all outcomes of interest). The estimated effects of historical immigration remain robust to the use of this alternative version of the instrument.

7.3. *Instrument validity checks*

As a further test for the validity of our instrument, we also undertake the following placebo test of whether past growth in manufacturing activity is correlated with future values of our instrument. We first split our sample into two periods of equal length: 1860–90 and 1900–20. We then consider all counties that up to 1890 were not yet connected to the railroad network. For these counties, their measure of predicted immigrant share is driven by variation in connection and immigrant inflows after 1890. We then test whether we observe a relationship between predicted immigrant share and average growth in manufacturing output per capita during the early period (1860–90). If we observe a correlation, this is evidence of early manufacturing growth affecting the timing of subsequent railway connectivity and the value of the instrument. The results of the test, which are reported in panel A of Online Appendix Table A24 show that the relationship between the two measures is weak and insignificant, which is consistent with the validity of our identification strategy. As reported in panels B and C, the findings are similar if 1870 or 1880 is used as a cut-off rather than 1890.

7.4. *Sensitivity of estimates to the set of covariates*

We next turn to an examination of the sensitivity of our estimates to the set of covariates that is included in our regression equation. We first check the robustness of our findings to the omission of our baseline set of covariates. The 2SLS estimates are reported in column 4 of Table 11 for per capita income, while the estimates for all economic and social outcomes are reported in Online Appendix Tables A25 and A26. We find that the estimates of interest remain very similar.[58]

The second exercise that we undertake is to expand our set of covariates beyond the baseline set. Since immigration is potentially associated with the westward expansion of the American

---

57. Since the predicted average immigrant share instrument for counties that are never connected to the railway network is zero, the specifications include an indicator variable for whether the county was never connected to the railway.

58. Formal tests of coefficient stability for the first-stage and reduced-form estimates for our economic outcomes are reported in panels B and C of Online Appendix Table A4.

frontier and given the recent findings of Bazzi *et al.* (2017) on the importance of frontier history, we check that our estimates are robust to controlling for their measure of the number of years a county was a "frontier" county. The estimates, which are reported in column 5 of Table 11 and Online Appendix Tables A27 and A28, show that the findings are robust to controlling for this measure.

Another potentially important factor is the Civil War, especially given the role that railways played in the War (Weber, 1952). We check the sensitivity of our findings to accounting for this by also controlling for a county-level indicator variable that equals one if a county became connected to the railway network during the decade of the Civil War (1860s). The estimates, which are reported in column 6 of Table 11 and Online Appendix Tables A29 and A30, are very similar to our baseline estimates.

The final covariate that we consider addresses the possibility that the railway may have had a direct effect on long-term development by connecting a county to international goods markets. To account for this, we allow the effect of the railway connection to vary differentially depending on the level of aggregate international trade at the time: $\text{Trade}_{t-1} \times I_{i,t-1}^{\text{RR Access}}$, where Trade_{t-1} is the annual average of the value of aggregate U.S. imports plus exports divided by GDP during the ten years prior to census year t. The interaction term captures differential long-run effects that connection to the railway network had depending on the level of international trade at the time. We find that our estimates are robust to the inclusion of the trade interaction control (see column 7 of Table 11 and Online Appendix Tables A31 and A32).

7.5. *Sensitivity to influential observations*

We also examine the robustness of our estimates to the removal of potentially influential observations, which are identified as observations with a Cook's distance (from the reduced-form regression) that is greater than $4/N$ (where N is the number of observations). The estimates with influential observations omitted, reported in column 8 of Table 11 and in Online Appendix Tables A33 and A34, show that our findings remain robust to the omission of influential observations.[59]

7.6. *Changing county boundaries*

For many counties in our sample, the current boundaries were established sometime after the initial period. As a consequence, our zero-stage panel is unbalanced, with counties entering over time as they are established.[60] Additionally, once counties are established, there can be changes to their boundaries.[61] Given these issues, we also test the sensitivity of our estimates to using a restricted sample of 1,489 counties, which consist of those that existed in 1860 and have the same boundaries from 1860 to 2000. As reported in column 9 of Table 11 and Online Appendix Tables A36 and A37, the results using the restricted sample are qualitatively identical to our baseline estimates.

---

59. An alternative strategy to address the potential of influential observations is to estimate quantile regressions. As reported in Online Appendix Table A.35, we obtain qualitatively identical conclusions from quantile IV estimates.

60. In 1860, there are 1,532 counties in our sample, there are 1,922 counties in 1870; 2,137 in 1880; 2,416 in 1890; 2,692 in 1900; 2,752 in 1910; and 2,935 in 1920.

61. For our baseline analysis, we match counties across time using the nominally integrated series available in the NHGIS data sets (Minnesota Population Center, 2011). For a detailed explanation of the NHGIS matching strategy, see `https://nhgis.org/documentation/time-series#geographic-integration`.

8. CONCLUSION

The goal of this study is to make progress on understanding the long-run effects of large-scale immigration. We examined the effects of the largest wave of immigration in U.S. history, the Age of Mass Migration, which occurred roughly from 1860 to 1920. To help identify the causal effects of immigrants on the locations in which they settled, we used an IV strategy that exploited the significant decade-by-decade fluctuations in aggregate immigrant flows to the U.S. that occurred during this era, the fact that immigrants typically used the railway to travel to their eventual destination, and the gradual expansion of the railway network over time.

We found that immigration resulted in large long-run economic benefits. Counties with more historical immigration, have higher incomes, less unemployment, less poverty, more education, and higher shares of urban population. We also found that these economic benefits do not have long-run social costs. Places with more historical immigrant settlement today have similar levels of social capital, civic participation, and rates of crime.

To better understand the underlying mechanisms, we examined the short-run effects of immigration. We found that immigrants resulted in an immediate increase in industrialization. Immigrants first contributed to the establishment of more manufacturing facilities and then to the development of larger facilities. We also found large positive effects of immigrants on agricultural productivity and innovation as measured by patenting rates. Having examined the short-run effects of immigration, we then turned to an examination of the dynamic effects of immigrants over the short-, medium-, and long-runs. Examining urbanization rates from 1920 to 2000, we found that large effects on urbanization were felt immediately and persisted over time. We also examined income and education, but for the more limited time period for which data are available (post WWII). We found a similar pattern for these outcomes as well.

Taken as a whole, our estimates provide evidence consistent with an historical narrative that is commonly told of how immigration facilitated economic growth. The less-skilled immigrants provided the labour force necessary for industrial development. A smaller number of immigrants brought with them knowledge, skills, and know-how that were beneficial for industry and increased productivity in agriculture. Thus, by providing a sizeable workforce and a (smaller) number of skilled workers, immigration led to early industrial development and long-run prosperity, which continues to persist until today.

Despite the specific conditions under which the largest episode of immigration in U.S. history took place, our estimates of the long-run effects of immigration may still be relevant for assessing the long-run effects of immigrants today. Our findings suggest that the long-run benefits of immigration can be large and need not come at high social cost. In addition, these benefits can be realized quickly and are highly persistent. These findings highlight the importance of taking a long-run view when considering the current immigration debate.

Supplementary Data

Supplementary data are available at *Review of Economic Studies* online.

REFERENCES

ABRAMITZKY, R. and BOUSTAN, L. P. (2017), "Immigration in American History", *Journal of Economic Literature*, **55**, 1311–1345.

—, — and ERIKSSON, K. (2012), "Europe's Tired, Poor, Huddled Masses: Self-Selection and Economic Outcomes in the Age of Mass Migration", *American Economic Review*, **102**, 1832–1856.

—, — and —. (2013), "Have the Poor Always Been Less Likely to Migrate? Evidence from Inheritance Practices during the Age of Mass Migration", *Journal of Development Economics*, **102**, 2–14.

—, — and —. (2014), "A Nation of Immigrants: Assimilation and Economic Outcomes in the Age of Mass Migration", *Journal of Political Economy*, **122**, 467–717.

ABRAMOVITZ, M. and DAVID, P. A. (2000), "American Macroeconomic Growth in the Era of Knowledge-Based Progress: The Long Run Perspective", in Engerman, S. E. and GALLMAN, R. E. (eds.), *Cambridge Economic History of the United States, Vol. 3* (Cambridge: Cambridge University Press), 1–92.

AGER, P. and BRUCKNER, M. (2013), "Cultural Diversity and Economic Growth: Evidence from the US during the Age of Mass Migration", *European Economic Review*, **64**, 76–97.

— and —. (2017), "Immigrants' Genes: Genetic Diversity and Economic Development in the US", *Economic Inquiry*, **56**, 1149–1164.

—, CINNIRELLA, F., and JENSEN, P. S. (2016), "The Kindergarten Movement and the U.S. Demographic Transition", (Mimeo, University of Southern Denmark).

ALEXANDER, J. G. (2007), *Daily Life in Immigrant America, 1870–1920* (Chicago: Ivan R. Dee).

ATKIN, D. (2016), "Endogenous Skill Acquisition and Export Manufacturing in Mexico", *American Economic Review*, **106**, 2046–2085.

BANDIERA, O., MOHNEN, M., RASUL, I. *ET AL.* (2016), "Nation-Building Through Compulsory Schooling During the Age of Mass Migration" (Working Paper, LSE).

BAZZI, S., FISZBEIN, M., and GEBRESILASSE, M. (2017), "Frontier Culture: The Roots and Persistence of "Rugged Individualism" in the United States" (NBER Working Paper No. 23997).

BERGQUIST, J. M. (2007), *Daily Life in Immigrant America, 1820–1870* (Westport, CT: Greenwood Press).

Blau, F. (1980), "Immigration and Labor Earnings in Early Twentieth Century America", *Research in Population Economics*, **2**, 21–41.

BLEAKLEY, H. and LIN, J. (2012), "Portage and Path Dependence", *Quarterly Journal of Economics*, **127**, 587–644.

BURCHARDI, K. B., CHANEY, T. and HASSAN, T. A. (2015), "Migrants, Ancestors, and Investments", *The Review of Economic Studies*, forthcoming.

CANCE, A. E. (1925), "Immigrants and American Agriculture", *Journal of Farm Economics*, **7**, 102–114.

CARD, D. (2012), "The Elusive Search for Negative Wage Impacts of Immigration", *Journal of the European Economic Association*, **10**, 1–21.

CARTER, S. and SUTCH, R. (1999), "Historical Background to Current Immigration Issues", in Smith, J. P. and Edmonston, B. (eds.), *The Immigration Debate: Studies on the Economic, Demographic and Fiscal Effects of Immigrants* (Washington, DC: National Academy Press), 289–366.

CHANDLER, A. (1977), *The Visible Hand: The Managerial Revolution in American Business* (Cambridge, MA: Harvard University Press).

DANIELS, R. (2002), *Coming to America: A History of Immigration and Ethnicity in American Life* (New York: Harper Perennial).

DAVIS, J. (2004), "An Annual Index of U.S. Industrial Production, 1790–1915", *Quarterly Journal of Economics*, 1177–1215, **119**, 642–658.

DENISON, E. F. (1974), *Accounting for United States Economic Growth, 1929–1969* (Washington, DC: Brookings Institution).

DESMET, K., NAGY, D. K., and Rossi-Hansberg, E. (2018), "The Geography of Development", *Journal of Political Economy*, **126**, 903–983.

DILLINGHAM, W. P. (1911), "Immigrants in Industries. Part 2: Iron and Steel Manufacturing", in The Immigration Commission, ed., *Reports of the Immigration Commission* (Washington, DC: Government Printing Office).

DROLLER, F. (2017), "Economic Development: Evidence from Settlements in the Pampas", *The Economic Journal*, **128**, 2321–2352.

ENGERMAN, S. L. and SOKOLOFF, K. (2000), "Technology and Industrialization, 1790–1914", in Engerman, S. E. and Gallman, R. E. (eds.), *Cambridge Economic History of the United States, Vol. 2* (Cambridge: Cambridge University Press), 367–401.

FAIRLIE, R. W. and LOFSTROM, M. (2015), "Immigration and Entrepreneurship" in Chiswick, B. and Millers, P. (eds.), *Handbook of the Economics of International Migration*, Vol. 1B. (Oxford, UK: Elsevier).

FAULKNER, H. U. (1960), *American Economic History* (New York: Harper and Row Publishers).

FAUST, A. B. (1916), *The Germans in the United States* (New York: German University League).

FISCHER, D. H. (1989), *Albion's Seed: A Cultural History* (New York: Oxford University Press).

FOERSTER, R. (1924), *The Italian Emigration of our Times* (Cambridge, Mass.: Harvard University Press).

—. (1969), *The American Immigration Collection* (New York: Arno Press Inc.).

FOGED, M. and PERI, G. (2016), "Immigrants' Effect on Native Workers: New Analysis on Longitudinal Data", *American Economic Journal: Applied Economics*, **8**, 1–34.

FULFORD, S. L., PETKOV, I., and SCHIANTARELLI, F. (2017), "Does It Matter Where You Came From? Ancestry Composition and Economic Performance of US Counties, 1850–2010" (Mimeo, Boston College).

GOLDIN, C. (1994), "The Political Economy of Immigration Restriction in the United States, 1890 to 1921", in Claudia Goldin and Gary D. Libecap, eds. *The Regulated Economy: A Historical Approach to Political Economy* (Chicago: University of Chicago Press), 223–257.

GRIPSHOVER, M. M. and BELL, T. L. (2012) "Patently Good Ideas: Innovations and Inventions in U.S. Onion Farming, 1883–1939", *Material Culture*, **44**, 1–30.

GROSJEAN, P. (2014), "A History of Violence: The Culture of Honor and Homicide in the US South", *Journal of the European Economic Association*, **12**, 1285–1316.

HAINES, M. R. and Inter-university Consortium for Political and Social Research, (2010) "Historical, Demographic, Economic, and Social Data: The United States, 1790-2002". ICPSR02896-v3. Ann Arbor, MI: Inter-university Consortium for Political and Social Research [distributor], 2010-05-21. http://doi.org/10.3886/ICPSR02896.v3.

HANDLIN, O. (1957), *Race and Nationality in American Life* (Boston: Doubleday Anchor Books).

HANNON, J. U. (1982), "Ethnic Discrimination in a 19th-Century Mining District: Michigan Copper Mines, 1888", *Exploration in Economic History*, **19**, 28–50.

HATTON, T. J. and WILLIAMSON, J. G. (1988), *The Age of Mass Migration* (New York: Oxford University Press).

— and —. (2005), *Global Migration and the World Economy: Two Centuries of Policy and Performance* (Cambridge, MA: MIT Press, 2005).

HEDGES, J. B. (1926), "The Colonization Work of the Northern Pacific Railroad", *Mississippi Valley Historical Review*, **13**, 311–342.

HIGHAM, J. (2011), *Strangers in the Land: Patterns of American Nativism, 1860–1925* (London: Rutgers University Press).

HIRSCHMAN, C. and MOGFORD, E. (2009), "Immigration and the American Industrial Revolution from 1880 to 1920", *Social Science Research*, **38**, 897–920.

HORNUNG, E. (2014), "Immigration and the Diffusion of Technology: The Huguenot Diaspora in Prussia", *American Economic Review*, **104**, 84–122.

HOUNSHELL, D. A. (1984), *From the American System to Mass Production, 1800–1932: The Development of Manufacturing Technology in the United States* (Baltimore: Johns Hopkins University Press).

HUNT, J. and GAUTHIER-LOISELLE, M. (2010), "How Much Does Immigration Boost Innovation?" *American Economic Journal: Macroeconomics*, **2**, 31–56.

JORDAN, T. G. (1966), *German Seed in Texas Soil: Immigrant Farmers in Nineteenth-Century Texas* (Austin: Yale University Press).

KARADJA, M. and PRAWITZ, E. (2016), "Exit, Voice and Political Change: Evidence from Swedish Mass Migration to the United States" (Working Paper, Uppsala University).

KENNEDY, J. F. (1964), *A Nation of Immigrants* (New York: Harper & Row Publishers).

KERR, S. P. and KERR, W. R. (2017), "Immigrant Entrepreneurship", in Measuring Entrepreneurial Businesses: Current Knowledge and Challenges, Haltiwanger, Hurst, Miranda, and Schoar.

KOLLMORGEN, W. M. (1942), *The Pennsylvania German Farmer* (Princeton, NJ: Princeton University Press).

LONG, J. and FERRIE, J. (2016), "British, American, and British-American Social Mobility: Intergenerational Occupational Change among Migrants and Non-Migrants in the Late 19th Century". (Working Paper, Northwestern University).

LUEBKE, F. C. (1977), "Ethnic Group Settlement on the Great Plains", *Western Historical Quarterly*, **8**, 405–430.

LUTERBACHER, J., DIETRICH, D., XOPLAKI, E. *ET AL.* (2004), "European Seasonal and Annual Temperature Variability, Trends, and Extremes Since 1500", *Science*, **303**, 1499–1503.

MADDISON, A. (2001), *The World Economy: A Millennial Perspective* (Paris: OECD Publications).

MALONE, D. (1935), "The Intellectual Melting-Pot", *The American Scholar*, **4**, 444–459.

MALONEY, W. F. and CACEIDO, F. V. (2017), "Engineering Growth: Innovative Capacity and Development in the Americas" (Mimeo, World Bank).

MCGOULDRICK, P. F. and TANNEN, M. B. (1977), "Did American Manufacturers Discriminate Against Immigrants Before 1914?", *Journal of Economic History*, **37**, 723–746.

MIGRATION POLICY INSTITUTE. (2016), "Data Hub: Tabulations of U.S. Department of Homeland Security, Office of Immigration Statistics, Yearbook of Immigration Statistics (various years)". http://migrationpolicy.org/programs/data-hub. Accessed July 2016.

MINNESOTA POPULATION CENTER. (2011), "National Historical Geographic Information System: Version 2.0". (Minneapolis, MN: University of Minnesota). http://www.nhgis.org.

MOSER, P., VOENA, A., and WALDINGER, F. (2014), "German Jewish Emigres and US Invention", *American Economic Review*, **104**, 3222–3255.

MULLER, T. (1993), *Immigrants and the American City* (New York: New York University Press).

OTTAVIANO, G. I. P. and PERI, G. (2006), "The Economic Value of Cultural Diversity: Evidence from U.S. Cities", *Journal of Economic Geography*, **6**, 9–44.

PAULING, A., LUTERBACHER, J., CASTY, C. *ET AL.* (2006), "Five Hundred Years of Gridded High-Resolution Precipitation Reconstructions over Europe and the Connection to Large-Scale Circulation", *Climate Dynamics*, **26**, 387–405.

PAZ, F. H. (2015), "The Long-Term Return to Early childhood Education: Evidence from the First U.S. Kindergartens" (Mimeo, University of California Los Angeles).

PERI, G. (2012), "The Effects of Immigration on Productivity: Evidence from US States", *Review of Economics and Statistics*, **94**, 348–358.

— and SPARBER, C. (2009), "Task Specialization, Immigration, and Wages", *American Economic Journal: Applied Economics*, **1**, 135–169.

RAMANKUTTY, N. and FOLEY, J. A. (1999), "Estimating Historical Changes in Global Land Cover: Croplands from 1700 to 1992", *Global Biogeochemical Cycles*, **13**, 997–1027.

ROCHA, R., FERRAZ, C., and SOARES, R. R. (2017), "Human Capital Persistence and Development", *American Economic Journal: Applied Economics*, **9**, 105–136.

ROSENBERG, N. (1972), *Technology and American Economic Growth* (White Plains, NY: M.E. Sharpe).

RUPASINGHA, A. and GOETZ, S. J. (2008), "U.S. County-Level Social Capital Data, 1990–2005". (The Northeast Regional Center for Rural Development, Penn State University, University Park, PA).

SALOUTOS, T. (1976), "The Immigrant Contribution to American Agriculture", *Agricultural History*, **50**, 45–67.

SOLOMOU, S. and WU, W. (1999), "Weather Effects on European Agricultural Output, 1850–1913", *European Review of Economic History*, **3**, 351–373.

SPITZER, Y. and ZIMRAN, A. (2013), "Migrant self-selection: Anthropometric evidence from the mass migration of Italians to the United States, 1907–1925". *Journal of Development Economics*, **134**, 226–247.

TABELLINI, M. (2018), "Gifts of the Immigrants, Woes of the Natives: Lessons from the Age of Mass Migration" (Mimeo, M.I.T.)

WEBER, T. (1952), *The Northern Railroads in the Civil War, 1861–1865* (Bloomington, IN: Indiana University Press).

WEGGE, S. A. (2002), "Occupational Self-Selection of European Emigrants: Evidence from Nineteenth-Century Hesse-Cassel", *European Review of Economic History*, **6**, 365–394.

WILLCOX, W. F. (1929), *International Migrations, Volume I: Statistics* (Cambridge: National Bureau of Economic Research).

WITTKE, C. (1939), *We Who Built America* (Ann Arbor: The Press of Western Reserve University).

WRIGHT, G. (1990), "The Origins of American Industrial Success: 1879–1940", *American Economic Review*, **80**, 651–678.